Reliable and Energy-Efficient Two Levels Unequal Clustering Mechanism for Wireless Sensor Networks

Alsnousi E Ali¹, Ahmed Y. Al-Dubai², Imed Romdhani³ and Mohamed A. Eshaftri⁴
School of Computing, Edinburgh Napier University, 10 Colinton Road, Edinburgh EH10 5DT, UK.
Email: A.Essa; A. Al-Dubai; I. Romdhani; M. Eshaftri@napier.ac.uk

Abstract—In Wireless Sensor Networks, clustering sensor nodes into disjoint groups is widely used to achieve load balance and increase network lifetime. In particular, traditional unequal clustering approaches where small clusters located close to the base station suffer from rapid energy depletion compared to others. To overcome the traffic load problem of near base station clusters, we propose a new two-level unequal cluster lightweight approach based on a threshold defined by the base station and using nodes’ residual energy and the distance from base station. Moreover, we developed a new re-clustering technique where cluster heads rotate locally in each cluster on a per level basis. Simulation results show that the proposed mechanism outperforms its counterpart algorithms.

Index Terms— WSNs; Distributed clustering; Routing; Network lifetime.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) is a key technology for the modern Internet of Things (IoT) applications. These networks play a key role in proliferating new IoT applications due to their high quality features in gathering information and flexibility. As rapid evolution shifts toward creating new IoT—WSN applications, many issues need to be addressed to fully integrate IoT with WSNs. Some of these challenges are mentioned in [1] [2]. On the other hand, significant research emphasis has been given to addressing sensor nodes scarce resources. WSNs consist of numerous tiny self-organised sensor nodes connected by wireless links. These nodes are capable of perceiving different physical world data. Sensor nodes are equipped with limited battery, sensing unit, processor and memory.

In some WSN applications [2], the nodes are unreachable when deployed, thus, it is impossible to replace or recharge their batteries. In relation to this limitation, the researchers were motivated to propose different approaches to extend network lifetime. Gathering data and transmitting it to the Base Station BS is often the most expensive paradigm in depleting node energy. Therefore, strict routing protocols must be applied to handle routing tasks in order to be energy efficient.

The scarce resources of randomly deployed sensor nodes make direct communication with BS or sending data in a multi-hop manner unfeasible for large-scale wireless sensor networks. The aggregation and elimination of redundant data is used in WSNs routing protocols to enhance the network lifetime [3]. Clustering is an effective mechanism by which to enhance network lifetime by forming manageable disjoint groups [4]. The nodes in these groups are either cluster head (CH) or cluster member (CM). Hence, each node assigned to be CH will perform local data aggregation from its CM to be routed to the BS. The CHs are selected based on some parameters to perform data gathering, compression and data transmission. Consequently, the energy of CH drains much earlier than that of CM. Managing the load on CH to balance the energy consumption is therefore necessary to maintain the balance.

There are two methods by which to construct clusters in WSNs: centralised or distributed [5]. In the centralised formation, the selection of CHs is handled by the BS or predetermined by the application. However, in distributed formation, the nodes independently select their CH based on the applied algorithm criteria. In this paper, we consider the distributed clustering technique. Several clustering algorithms have been proposed with different objectives, such as fault-tolerance, load-balance, increasing network connectivity and reducing data redundancy etc [6]. This paper introduces the new Reliable and Energy-efficient Clustering Mechanism (REUCS) for WSNs. The paper is organised as follows. Section I describes the motivation and contribution of the study. Section II outlines related works. Section V provides the preliminaries. Section V presents our proposed algorithm. Section VI presents performance evaluation results, and future work and recommendations for further improvement of the proposed approach are discussed in Section VII.

II. MOTIVATION AND CONTRIBUTION

The fundamental motivation for this study is to enhance network lifetime, achieving scalability and reliable node communication, whilst considering the nodes resources and the nature of random deployment. The construction of clusters in randomly deployed nodes can be challenging, especially when applying a distributed algorithm [5]. The load on selected CHs in each round is burdened by managing intra/inter-cluster communication. Therefore, CHs deplete their energy more than CMs and they cannot send their data until a new clustering process starts [5]. This problem is worse when applying periodic selection of CHs, as in [7], where fewer energy nodes can be selected to be CH. Other algorithms use node residual energy as the prime criteria for nodes to become cluster head; this approach is effective for mitigating the issue of selecting fewer energy nodes as CH in certain applications. In large-scale clustered networks, it
is efficient to employ multi-hop communications between CHs to transmit gathered data to BS and overcome signal propagation problems [8]. Nevertheless, the issue near the BS clusters persists due to incoming traffic load from far clusters. This issue will result in the near BS Cluster Heads potentially deplete their energy much earlier [9]. The loss of these critical CHs will significantly affect the performance of the network and could isolate the BS.

**First Motivation:** Cluster formation process issue. The aforementioned facts have an influence on the process of selecting suitable nodes to become CH in the set-up phases. Several criteria have been studied in the literature by which to select optimal nodes to become CHs. However, most of these algorithms apply periodic selection or are based on various parameters, where messages exchanged to gather other nodes information affects the network lifetime.

**First Contribution:** we reviewed different techniques proposed in the literature for selecting CHs among randomly deployed nodes; we concluded our review by developing a novel lightweight selection of CHs in the first round and reducing the set-up phases to construct clusters, as shown fig 1(A) for our contribution compared with 1(A) for traditional clustering algorithms. This was achieved by eliminating neighbour discovery process and gathering other nodes information affects the network lifetime.

**Second Motivation:** Re-clustering process. The vital objective of clustering is to maximise network lifetime, stabilising network topology, data aggregation and scalability. The conflict between the energy consumption and frequent re-clustering of the entire network results in excessive energy waste; this impact is mainly caused by control messages exchanged to re-cluster the network in each round.

**Second Contribution:** Second Contribution: levels re-clustering scheme proposed to stabilise network topology and conserve network energy. Frequent re-clustering or other re-clustering schemes explained in Section III are inefficient when compared with REUCS. The re-clustering of the entire network in each round, or when one of the clusters depletes their energy is not affective because of the extensive overhead and re-clustering process. In REUCS, we applied local CH rotation inside the cluster based on energy threshold, and when the entire cluster reaches the defined threshold, the BS will trigger a new clustering process only for that level.

III. RELATED WORKS

**Low Energy Adaptive Clustering Hierarchy (LEACH)** [7] was one of the first clustering algorithms introduced for WSNs. The clusters are constructed by periodically selecting CHs among network nodes and the remaining nodes join the nearest CHs based on RSSI. This procedure rotates in each round to balance the energy consumption in the network. However, the communication between CHs and BS is direct, thus, the energy cost for clusters that are far away from BS is significant. The probabilistic approach used in LEACH for CH selection, which results in selecting suboptimal nodes to perform CH roles, leads to poor network performance. Several modified versions of LEACH protocol have been proposed in the literature: LEACH-C [10], LEACH-B [11], EE-LEACH [12].

**A Hybrid Energy-Efficient, Distributed protocol (HEED)** , can be used to overcome some of LEACHs protocol drawbacks; HEED proposes multi-hop communications between CHs in the network. The selection of CHs is based on the remaining energy of the node, so the node with a high energy level is more likely to become CH. However, the construction of clusters in HEED shows that clusters close to the BS suffer from the incoming load from other clusters. This process of selecting the CHs creates a significant overhead.

**An Energy-Efficient Unequal Clustering Mechanism for Wireless Sensor Networks (EEUC)** [9] proposes an unequal cluster size algorithm; the main idea here is that the clusters close to BS should always be small in size to allow the CHs in these clusters to potentially preserve their energy and handle the incoming traffic from other clusters. The EEUC selects CHs mainly based on the nodes residual energy. However, the energy in CHs near the BS is greatly affected by the imposed data traffic and significant overhead in set-up phase.

**In A New Weight based Rotating Clustering Scheme for WSNs (WEERC)** [13], the protocol is weight-based CH selection protocol. The selection of CHs is based on two steps: first, the CH candidates are selected from the nodes based on three metrics the transmission range, degree and remaining energy; second, the candidates will compete to become the final CH using centrality of the node among its neighbours and the transmission diameter if there are any other candidates in the same range. The WEERC uses rotation inside the cluster until all the nodes drain their energy and then the BS will trigger the network re-clustering process.

**Weight Driven Cluster Head Rotation for Wireless Sensor Networks (WDCRR)** [14] is a weight-based protocol based on selecting the CH with the highest weight node. This weight is a combination of three metrics: the residual energy of the node, the cluster head frequency and the node distance.
to the BS. The rotation in WDCRR is based on node weight if the current CH falls below the threshold, the next highest weight member will become CH.

**An Efficient Dynamic Load Balancing Aware Protocol (DLCP)** [15] selects the nodes with the highest residual energy as CHs in the network, which is similar to the selection process in the HEED protocol. The clusters are formed in unequal size, as in EEUC. The CHs rotate inside clusters applied in DLCP and when the energy of CHs fall below a fixed threshold, the BS will re-cluster the entire network.

**Rotated Hybrid, Energy-Efficient and Distributed Clustering Protocol in WSN (R-HEED)** [16], protocol is an enhancement of the HEED protocol, adding rotation inside the cluster. Each selected CH will sort its member and broadcast a schedule table to its members, alerting every node of its turn to be a CH. The CH waits for a fixed time in order to check whether there has been any message from the BS to re-cluster the entire network, otherwise the rotation continues.

**IV. PRELIMINARIES**

**A. Network structure**

In this paper, the clustering structure has the following assumptions: \( N \) heterogeneous sensor nodes randomly deployed in area of dimension \( M \times M \) the BS is located outside sensing area. \( n_i \) nodes are selected as CHs where \( n_i \in N \) then \( C \) clusters are constructed and \( n_i \in C \), the \( C \) are defined as \( C = (C_1, C_2, \ldots, C_k) \), the \( n \) route its data packet through associated CH. The CH aggregate and process member data before forwarding it as single packet to next hop CH or to the BS in case its within defined first level if multi-hop condition not satisfied for first level multi-hop. The following are system assumption for the nodes and BS:

- The nodes have limited power source and the base is not energy constrained
- The deployed nodes have different initial energy. However, the processing, computational and communication capabilities are similar
- The nodes are stationary throughout network lifetime and deployed randomly in area of interest to form WSN.
- Nodes are location unaware
- Links are symmetric between nodes.
- BS is based outside the sensing area.
- The nodes can compute approximate distance to another node by Received Signal Strength Indication (RSSI)

**B. Energy models**

The energy model considered in this paper for energy consumption and lifetime is based on [9] model. It uses both free space \((d^2 \text{ power loss})\) and the multi path fading \((d^4 \text{ power loss})\) for the channel models. The energy consumption is decided depending on the distance between the transmitter and the receiver. The equation 1 illustrates The energy spent to transmit \( l \)-bit packet over distance \( d \).

\[
E_{Tx}(l,d) = \begin{cases} 
    lE_{elec} + l\varepsilon_f d^2, & d < d_o \\
    lE_{elec} + l\varepsilon_m d^4, & d \geq d_o 
\end{cases} 
\]

The radio expends energy when receiving that message as

\[
E_{Rx}(l) = lE_{elec}
\]

**V. PROPOSED ALGORITHM**

Distributing energy consumption among the deployed nodes is a key element in sustaining network lifetime regardless of the computational process inside the individual node, which is unavoidable in most WSN applications. In random deployment, we cannot guarantee equal distribution of nodes, where the nodes could be in different densities. In this section, we describe the REUCS mechanism to construct unequal clusters using an affective lightweight approach to select CHs.

**A. REUCS Basic Idea**

The proposed algorithm operates on rounds basis, as in most WSN clustering algorithms. Once the nodes are deployed randomly, the BS broadcast a single packet to the nodes. Upon receiving BS packet, nodes will determine approximate distance from BS based on the received RSSI. The clustering process comprises of two phases: the set-up phase and the steady phase. In the set-up phase, number of nodes selected to be CHs and the non-CH nodes join the closest CH based on the strongest received signal. The steady phase inter/intra data transmission in the network then takes place. In contrast to counterpart algorithms reviewed in Section III, the REUCS initial set-up phase, neighbouring node information is not required in order to select CH. The decision to select nodes as CHs is based on effective utilisation of node local information:residual energy and the distance from base station. REUCS forms two levels of network clusters based on the threshold identified by the BS.

**B. Base Station operation**

In REUCS we assume that the BS knows the size of the deployment field and can adjust its transmission levels into two levels maximum and minimum. Fig 2 illustrates the two transmission levels of BS the \( T_{r_{max}} \) in yellow lines and the \( T_{r_{min}} \) in blue squares. Using this assumption, the BS can set a regional threshold for the nodes to identify in which level they belong to. In our strategy, the threshold value of the network levels can be calculated by equation 3.

\[
\text{threshold} = \frac{\text{area}_{max} - T_{r_{min}}}{T_{r_{max}} - T_{r_{min}}}
\]

where \( \text{area}_{max} \) is the edge of the network as shown in fig 2. The BS will broadcast one single control packet containing BS ID, Level threshold, \( T_{r_{max}} \), \( T_{r_{min}} \) and other control data.
C. Cluster formation

The size of cluster in the network has a significant impact on network lifetime $L$. Thus, Our goal is to form unequal clusters where clusters near $BS$ level smaller than those in second level. The idea of unequal cluster was derived from [9], with some modification to implement our strategy of unequal two levels clusters. In order to achieve our goal, the nodes firstly determine in which level they belong when they receive $BS$ control packet. Upon receiving $BS$ control packet which contain, $BS$ ID, Level threshold, $TR_{max}$ and $TR_{min}$, the nodes can determine approximate distance from $BS$ by RSSI. Subsequently, all the nodes need to determine their level $n_{level}$ by employing equation 4.

$$n_{level} = \frac{d(n_i, BS) - NTR_{max}}{NTR_{max} - TR_{min}}$$  \hspace{1cm} (4)$$

where $d(n, BS)$ is the distance from $BS$ and $TR_{min}$ is $BS$ minimum transmission range, $NTR_{max}$ is the node maximum transmission which can be obtained by equation 5.

$$NTR_{max} = d(n_i, BS) - TR_{min}$$  \hspace{1cm} (5)$$

Fig 3, demonstrates an example of $NTR_{max}$ and $TR_{min}$. BS will be influenced by the largest energy level to be the highest $f_{n, max}$ value, therefore will have more chance to become $CH$, whereas the nodes far away from the $BS$ the distance will influence the value of $f_{n, max}$. To ensure the selection of optimal nodes for the role of $CH$. Thus, $f_{max} = n_{i1}, n_{i2}, \cdots, i_N$ become potential $CH$ in their level. The $E_i$ calculated as in equation 6

$$E_i = (1 - \frac{E_{res}}{E_{max}})X.$$  \hspace{1cm} (6)$$

where $E_{res}$ is the residual energy of $i$ and $E_{max}$ is the maximum energy of $i$, $X$ is random number between (0 and 1) this to ensure ($E_i \neq E_j$). The $d(n_i, BS)$ is the distance from $n_i$ to $BS$. The $d(n_i, BS)$ measured by the power of received signal strength $PRa$ which was represented by [17] and illustrated in equation 7.

$$PRa = \frac{PRa - \lambda}{4\pi d}$$  \hspace{1cm} (7)$$

where $PRa$ and $d$ is the distance from the sender. the $RSSI$ is the ration of power received to $PRa$.

$$RSSI = 10 \log(\frac{PRa}{PRa})$$  \hspace{1cm} (8)$$

The final stage in constructing cluster, the nodes in each level compete to become final $CH$ according to the equation 9.

$$nR_{adu} = 1 - g \cdot \frac{NTR_{max} - d(n_i, BS)}{NTR_{max} - TR_{min}}Rd$$  \hspace{1cm} (9)$$

where $NTR_{max}$ is the node maximum transmission, and $TR_{min}$ is the $BS$ minimum transmission range. $g$ is the constant number from (0,1). Unlike EEUC algorithm were the $d_{max}$ and $d_{min}$ are pre-defined, in REUCS the $NTR_{max}$ and $TR_{min}$ can be obtained during network configuration.

Once $n_{level}$ obtained the deployed nodes broadcast cost function $f_{n, max}$ for CHs selection. $f_{n, max}$ is a combination of node residual energy $E_i$ and the distance from the BS $d(n_i, BS)$. Each node will broadcast the value of $f_{a}(E_i) + d(n_i, BS)$ which can be received only by the nodes within same $n_{level}$. We need to ensure in first round the CHs near the $BS$ are selected as the highest energy level to sustain with incoming traffic whereas those far away will depend on the distance from $BS$ and the competition radius. According to the value obtained by $f(E_i + d(n_i, BS))$ the nodes near $BS$ will be influenced by the largest energy level to be the highest $f_{n, max}$ value, therefore will have more chance to become $CH$, whereas the nodes far away from the $BS$ the distance will influence the value of $f_{n, max}$. To ensure the selection of optimal nodes for the role of $CH$. Thus, $f_{max} = n_{i1}, n_{i2}, \cdots, i_N$ become potential $CH$ in their level. The $E_i$ calculated as in equation 6

$$E_i = (1 - \frac{E_{res}}{E_{max}})X.$$  \hspace{1cm} (6)$$

where $E_{res}$ is the residual energy of $i$ and $E_{max}$ is the maximum energy of $i$, $X$ is random number between (0 and 1) this to ensure ($E_i \neq E_j$). The $d(n_i, BS)$ is the distance from $n_i$ to $BS$. The $d(n_i, BS)$ measured by the power of received signal strength $PRa$ which was represented by [17] and illustrated in equation 7.

$$PRa = \frac{PRa - \lambda}{4\pi d}$$  \hspace{1cm} (7)$$

where $PRa$ and $d$ is the distance from the sender. the $RSSI$ is the ration of power received to $PRa$.

$$RSSI = 10 \log(\frac{PRa}{PRa})$$  \hspace{1cm} (8)$$

The final stage in constructing cluster, the nodes in each level compete to become final $CH$ according to the equation 9.

$$nR_{adu} = 1 - g \cdot \frac{NTR_{max} - d(n_i, BS)}{NTR_{max} - TR_{min}}Rd$$  \hspace{1cm} (9)$$

where $NTR_{max}$ is the node maximum transmission, and $TR_{min}$ is the $BS$ minimum transmission range. $g$ is the constant number from (0,1). Unlike EEUC algorithm were the $d_{max}$ and $d_{min}$ are pre-defined, in REUCS the $NTR_{max}$ and $TR_{min}$ can be obtained during network configuration.
Algorithm 1  Cluster Formation algorithm

1: procedure BS broadcast
2: for each \( n \in N \) do
3: \( n \leftarrow d(n_i, BS) \) Computed in equation (5)
4: \( n \leftarrow (n_{level}) \) Determine (equation (4)
5: \( n \leftarrow f_{n\text{max}} \)
6: end for
7: if \( n_{level} < \text{Threshold} \) then
8: \( n_i \leftarrow \text{Level1} \)
9: costFunctionMsg (\( f_{n\text{max}} \))
10: else
11: \( n_j \leftarrow \text{Level2} \)
12: costFunctionMsg (\( f_{n\text{max}} \))
13: end if
14: On receiving costFunctionMsg
15: if \( n_{level} == n_{level} \) then
16: Add \( n_j \) to \( n_i\text{.compteteList} \)
17: FinalHeadMsg (ID; \( nRdu \))
18: end if
19: end procedure

D. Re-Clustering process

To avoid traditional frequent re-clustering of the whole network in fixed interval, which negatively affects network lifetime. Therefore, we developed a new scheme that allows nodes to rotate CH role locally inside cluster until all the nodes perform CH role. The local re-clustering is triggered only when current CH energy fall below energy threshold. Once all the cluster nodes performed CH role, Then the last CH sends re-clustering message and the cluster average energy to the BS in order to trigger new clustering process. Upon receiving re-clustering message the BS checks in which level re-clustering message coming from and the average energy level in that cluster. If the average energy level in cluster not below threshold, then BS trigger new re-clustering process only for that level otherwise will trigger new network clustering. The following pseudocode summaries Local and levels CH rotation algorithm.

Algorithm 2  Local and levels CH rotation algorithm

1: procedure CH Broadcast rotate
2: if \( n \in C_{CH} \) then
3: if \( E_{i\text{res}} > E_{\text{ClusterMember}} \) then
4: \( n_i = \text{newCH} \)
5: end if
6: end if
7: if \( E_{i\text{res}} < E_{\text{ClusterThreshold}} \) then
8: BS broadcast level; re-cluster
9: end if
10: end procedure

VI. PERFORMANCE EVALUATION

Simulation experiments have been conducted to evaluate the performance of the proposed REUCS using a Castalia simulator. We compared the network lifetime to the main performance parameter. The network lifetime is often addressed by three criteria: 1) When First Node Dies (FND); 2) When Half nodes Die (HND); 3) When last node dies (LND); In order to validate our simulation results we compared REUCS with well-known algorithms EEUC protocol, as well as state of art algorithm such as WDCR [14] and DLCP [15]. Table I shows the parameters and values considered for simulation experiments. Fig. 5 shows the rounds until the first node dies (FND) and the number of nodes in the network, which ranges from 100 to 400 nodes. From this figure, it can be noted that FND in EEUC outperforms all algorithms including our REUCS in all sets of nodes. This is because EEUC nodes not all involved in CH election process, whilst, re-clustering in EEUC is based on time interval where energy load is distributed between the nodes in only first rounds. However, This process negatively impacts the total network lifetime.

![Fig. 5. First Node Dies (FND)](image)

Table I shows the parameters and values considered for simulation experiments.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>250 x 250</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>100 - 400</td>
</tr>
<tr>
<td>Base Station Location</td>
<td>275 x 125</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>25J</td>
</tr>
<tr>
<td>Data Packet Size</td>
<td>200 bit</td>
</tr>
<tr>
<td>Deployment</td>
<td>Random</td>
</tr>
<tr>
<td>Radio Model</td>
<td>CC2420</td>
</tr>
</tbody>
</table>

![Fig. 6. HND for the four protocols in different network size](image)
In fig 7 shows the round until the last node dies (LND) in the four protocols tested by different network size ranging from 100 to 400 nodes. According to the results obtained, the REUCS performs better than the other protocols in all network sizes. In REUCS the 100, 200 and 400 nodes network has the highest performance in LND at 850, 899 and 850 rounds respectively, whereas in 300 node networks it reaches around 830 rounds. In comparison with the other three protocols in their best performance network size, REUCS is more efficient by 4.8% than DLCP in 300 network nodes. And more efficient by 14.5% than WDCR in 200 nodes and 38% than the EEUC in 300 nodes respectively.

VII. CONCLUSION

In this study, we proposed a novel clustering mechanism in order to increase network lifetime. We showed that dividing the network into two levels of clusters can reduce energy consumption and balance the load among nodes. In REUCS the selection of cluster heads is based on two local metrics: node residual energy and distance to the base station. Using only these two metrics helped to reduce overheads in the set-up phase. We also eliminated the frequent re-clustering process in traditional clustering algorithms, which has a significant impact on energy consumption. The simulation results show that the REUCS algorithm outperforms its counterparts in terms of network lifetime and scalability.

REFERENCES