A comparison of HEED based clustering algorithms - introducing ER-HEED

Zaib Ullah, Leonardo Mostarda, Roberto Gagliardi, Diletta Cacciagrano, Flavio Corradini
Computer Science Division, University of Camerino, Italy
Email: {zaibullah.zaibullah,leonardo.mostarda,roberto.gagliardi,diletta.cacciagrano,flavio.corradini}@unicam.it

Abstract—A Wireless Sensor Network (WSN) is composed of distributed sensors with limited processing capabilities and energy restrictions. These unique attributes pose new challenges amongst which prolonging the WSN lifetime is one of the most important. Clustering is an energy efficient routing technique that has been widely applied to report data from the WSN nodes to a centralised Base Station. A plethora of different clustering protocols have been proposed. Some protocols are based on equal-sized clusters while others use clusters of unequal size. Some others make use of rotation techniques to reduce the amount of cluster head elections. When different clustering approaches are presented different simulation settings are used. In this paper we perform a comparison study of HEED based clustering protocols that are HEED, UHEED, RUHEED and a novel variation of R-HEED that is ER-HEED. We have considered the same network model, the same energy consumption model and we have compared the lifetime of the protocols by considering various case studies. Our comparison study shows that the selection of the protocol to be used depends on the case study and the WSN lifetime measure that is considered.

Index Terms—Wireless sensor networks; Clustering protocols; Energy efficiency;

I. INTRODUCTION

A wireless sensor network (WSN) consists of hundreds or thousands of sensor nodes that can sense the environment and report data to a base station (BS) [1], [2], [3], [4]. Since these sensor nodes are equipped with batteries and could be deployed in an unattended environment, the energy efficiency and enhancement of lifetime are important issues. Sensor node deployment in an environment can be deterministic, random, stationary or mobile, based on the nature of the application [12]. Nodes sense the environment either in a continuous or event based fashion and transmit information to a centralised BS. This is a controller that processes the collected data from the nodes and makes decisions accordingly [10], [11].

Various approaches have been proposed to enhance the network lifetime and efficiency. Cluster based design of WSNs is a widely used technique to preserve sensor node energy and enhance network lifetime. Each cluster consists of a set of sensor nodes and has a representative node that is known as cluster head (CH). A CH aggregates data from its member nodes (i.e., intra-cluster communication) and can cooperate with other CHs to report the data to the BS (i.e., inter-cluster communication). To date hundreds of clustering protocols have been proposed and it is very difficult to determine the best protocol to apply when some specific applications and needs are encountered. For instance different applications can require different lifetime goals, i.e., the optimisation of the first node dies or last node dies. Different systems can have different features, i.e., number of nodes, size of the deployment area.

In this paper we perform an equal comparison of different clustering approaches in order to allow an informed choice for the selection of the right clustering approach. While we consider different types of cluster enhancements (i.e., equal, unequal and cluster head rotation), we decided to focus on a widely cited and implemented cluster election algorithm that is HEED. More precisely we consider HEED, UHEED [15], RUHEED [5] and a novel variation of R-HEED that is ER-HEED. We have considered the same network model, the same energy consumption model and we have compared the lifetime considering various case studies. Different measures of WSN lifetime have been considered such as first node dies and half node dies. Our comparison study shows that the selection of the protocol to be used depends on the case study and the WSN lifetime measure that is considered. The main contribution of this paper is listed in the following: (i) a fair comparison of various HEED based protocols to make an informed choice about the clustering to be selected; (ii) a novel clustering protocol ER-HEED that is a variation of R-HEED.

The rest of the paper is organised as follows: section II describes HEED, UHEED, RUHEED and ER-HEED; Section III discusses the network model, the energy model and the simulation results; Section IV compares the presented work with existing related works in this area. Finally, Section VI provides a conclusion and outlines future work.

II. HEED BASED CLUSTERING ALGORITHMS

In this section we explain HEED, UHEED, RUHEED and ER-HEED.

A. HEED

Hybrid, Energy-Efficient, Distributed (HEED) [17] is an equal-sized clustering protocol that produces clusters of equal size. CHs selection in HEED is based on the residual energy of sensor nodes and one of the following parameters: node degree or distance of neighbouring nodes to the CHs. Cluster formation in HEED is performed by means of three phases: initialisation; iteration and finalisation.

The initialisation phase assigns to each node the probability of becoming a tentative cluster head. This is done according to the following formula (see [17] for details about the formula):
\[ CH_{\text{prob}} = C_{\text{prob}} \times \frac{E_{\text{residual}}}{E_{\text{max}}} \]  

(1)

where \( C_{\text{prob}} \) is the initial probability (i.e., a predefined value), \( E_{\text{residual}} \) is the residual energy and \( E_{\text{max}} \) is the maximum energy of the sensor nodes.

In the iterative phase of HEED, some nodes will become tentative cluster head. When a node is in the communication range of some tentative cluster heads it will choose the one with the least cost. When a node is not in the communication range of a tentative cluster head it will eventually become cluster head. In finalisation phase, sensor nodes that did not select any cluster head will become one.

HEED generates equal-sized balanced clusters. The size of a cluster is independent of its distance from the BS. As reported in [17], due to the inter-cluster communication (i.e., relay traffic amongst CHs), sensor nodes near to the BS drain their energy quickly. In fact while they relay the normal intra-cluster traffic (i.e., the traffic inside a cluster) they are also burdened with heavier relay traffic from the rest of the network nodes. This relay traffic not only affect the network lifetime but also leads to network partitions (close to the base station). This is also known as the hot spot problem. Unequal clustering has been devised to counter this issue.

B. UHEED

UHEED [15] is an unequal clustering protocol that combines HEED [17] and EEUC [23]. More precisely, it uses the leader election algorithm that is defined in HEED and each cluster head computes the competition radius by using the formula defined by EEUC [23]. This is described as follows (see [23] for a full description):

\[ R_{\text{comp}} = (1 - c \left( \frac{d_{\text{max}} - d(s_i, BS)}{d_{\text{max}} - d_{\text{min}}} \right)) R_{\text{comp}} \]  

(2)

where \( R_{\text{comp}} \) is a predefined competition radius that is the maximum transmission radius of a sensor node. The constant \( c \) is a coefficient having values between 0 and 1, \( d_{\text{max}} \) and \( d_{\text{min}} \) are the maximum and minimum distances of sensor nodes from the BS, \( d(s_i, BS) \) is the distance of \( i \)-th sensor node \( s_i \) from the BS.

UHEED creates unequal-sized clusters whose size depends on the distance of the CH from the BS. The farther away a cluster head is from the BS, the larger will be its competition radius. In other words, clusters farther away from the BS will have a larger radius when compared to clusters nearer to the BS. By creating unequal sized clusters, the amount of inter-cluster traffic is considerably reduced for the CHs nearer to the BS. Thus energy hole problem discussed for HEED is mitigated. Although UHEED improves the HEED protocol, enhancements can still be made. Energy can be saved with the introduction of cluster head rotation. More precisely, the cluster election phase is sometime replaced by a designation phase where the old cluster head directly elects a new cluster head. This can reduce the number of cluster election phases thus reducing the number of control messages.

C. RUHEED

RUHEED [5] improves UHEED with the introduction of rotation [24]. RUHEED is composed of three phases that are cluster head election, cluster formation and rotation. In the cluster head election phase RUHEED algorithm is used. In the cluster formation phase the competition radius formula of EEUC [19] is used. In the rotation phase the current CH designates one of its cluster members as the new CH without performing any election protocol. The new cluster head is selected based on the highest residual energy. The rotation phase is performed until one of the WSN nodes completely depletes its energy. When this happens the BS will inform all nodes to perform a new cluster head election and cluster formation phase. The advantage of RUHEED is to reduce the number of cluster head election and cluster formation phases thus reducing the number of control messages.

D. A novel protocol - Energy based Rotated HEED (ER-HEED)

ER-HEED improves HEED with the introduction of rotation in equal-sized clusters. ER-HEED is composed of the following phases:

1) cluster head election and cluster formation: cluster head election and cluster formation are performed according to the HEED protocol;
2) rotation phase: the current CH designates the cluster member with the highest energy as the new CH without the need of performing any election protocol;
3) When any of the node dies re-clustering is performed by repeating step 1.

Rotation of CH role in equal-sized clusters reduces the number of HEED clustering election phases. It is assumed that each data packet received by the CH contains energy information of its member nodes. This is used for the designation phase.

III. NETWORK MODEL, ENERGY MODEL AND SIMULATION RESULTS

In this section we present the network model, energy model and simulation results. We make sure that all the clustering protocols are compared under the same conditions and by using the same assumptions.

A. Network Model

Sensor nodes are uniformly deployed in a two dimensional field with the following assumptions:

- All sensor nodes have the same initial energy (homogeneous), same data communication and processing capabilities;
- Each sensor node is identified with a unique ID;
- Sensor nodes can transmit at various power levels depending on the distance of the receivers;
- Sensor nodes are not mobile that is they remain stationary;
- Sensor nodes are equally distributed in the field and transmit at the same rate.
The BS is located away from the sensing field with no energy constraints. It is considered to be a node with enhanced communication and computation capabilities. The BS is stationary. All data that are collected by the WSN is highly correlated, therefore it can be aggregated. More precisely, we assume that each CH (after receiving one message from each cluster member) aggregates the intra-traffic into a single message. Inter-traffic is not aggregated that is a CH forwards (towards the BS) messages sent from other CHs with no aggregation 1.

We use a quite common network operation model that is presented in [17], [15], [5], [23]. A clustering protocol usually includes the following phases: (i) cluster election and formation; (ii) data exchange phase; (iii) rotation (if any); (iv) re-election and formation. During the data exchange phase a TDMA is composed of the following two activities: (i) each member node sends one message to its cluster head; (ii) all CHs’ data reaches the BS. In other words, a TDMA starts from the collection of data from the member nodes and ends when all the data reaches the base station. A round is composed of multiple TDMAs. In most of the simulations we consider that a round is composed of 5 TDMAs.

The multi-hop transmission is treated in a very fair way. When the simulation starts the same competition radius has been defined for all the nodes. We tried different simulation with a different initial competition radius that are 40 meters, 55 meters, 70 meters, 75 meters and 85 meters (see Section III-C for details). When nodes close to the BS start dying (or cannot reach the BS) we can end up in a situation where no nodes can relay data to the BS. In this case the closest nodes to the BS increase their power in order to reach it.

### B. Energy Model

The radio model employed uses both the free space and the multi-path channel model and assumes error-free communication links. The simulation parameters are those used in [17] and summarised in Table 1. The network field (in the table referred to as network grid) is 100 metres by 100 metres. The position of the base station is (50,175). A sensor node spends $E_{elec} = 50nJ/bit$ to run the transceiver circuitry. The energy spent by the amplifier $E_a$ will depend on the distance $d$ between the sender and the receiver. More precisely $E_a = E_{fs} = 10pJ/bit/m^2$ when $d < d_0 = 75m$ (in this case a free space model is assumed) while $E_a = E_{mp} = 0.0013pJ/bit/m^4$ when $d \geq d_0 = 75m$ (in this case a multipath model is assumed). The transmission energy $E_{Tx}$ spent to send a packet can be calculated by using the following formula (see [16] for details):

$$E_{Tx} = (E_{elec} \times k) + (E_a \times k \times d^n),$$  

(3)

where $k$ specifies the number of bits that are sent, $d$ is the distance of the receiver and $n = 2$ for the free space model and $n = 4$ for the multipath model.

The amount of energy $E_{Rx}$ spent to receive a $k$-bit size message can be calculated as follows (see [16] for details):

$$E_{Rx} = (E_{elec} \times k)$$  

(4)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network grid</td>
<td>From (0,0) to (100,100)</td>
</tr>
<tr>
<td>BS/Sink</td>
<td>(50,175)</td>
</tr>
<tr>
<td>$E_{elec}$</td>
<td>50nJ/bit</td>
</tr>
<tr>
<td>$E_{fs}$</td>
<td>10pJ/bit/m^2</td>
</tr>
<tr>
<td>$E_{mp}$</td>
<td>0.0013pJ/bit/m^4</td>
</tr>
<tr>
<td>Threshold Distance (do)</td>
<td>75m</td>
</tr>
<tr>
<td>$R^0_{comp}$ (competition radius)</td>
<td>40m, 55m, 70m, 85</td>
</tr>
<tr>
<td>Data Packet Size</td>
<td>2000bits</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>2J</td>
</tr>
<tr>
<td>Control parameter</td>
<td>$c = 0.1, 0.9$</td>
</tr>
<tr>
<td>Number of sensor nodes</td>
<td>100, 200, 300, 400</td>
</tr>
</tbody>
</table>

### C. Simulation Study and Results

We simulated the LEACH, HEED, UHEED, ER-HEED and RUHEED clustering algorithms by using a grid with dimension of 100 by 100 meters and by considering nodes that are uniformly deployed and have continuous data collection capacity (i.e., a constant rate). Each experiment result is obtained by averaging over 50 runs. We have compared the clustering algorithms by varying the following parameters: (i) network density; (ii) node coverage and (iii) leader election frequency.

**Network density** This has been tested by using an increasing number of nodes that is 100, 200, 300 and 400 nodes. An increasing number of nodes results in an increasing density since the grid dimension is fixed.

**Node coverage** In the case of equal clustering (i.e., LEACH, HEED and ER-HEED) we have set an increasing value for the competition radius (in the following referred to as $R_0$) that is 40, 55, 70 and 85 meters. An increasing radius results in an increasing cluster size (i.e., less clusters inside the WSN). The same competition radius $R_0$ is also used to calculate the cluster size in the unequal clustering case (i.e., UHEED and RUHEED). This is done by setting the maximum competition radius $R_{comp}^o$ of formula (2) to the competition radius $R^0$. Higher values of $R_{comp}^o$ will increase the size of the clusters (although clusters near to the BS are always smaller than the farthest ones).

**Leader election frequency** Various leader election frequencies have been tested by varying the number of rounds (i.e., TDMA frames, see Section III-A for details) before leaders get re-elected. More precisely we have performed 4, 5, 6, 7 or 8 rounds of data collection before re-electing the leaders.

We have compared the clustering algorithms according to three different definitions of network lifetime that are first node dies and half of the nodes dies.
1) First node dies and packet arrival at the base station: Figure 3 shows a comparison of first node dies (FND) for ER-HEED, RUHEED, UHEED, HEED and LEACH. This comparison is performed for an increasing number of nodes and a competition radius $R_0$ equal to 40. In the case of 100 nodes LEACH and ER-HEED have a comparable performance but as the number of nodes increases ER-HEED outperforms all the algorithms. Figure 4 shows a first-node-die comparison for a WSN composed of 300 nodes. This comparison is performed with an increasing competition radius i.e., 40, 55 and 70 meters. ER-HEED is outperforming all the clustering algorithms but when the competition radius is higher ER-HEED and RUHEED have a comparable performance. We have developed some simplified mathematical models and some further simulations in order to understand the aforementioned results. We found out that RUHEED and ER-HEED will always outperform HEED and UHEED. The main reason lies on the cluster election strategy. HEED and UHEED will perform always more cluster head elections (1 each round) with respect to RUHEED and ER-HEED. Here the cluster election is replaced by rotation till first node dies thus reducing the overhead energy needed for cluster election. This confirmed by the Figure 5 which shows the energy consumption for cluster head election which is higher in HEED and UHEED.

The results related to RUHEED and ER-HEED are slightly more complicated to be explained. Although ER-HEED outperforms RUHEED when the competition radius $R_0$ increases RUHEED energy performance is quite close to ER-HEED. This can be explained by considering the different clustering strategies they use for small and high values of $R_0$:

Small competition radius $R_0$. In this case RUHEED has more clusters than ER-HEED (especially next to the base station where RUHEED clusters are smaller). Since RUHEED has more clusters it is probable that some cluster head close to the BS will get more inter-traffic (i.e., traffic from other cluster head nodes) with respect to ER-HEED case (where there are less clusters). This can be confirmed by looking at Figure 1. This shows that CHs of RUHEED that are close to the BS can get up to four inter-cluster messages (there are four layers L1, L2, L3 and L4) while CHs of ER-HEED gets three inter-cluster messages (there are three layers).

Large competition radius $R_0$. In this case the number of clusters in RUHEED and ER-HEED are comparable (see Section III-C the Node coverage paragraph for details) thus the two approaches have a similar performance.

We have further compared the protocols with a variable number of TDMA frames before the re-election (for HEED and UHEED) and rotation (for RUHEED and ER-HEED) are performed (see Figure 6). We can conclude that (when HEED election is considered) for first node dies an equal-sized clustering approach that is based on rotation is the most energy efficient.

2) Half of the nodes die (HND): Figure 8 shows the number of nodes alive as the number of rounds increases. This result is related to a network of 300 nodes, a competition radius equal to 40 and a TDMA frame equal to 5. We can clearly see that for HND lifetime measure the RUHEED clustering approach outperforms all the algorithms. The same conclusion is reached when an increasing number of nodes (see Figure 9) and an increasing competition radius (see Figure 10) is considered. We have performed some further simulations in order to understand the aforementioned results. More precisely we have observed the distribution of the nodes that have residual energy when the first node dies in ERHEED and RUHEED. We noticed that the amount of nodes still alive that are around the BS is more in RUHEED than ERHEED. This is shown in Figure 2 that is a snapshot of the WSN taken after the first node dies for ER-HEED. We can clearly see that while for ER-HEED all nodes that are close to the BS are dead (the red clusters of Figure 2 define areas of the grid where no nodes are alive) RUHEED has still some nodes close to the BS that are alive (the green clusters of Figure 2 define areas of the grid where there are nodes that are still alive).

It is worth mentioning that rotation over equal-sized clustering (i.e., ER-HEED) has better lifetime when compared with no-rotation over unequal-sized clustering (i.e., UHEED). In other words reducing cluster head election (i.e., rotation) is more effective than introducing unequal size. We can conclude that (in our simulation settings) for half of the node die lifetime measure unequal-sized clustering with rotation is the most energy efficient approach.

---

3 For the sake of presentation we do not introduce the mathematical models. The on-line technical report [27] contains details about the mathematical models.

3 The number of clusters is also comparable when a very small value of the parameter $C$ in the formula of RUHEED is set.
Fig. 2. Cluster left after first node dies for ER-HEED.

Fig. 3. First node dies for LEACH, HEED, UHEED, RUHEED and ER-HEED when competition radius $R_0=40$ and TDMA=5.

Fig. 4. First node dies for LEACH, HEED, UHEED, RUHEED and ER-HEED with variable cluster radius, 300 nodes and TDMA=5.

Fig. 5. Ratio of consumed clustering energy w.r.t total energy of WSN at various number of sensor nodes.

Fig. 6. Comparative analysis of WSN protocols at various TDMA frames with 300 nodes and a competition radius of 40.

Fig. 7. Data Packets of ER-HEED, RUHEED, UHEED and HEED reached to the BS at various number of sensor nodes.
3) Correctness of the simulation results: The simulation programs for HEED, RUHEED, UHEED and LEACH have been validated by using the numerical results presented in the existing literature. After the correctness of the algorithms was ensured we run the simulations presented in this paper.

IV. LITERATURE REVIEW

A lot of literature is available on equal and unequal size clustering techniques for wireless sensor networks. Low Energy Adaptive Clustering Hierarchy (LEACH) [16] is one of the primary adaptive hierarchical clustering algorithm. In LEACH, cluster head election happens probabilistically rather than based on residual energy of sensor nodes [15] and data transmission between CHs and BS takes place using single hop communication. Once a node has been elected as a cluster head it cannot take the same role in the next round. LEACH proposed randomized rotation of CH in the network and data compression at the cluster head. Hybrid Energy Efficient Distributed (HEED) [17] algorithm produces clusters of equal size and the residual energy of nodes and intra-cluster communication cost play a significant part in CH selection. Rotated Hybrid, Energy-Efficient and Distributed (R-HEED) [24] introduces rotation in HEED. This is based on predefined turn schedule before entering rotation. In ER-HEED, CH selection inside cluster during rotation phase is based on the residual energy information that is sent by the member’s nodes over the time. Distributed Weight-based Energy-efficient Hierarchical Clustering protocol (DWEHC) [13] proposes an optimization of the intra-cluster topology of the HEED protocol. This optimization produces more balanced, equal sized clusters and thus resulting in a better network lifetime. Like HEED, every node execute DWEHC protocol independently until and unless may not decide its status of either member node or CH. Manju Bala et al. (2012) proposed Deterministic HEED (D-HEED)[26] that primarily enhance the heterogeneous sensor nodes network lifetime and improves stability as well. D-HEED follows multi-level heterogeneous HEED (MH-HEED) protocol. D-HEED focuses on distributed clusters and optimal number of cluster head selection on the basis of probability that describes threshold T (Si) in heterogeneous sensors nodes network.

Extensive studies have been carried out on designing unequal sized clustering protocols. Energy Efficient Unequal Clustering (EEUC) protocol [19] is based on the idea that sensor nodes should join the unequal clusters and generate smaller unequal clusters near to the BS. Thus CHs nearer to the BS lives longer and avoid energy holes in the WSN.

In the unequal layered clustering approach (ULCA) [20],...
author partition the WSN into layers. The layers nearer to the BS are of smaller size when compared to the layers away from the BS. Cluster heads of middle layers retains more energy for inter-cluster data relay traffic [10]. ULCA performance is better than the EEUC in terms of network lifetime.

In [23], the authors describe an improved energy efficient unequal clustering (IEEU) protocol. This compute competition radius using node degree. Cluster heads that are closer to the BS have less number of member nodes in comparison to farthest one. Thus, CHs near to the BS have more residual energy for inter-cluster data transmission.

In [15], the authors introduce Unequal Clustering Algorithm (UHEED). This combine the unequal clustering size of EEUC and the leader election algorithm proposed in HEED. Smaller clusters are created near to the BS and larger clusters are created at the farthest distance. UHEED outperform LEACH, HEED and EEUC.

V. ACKNOWLEDGEMENTS

This work has been partially supported by the MIUR PRIN project CINA (2010LHT4KM)

VI. CONCLUSION AND FUTURE WORK

This paper provides the following two main contribution (i) a fair comparison of various HEED based protocols to make an informed choice about the clustering approach to be selected; (ii) a novel clustering protocol ER-HEED that is a variation of R-HEED. We compared LEACH, HEED, UHEED, RUHEED and ER-HEED by varying the network density, the node coverage and the leader election frequency. Our simulation results show that when first node dies is considered the equal-sized clustering approach that is based on rotation (i.e., our novel ER-HEED) is the most energy efficient. When half of the node die life time measure is considered unequal-sized clustering with rotation (RUHEED) is the most energy efficient approach.

As a future work we plan to define various mathematical models that can prove the aforementioned results. This should be used to validate the results for the HEED based protocols but also to generalise the results to a wider range of deterministic clustering based techniques.

REFERENCES