UHEED - AN UNEQUAL CLUSTERING ALGORITHM FOR WIRELESS SENSOR NETWORKS

E.Ever,1 R. Luchmun,1 L. Mostarda,1 A. Navarra2 and P. Shah1

1 CCM Department, School of Engineering and Information Sciences, Middlesex University, London NW44BT, UK.
{E.Ever, RL391, L.Mostarda, P.Shah}@mdx.ac.uk

2 Department of Mathematics and Computer Science, University of Perugia, 06123 Perugia, Italy navarra@dmi.unipg.it

Keywords: clustering, multi-hop, power management, residual energy

Abstract: Prolonging the lifetime of wireless sensor networks has always been a determining factor when designing and deploying such networks. Clustering is one technique that can be used to extend the lifetime of sensor networks by grouping sensors together. However, there exists the hot spot problem which causes an unbalanced energy consumption in equally formed clusters. In this paper, we propose UHEED, an unequal clustering algorithm which mitigates this problem and which leads to a more uniform residual energy in the network and improves the network lifetime. Furthermore, from the simulation results presented, we were able to deduce the most appropriate unequal cluster size to be used.

1 INTRODUCTION

Wireless Sensor Networks (WSNs) are usually self-forming, self-healing networks that interact with their environment to monitor or sense physical parameters such as temperature, acoustics, vibration and humidity among others. They are usually composed of fixed, spatially distributed sensors and a base station (BS). The main functions of a sensor node in a WSN are sensing the environment, processing the raw values and transmitting them to a nearby node until they reach the base station. The role of the base station is to collect all those data received over time, analyse them and ultimately make decisions based on whether certain thresholds have been exceeded or not.

WSNs can operate in two modes: continuous periodic sensing and transmission or event-triggered sensing followed by transmission. To decide on which mode of operation to use is highly application dependent. WSN, being a relatively new technology, leads to many challenges, some of which have still not been met completely. These are the real time, power management, security and privacy factors (Karl and Willig, 2005). The energy challenge is considered to be very important because in most typical usages, WSN nodes are deployed with a limited, non-renewable source of energy, on which the lifetime of the network will depend. One of the solutions put forward by researchers is clustering.

In the clustering operation of WSNs, nodes are partitioned into a number of small groups called clusters. Each cluster has a coordinator, known as a Cluster Head (CH), and a number of member nodes which communicate only to their CH in order to transmit data. Clustering offers some advantages such as data aggregation done at the CH level, distribution of load across all nodes since the role of the CH is not permanently fixed to one particular node; hence rotation of CH is present. CH handles two types of traffic: intra-cluster and inter-cluster communication; the former being communication between member nodes of a cluster and the CH and the latter being the transmission/relay of packets from CH to CH until it reaches the BS. Inter-cluster communication can make use of either single hop or multi-hop forwarding (Zhao and Wang, 2010). In single hop forwarding, each CH directly transmits to the BS, which can cause excessive use of energy for the CH furthest away from the BS making them critical nodes. However, in multi-hop clustering, nodes nearest to the BS tend to deplete their energy the fastest since they are burdened with heavy relay traffic from the rest of the network in addition to their own intra-cluster traffic share. Those nodes closer to the BS tend to die earlier than the rest and as a result, sensing coverage gets reduced and network partitioning becomes apparent, (Zhao and
Wang, 2010; Li et al., 2005; Xuhui et al., 2009; Kim et al., 2008) which is defined as the hot spot problem. Nevertheless multi-hop data transmission from source to BS is usually more energy efficient due to the nature of the wireless channel (Zhao and Wang, 2010).

In this paper, an unequal clustering algorithm (UHEED), based on the HEED algorithm (Younis and Fahmy, 2004), is proposed. HEED is a hybrid energy efficient distributed algorithm which uses 2 parameters to form equal sized clusters: residual energy of a node and node degree or the node proximity to its neighbours. HEED does not make any assumptions about network topology, size and distribution or density of nodes. UHEED creates unequal sized clusters based on the distance of the CH from the BS. The further away a cluster head is from the BS, the larger will be its competition radius and hence the cluster size will be bigger compared to those clusters formed nearer to the BS. By creating unequal sized clusters, the amount of intra-cluster traffic is considerably reduced for the CH’s nearer to the BS.

In this paper, we also attempt to find the right cluster size based on distance from the BS. We first demonstrate that the hot spot problem actually exists in equal size clusters (HEED) and use this as a comparison basis with UHEED. From the analysis of the simulation results; UHEED effectively mitigates the hot spot problem of equal sized clusters and thus balances the energy levels of CHs in the network, provided appropriate sized cluster have been formed. Our simulation results also indicate that the network life-time is significantly improved when compared with HEED (Younis and Fahmy, 2004), LEACH (Heinzelman et al., 2000) and unequal LEACH (Ren et al., 2010) based clustering algorithms.

The rest of this paper is organised as follows: Section 2 presents most recent research within the area of clustering algorithms in WSNs; Section 3 describes the HEED algorithm and how we changed it to develop UHEED; Section 4 describes the UHEED algorithm, the radio and the network models used, the competition radius formula and the simulator built in Java to test this algorithm. Section 5 presents the simulation study conducted for different test cases and a comparison of the results with recent advances in clustering in WSNs. Finally, the paper is concluded along with an opportunity to extend the work further in Section 6.

2 RELATED WORKS

There have been a number of (equal and unequal) clustering algorithms proposed for wireless sensor networks in recent years. Existing studies on unequal clustering approaches are considered in this section.
An unequal clustering model was first proposed in (Soro and Heinzelman, 2005) based on unequal clustering size (UCS) in order to balance the energy level or energy consumption of cluster heads due to heavy inter-cluster relay traffic. Their simulations assumed that cluster heads were located at pre-determined locations and involved using heterogenous node structure. In the case of multi-hop networks, it was demonstrated that UCS was 10-30% better than existing equal clustering models.

In the energy efficient unequal clustering (EEUC) algorithm (Li et al., 2005), the authors propose another unequal clustering algorithm where nodes join clusters of unequal size. However, according to (Gong et al., 2008), EEUC may produce lone nodes since the cluster head election is probabilistic. Zhao et al. propose an unequal layered clustering approach for large scale wireless sensor network (ULCA) (Zhao and Wang, 2010) which assumes a BS at the centre of the grid and creates layers. The layers closer to the base station are smaller in size giving the inner layers more residual energy for inter-cluster traffic. When compared to EEUC (Li et al., 2005), ULCA has a better network lifetime and the overhead for clustering the network is much lower because of the inherent local join and local broadcast mechanism.

In (Gong et al., 2008), a Multi-hop Routing Protocol with Unequal Clustering (MRPUC) is proposed which bears similar characteristics of the algorithm mentioned in (Karl and Willig, 2005; Li et al., 2005). A comparison study conducted in (Gong et al., 2008) demonstrates that MRPUC outperforms an equal clustered version of itself by extending the network lifetime by 34.4%.

In (Heinzelman et al., 2000), the Lower Energy Adaptive Clustering Hierarchy (LEACH) protocol is presented. The algorithm elects cluster heads solely based on probability. No residual energy is taken into account. Moreover, cluster heads use the single hop communication model to forward packets to the base station. A refined version of LEACH can be found in (Xuhui et al., 2009). The lifetime of the sensor networks is maximised by first forming unequal clusters, and then a new threshold algorithm, based on residual energy, is used to elect cluster heads.

In (Yu et al., 2011), an energy-driven unequal clustering (EDUC) algorithm is proposed which discusses the rotation of the role of CH based on either time-driven CH rotation or energy-driven CH rotation approach. In EDUC, it is discussed that the energy-driven CH rotation is better since a new CH is elected only when the energy of the current CH has fallen below some set threshold value and since the election is local, this avoids global topology reconstruction and downtime. In EDUC unequal clusters are formed by having unequal competition ranges and each node can be a cluster head only once during the sensor net-
Figure 3: Network lifetime for 500x500 grid (1000 nodes)

work lifetime. One drawback of EDUC is that it uses single-hop inter-cluster transmission and according to (Zhao and Wang, 2010), multi-hop is better.

A totally new approach is proposed in (Jaichandran et al., 2010) to mitigate the hot spot problem in WSNs in which an area $S$ is divided into $N$ number of cells, each having at least 1 sensor node and the sensor nodes cooperate with neighbour nodes to forward sensed data to the base station. Sensor nodes near the base station act as gateway nodes ($G$ nodes) which tend to die earlier as they have to relay heavy traffic. The novel approach in (Jaichandran et al., 2010) is to introduce additional sensor nodes in the gateway area to help in the relay of traffic to the base station. Although, results indicate that adding an arbitrary number of $G$ nodes does not improve performance, instead a calculation of optimal number of $G$ nodes is first required which, then needs to be added to improve and extend the lifetime of the network.

Recent research in (Bagci and Yazici, 2010) proposes an energy aware fuzzy unequal clustering algorithm (EAUCF) which uses 2 parameters in order to calculate the competition range of the cluster head. From the results obtained, EAUCF outperforms EEUC (Li et al., 2005), ULCA (Zhao and Wang, 2010) and LEACH (Xuhui et al., 2009) for all the performance measures.

In (Pin et al., 2010), the authors propose IEEUC, which is similar to the study in EEUC (Li et al., 2005), since it creates unequal sized clusters as they are further away from the base station, to mitigate the hot spot problem. The main difference between IEEUC and EEUC lies in the competition radius calculation. IEEUC uses the node degree factor, which is based on the number of hops to the base station, to calculate the competition radius. In EEUC, even clusters equidistant from the BS may have different number of member nodes, either too many or too few. On the other hand in IEEUC, we do not have this problem.

The authors of (Nam et al., 2010) propose a variant of the LEACH algorithm by including the residual energy parameter in the calculation of the threshold $T$. In the original LEACH algorithm, a node generates a random value $U \in [0,1]$ and if that value is less than $T$, it becomes a cluster head and also, if $p$ is the ratio of cluster heads, a node can be a cluster head only once during the $1/p$ round. In (Nam et al., 2010) a new threshold formula is used to allow a cluster head to be re-elected based on residual energy and it is shown that this new threshold increases the network lifetime. In (Ren et al., 2010), yet another version of LEACH is presented. Two parameters are used in the setup phase of electing cluster heads: energy ratio (current energy to initial energy) and competition radius. Similar to the study in (Nam et al., 2010), this technique allows cluster heads with more residual energy to be elected. In addition, (Ren et al., 2010) attempts to solve the hot spot problem by creating unequal sized clusters by varying the competition radius. Smaller clusters will be formed near the base station while larger ones will be created as they are further away from it.

In this paper, we introduce an unequal variant of
HEED (UHEED). While HEED defines an equal cluster size, UHEED makes use of a competition radius formula which creates unequal clusters. We compare UHEED with the HEED, LEACH and unequal LEACH algorithms described in (Younis and Fahmy, 2004), (Heinzelman et al., 2000), and (Ren et al., 2010), respectively.

3 FROM HEED TO UHEED

Hybrid Energy Efficient Distributed (HEED) is a distributed clustering algorithm in which two parameters are used to determine the eligibility of a node to become a cluster head. Since prolonging the network lifetime is the main goal, residual energy is used as the first parameter, which allows those nodes with higher residual energy to become cluster heads, thus balancing the overall energy of the network. The second factor intra communication cost, which can be cluster density, allows a node to join a CH with the least number of nodes so as to reduce the load of the intra-cluster traffic on the CH. HEED does not make any assumption about the network such as density or size.

The HEED algorithm is run by each node and is in 3 stages:

In the Initialisation stage, an initial percentage of CH among N nodes is set ($C_{prob}$) which has no impact on the final number of CH to be formed at the end of the algorithm and as such is only necessary to limit the initial number of broadcast. Each node calculates its probability ($CH_{prob}$) of becoming a CH. The $CH_{prob}$ is not allowed to fall below a certain threshold $p_{min}$ in order for the algorithm to terminate in $O(1)$ iterations.

In the Repeat stage, those nodes that could not join a CH, elect to become a tentative CH and send an announcement. This phase iterates itself and each time the $CH_{prob}$ value doubles until it becomes 1. During the iterations, the node can also decide to find a CH instead of becoming one itself.

In the Finalise stage, a node decides its status to become a final CH for the current round or joins the least cost cluster.

Once the clustering process is over, the network enters a data transfer phase. Clustering will occur again after some time in order to rotate the role of the CH and thus balance the energy levels in the network. In this phase, each node of a cluster forwards data to the CH which in turn forwards the aggregated data of its members in a multi-hop fashion (CH to CH) until the base station (BS) is reached.

The problem here is that those nodes nearer to the BS deplete their energy faster than those located further away. This excessive inter-cluster traffic near the BS causes the nearby nodes to die earlier reducing the overall network lifetime.

In UHEED, we attempt to solve the problem of
nodes nearer to the BS dying earlier. In HEED, each CH uses the same competition radius, irrespective of its distance from the BS, hence on average having the same number of nodes. UHEED uses the competition radius formula from EEUC (Li et al., 2005), which creates smaller clusters as the BS is nearer. This overall can improve the network lifetime for multi-hop WSNs which will be shown using simulations. This allows less intra-cluster traffic for the CHs near the BS and hence more energy is allocated to the inevitable higher load of relay traffic.

4 UHEED

In this section, the proposed algorithm is described in detail. Although uneven clustering methods have extensively been discussed in the related works, the uneven clustering approach have not been used together with the well known HEED algorithm. It is necessary to analyse the system parameters in order to provide right cluster sizes for the HEED algorithm with uneven cluster sizes (UHEED).

4.1 Network model

The network model introduced uses a two dimensional representation of the environment and the nodes are deployed randomly following a uniform distribution. We make the following assumptions about nodes: (i) all nodes are homogeneous in terms of energy, communication and processing capabilities; (ii) each node is identified with a unique ID; (iii) nodes can transmit at various power levels depending on the distance of the receivers; (iv) nodes are not mobile that is they remain stationary after the uniformly random deployment; (v) communicating nodes can establish the distance among them; (vi) all nodes know their distance from the base station.

The BS is located away from the sensing grid with no energy concerns at all, and it is considered to be a node with enhanced communication and computation capabilities. The BS is not mobile. The data captured in a cluster is highly correlated, therefore it can be aggregated before being transmitted to the base station.

A network operation model similar to that of (Younis and Fahmy, 2004) consisting of multiple rounds is used. A round starts by triggering the clustering mechanism and after clusters have been formed, the network goes into a data exchange phase. This includes intra-cluster communication where each sensor node sends exactly one message to its cluster head and inter-cluster communication where each aggregated data is sent by the cluster head to the BS (multi-hop data transmission among cluster heads is performed). The round ends when all aggregated data sent by the cluster heads are received at the base station.

The radio model employed uses both the free space and the multi-path channel model and assumes error-free communication links. The simulation parameters used are similar to those in (Younis and Fahmy, 2004). A sensor spends $E_{elec} = 50nJ/bit$ (Younis and Fahmy, 2004) to run the transmitter or receiver circuitry. The energy spent by the transmitter amplifier $E_d$ will depend on the distance $d$ between the sender and the receiver: $E_a = E_{fs}$ assuming a free space model when $d < d_0$ and $E_a = E_{mf}$ assuming a multipath model when $d \geq d_0$, where $d_0 = 75m$ is a constant distance. $E_{fs} = 10pJ/bit/m^2$ and $E_{mf} = 0.0013pJ/bit/m^4$. In order to transmit a $k$-size packet over a distance of $d$ using the above radio model, the amount of energy consumed for transmission $E_{tr}$, can be calculated as:

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

where, $n = 2$ for the free space model and $n = 4$ for the multipath model. The amount of energy $E_{R_i}$ spent to receive a $k$-bit size message is:

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

4.2 Simulation model

UHEED is based on the HEED algorithm (Younis and Fahmy, 2004), however, unlike HEED it uses the competition radius formula given below, in order to create unequal clusters. Since the lifetime of the leaders closer to the BS is more critical, the clusters further away have larger sizes compared to the clusters close to the BS.

$$R_{comp} = \left(1 - c \left(\frac{d_{max} - d(x_i, BS)}{d_{max} - d_{min}}\right)\right) R_{comp}^0 \tag{3}$$

$R_{comp}$ is the maximum competition radius which is predefined. In this work it is defined as the diagonal distance of the sensing grid area divided by 10.

$d_{max}$ and $d_{min}$ are the maximum and minimum distance between sensor nodes and the base station; $c$ is a constant coefficient between 0 and 1.

For the simulation we perform various consecutive rounds (explained in Section 4.1). These are performed until the network is dead. The network is considered dead when all nodes have depleted 99.9%
Network lifetime is based on rounds rather than clock time, and the simulation model used is event triggered. In other words, the simulation clock is always set to the time of the next event until the network dies.

5 Simulation Study

In this study, a simulation program is employed in order to evaluate the proposed UHEED algorithm. The simulation program is first validated by using the numerical results presented in the existing literature (Younis and Fahmy, 2004). For the validation, a grid with dimensions 2000 $\times$ 2000 metres is considered and 1000 nodes are deployed. The cluster radius is taken from 20m to 400m and each experiment value is obtained for an average of 100 runs. The numerical results shows that our implementation of the HEED algorithm exhibit similar behaviour as in (Younis and Fahmy, 2004).

5.1 Existence of hot spot problem in equal size clusters

The results presented in this section are related to the HEED equal clustering algorithm. More specifically, the HEED algorithm has been run for one round, and figures 1 and 2 clearly show that cluster heads nearer to the base station have lower residual energy compared to that of cluster heads further away. The results presented are for cluster radiuses of 20m and 50m, however the behaviour is the same for different cluster sizes as well.

5.2 Network lifetime

In this section the lifetime of UHEED is evaluated by running simulations with different parameters of grid size and number of nodes. When comparing UHEED and HEED the same parameters from (Younis and Fahmy, 2004) have been used. The base station is located at lower right side of the grid, $E_{elec} = 50nJ/\text{bit}$, $E_{a} = 10pJ/\text{bit}/m^2$, number of nodes 300 to 1000 and the initial Energy = 2J. More specifically, we have used two settings: (A) a grid size of 500m x 500m with 1000 nodes; (B) a grid size of 100m x 100m with 300 nodes. We have ensured there is the same number of cluster heads in both UHEED and HEED. This ensures that the two algorithms perform the same number of hops in the inter-cluster communication. Figures 3 and 4 show the best and worst case scenario, for network life time, for UHEED and HEED for:

Case(A): The results show that UHEED outperforms HEED in the best scenario with $c = 0.8$ for UHEED and $r = 35m$ for HEED; but for the worst case, UHEED and HEED both follow a similar pattern with the last node dying after around 500 rounds.
In the best case scenario, the last node for UHEED dies after round 3325 and for HEED, it dies after 900 rounds which is evident in Figure 3. Overall, there is a 250% increase in network lifetime for UHEED in the best case scenario.

Case (B): UHEED outperforms HEED in both the worst and best scenarios. In the best case scenario, the last node in UHEED dies after 4750 rounds, whereas in HEED, it dies after 3340 rounds, which can be observed in Figure 4. Overall, there is an increase of more than 40% network lifetime for UHEED in the best case scenario.

When comparing UHEED, LEACH and unequal LEACH the same parameters from (Heinzelman et al., 2000) and (Ren et al., 2010) have been used. More specifically, the base station is outside the grid, $E_{\text{elec}} = 50 \text{nJ/bit}$, $E_a = 10 \text{pJ/bit/m}^2$, number of nodes 400 and the initial Energy = 0.3J.

UHEED is compared to Unequal LEACH (Ren et al., 2010) and results are shown in Figure 5. The best and worst case for UHEED and Unequal LEACH are considered. For Unequal LEACH we have used the parameters found in (Ren et al., 2010) with a grid size of 200 by 200 and 400 nodes. In order to compare UHEED to Unequal LEACH, the data exchange phase of UHEED has been modified to a single-hop data transmission since Unequal LEACH is a single-hop protocol.

In the simulation study between UHEED and Unequal LEACH, it can be observed from Figure 5 that UHEED outperforms Unequal LEACH by a factor of more than 100% when network lifetime is considered.

Figure 6 shows the residual energy for UHEED, Unequal LEACH and LEACH with respect to First Node Dead (FND) and Half Node Alive (HNA). The residual energy is obtained by calculating the residual energy of the entire network. As can be seen from the Figure 6, in UHEED after the first node is dead, the overall residual energy level for all the cases from $c = 0.1$ to $c = 0.9$ is much higher than LEACH or Unequal LEACH. Also, it is observed that when half of the nodes are alive, the residual energy level in case of UHEED is comparatively higher than LEACH and Unequal LEACH. Hence, from the results seen in Figure 6 for residual energy levels and Figure 5 for the network lifetime, it is seen that the lifetime degradation of UHEED is graceful. This means that not many nodes die very quickly and then the network has very few nodes which are alive for a longer duration, but, as observed in Figure 6, after half the number of nodes are dead in the case of UHEED, there is still higher residual energy level available for the rest of the nodes to continue operation with respect to LEACH and unequal LEACH.

6 CONCLUSIONS

In this paper, we proposed an unequal clustering algorithm for wireless sensor network based on HEED (Younis and Fahmy, 2004). A common prob-
Problem in equal based cluster in sensor networks is the hot spot problem. Our approach to provide a solution was first to implement the HEED algorithm, show that the hot spot problem really exists, and finally attempt to mitigate it by creating UHEED. This algorithm uses a competition radius formula which creates unequal clusters as they are further away from the base station. This effectively allows more inter-cluster or relay traffic and less intra-cluster communication for nodes nearer to the base station, hence preventing their early death. Simulation performed on the UHEED algorithm demonstrated that the lifetime of the network was increased in all test scenarios compared to HEED, LEACH and Unequal LEACH. An interesting study also conducted was regarding the value of the constant $c$ in the competition radius formula. Simulation results showed that a value of $c=0.8$ achieved up to almost 250% improvement in the network lifetime when compared to HEED and almost 100% improvement when compared to unequal LEACH. During the course of this investigation, we found out that values like first node dead (FND), half node alive (HNA) and last node dead (LND) are somewhat affected by the density of the network. Hence, our future work will be to investigate the relationship between the network density and the aforementioned parameters. A mathematical model will also be developed in order to support the simulation results presented in this paper.

REFERENCES


