2016 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)

# Evaluation Study for Clustering in Wireless Sensor Networks

Michael Stein*, Dominic Lerbs, Mohamed Hassan†, Mathias Schnee†,
Immanuel Schweizer*, Karsten Weihe†, Max Mühlhäuser*
*Telecooperation Group, Technische Universität Darmstadt
{michael.stein, schweizer, max}@tk.informatik.tu-darmstadt.de
†Algorithms Group, Technische Universität Darmstadt
{hassan, schnee, weihe}@algo.informatik.tu-darmstadt.de

***Abstract*—Typically, wireless sensor nodes are battery-powered. The network's lifetime depends on the energy consumption of the sensor nodes. Transmitting messages causes a good portion of this energy consumption. Clustering the sensor nodes may reduce energy consumption through local communication and aggregation. Many clustering algorithms exist, but corresponding simulation results are hardly comparable.**

**This paper conducts an extensive simulation study. We compare five popular clustering algorithms in four different scenarios under strictly uniform conditions. Our results indicate that two criteria for clustering algorithms are particularly important: considering residual energy for cluster head selection, and small communication overhead during cluster formation.**

***Index Terms*—Wireless sensor networks, Clustering algorithms, Network architecture, Energy aware design.**

## I. Introduction

Advances in micro-electro-mechanical systems (MEMS) technology, digital electronics and wireless technology allow the development of small, cost-effective and energy-saving sensor nodes. These nodes are able to detect environmental phenomena on a given field, process data, and transmit the results wirelessly to a central processing unit or a base station. These so-called wireless sensor networks (WSNs) [1] may consist of thousands of nodes. The applications of WSNs are manifold, ranging from industrial automation [2] to volcano monitoring [3].

Typically, sensor nodes are battery-powered [1]. Therefore, the nodes have a limited lifetime, a low data rate, increased risk of failure, and limited computing and storage capacity. In recent years, much progress has been made in the development of efficient sensor nodes, but the limited lifetime of sensor nodes is still a big problem. To ensure a long lifetime of a WSN, conducting communication from sensor nodes to the base station in an efficient way is a core requirement.

The longer the distance to the base station, the higher the energy required for transmitting messages. Due to vast fields to be covered [4], the base stations in various scenarios are located far away from the sensor nodes. Examples of such scenarios are wildlife [5] or glacier monitoring [6]. As a consequence, the transmission power required to communicate with the base station is responsible for a major proportion of the nodes' energy consumption. The batteries of the nodes far away from the base station are exhausted earlier, and the observed field is no longer uniformly covered with sensors.

One approach for decreasing the communication cost is to cluster the network into groups of nodes (so-called clusters). All nodes within a cluster transmit their data to a selected cluster head. The cluster head coordinates the communication within the cluster, receives the individual data of its cluster members, aggregates this data and forwards the aggregated data to the base station. Clustering is able to decrease the overall cost of the communication with the base station; it replaces the costly direct communication with the base station with the selection of few cluster heads to communicate with the base station directly. It is important to choose a cluster head that keeps the intra-cluster communication costs low. Therefore, the cluster heads should be chosen to be well connected to other nodes in their cluster. Moreover, it is the task of the clustering algorithm to regularly rotate the cluster head role between the sensor nodes of a cluster to ensure a uniform energy consumption. Doing so, a uniform coverage of the monitored field can be achieved.

A considerable amount of clustering algorithms have been proposed [4]. However, these algorithms have been evaluated with different simulation parameters, which makes a fair comparison of these clustering algorithms difficult.

In this paper, we compare five popular clustering algorithms (LEACH [7], EECS [8], HEED [9], DHAC [10], and AP [11]) under strictly uniform conditions. In our simulation study, we consider four different WSN scenarios and evaluate the lifetime of the WSN and the energy consumption of the sensor nodes for each clustering algorithm.

The rest of the paper is structured as follows. First, Section II outlines the considered clustering algorithms. Section III describes the used energy models, followed by the description of our experiments (Section IV). Then, we present (Section V) and discuss (Section VI) the corresponding results. Section VII concludes this paper.

## II. Examined Algorithms

A base station is a central point that collects the measurements of the sensor nodes within a WSN. The communication of the sensor nodes with the base station is essential for providing results and actual supervision for the WSN. We denote the approach where all sensor nodes communicate with the base station directly as direct sending (DS). As mentioned in Section I, base stations in various scenarios are located far away from the sensor nodes, resulting in high communication costs. Clustering algorithms avoid DS to the base station and perform the communication via selected cluster heads [4].

Low-Energy Adaptive Clustering Hierarchy (LEACH) [7] is a clustering protocol that randomly selects the cluster heads periodically. At the beginning of each round, each node decides based on a stochastic algorithm if it will become a cluster head. Then, each other node selects a cluster head that minimizes the communication energy. This process is repeated periodically to fairly distribute the cluster head task (and its associated communication costs) across all WSN nodes. This is expected to lead to an increase of the lifetime of the WSN. However, the random selection may reduce energy efficiency in some scenarios. In the worst case, the sensor nodes do not choose to communicate via the cluster heads to the base station because the cluster heads are not in a central position for these nodes. This direct communication with the base station results in a higher energy consumption.

Hybrid Energy Efficient Distributed clustering (HEED) [9] uses two parameters to determine the cluster heads. The primary parameter is based on the residual energy of a sensor node. The secondary parameter reflects the cost of intra-cluster communications, e.g., the average distance to neighbors. Nodes with a high residual energy are considered as candidates for becoming cluster heads. The secondary parameter selects the final cluster heads from these candidates and determines which sensor nodes will connect to the cluster heads. HEED ensures a good distribution of cluster heads in the field. Thus, the algorithm promises a balanced energy consumption and a long network lifetime. The biggest drawback of HEED is the iterative process. HEED takes considerably more time for cluster formation than many other algorithms.

Energy Efficient Clustering Scheme (EECS) [8] builds clusters with little messaging overhead and evenly distributed cluster heads. A cost function for a balanced energy consumption for all the sensor nodes is used to form the clusters. Each node that is no cluster head decides on the basis of the weighted cost function which cluster head to choose. One part of the cost function consists of the distance of the sensor nodes from each cluster head, the other part takes into account the distance of the cluster head to the base station. The first part minimizes the energy consumption of the cluster members, while the second part chooses the cluster heads to be positioned closer to the base station to consume less energy during transmission. One disadvantage of EECS is the necessity of knowledge of the distance of each node to the base station.

Unlike most of the other algorithms, Distributed Hierarchical Agglomerative Clustering (DHAC) [10] first builds the clusters and then determines the cluster heads. DHAC avoids re-clustering. The choice of the cluster heads is periodically rotated on all members of a cluster to ensure a balanced energy consumption. A problem of this rotation scheme is that cluster heads far away from the base station fail earlier than closer ones. Furthermore, the energy consumption in the intra-cluster communications are expected to be comparatively high as the cluster heads may be located at the edge of a cluster.

TABLE I
RANGES WITH ENERGY CONSUMPTION FOR BOTH ENERGY MODELS

| | | LEACH-C | | MICAz | |
|---|---|---|---|---|---|
| Index | Transmission power | Distance | Energy consumption | Distance | Energy consumption |
| 0 | -25 dBm | 7 m | 5.5 mA | 3 m | 8.5 mA |
| 1 | -15 dBm | 23 m | 6.2 mA | 12 m | 9.9 mA |
| 2 | -10 dBm | 43 m | 7.7 mA | 22 m | 11.2 mA |
| 3 | -7 dBm | 61 m | 9.8 mA | 31 m | 12.5 mA |
| 4 | -5 dBm | 78 m | 11.3 mA | 39 m | 13.9 mA |
| 5 | -3 dBm | 99 m | 20.2 mA | 50 m | 15.2 mA |
| 6 | -1 dBm | 125 m | 42.4 mA | 64 m | 16.5 mA |
| 7 | 0 dBm | 140 m | 63.2 mA | 72 m | 17.4 mA |
| 8 | 1 dBm | 158 m | 98.5 mA | 80 m | 21.7 mA |
| 9 | 3 dBm | 199 m | 237.2 mA | 102 m | 42.3 mA |
| 10 | 5 dBm | 251 m | 587.3 mA | 130 m | 95.4 mA |
| 11 | 7 dBm | 317 m | 1475.9 mA | 163 m | 224.7 mA |
| 12 | 10 dBm | 445 m | 5776 mA | 232 m | 870.8 |

Affinity Propagation AP [11] iteratively exchanges messages over time between the sensor nodes. These messages are updated based on formulas that search for the minima of a certain chosen energy function. The clusters gradually emerge during the message exchange process. One drawback of this algorithm is the high messaging overhead required for cluster formation. AP in some scenarios may require more energy than performing direct communication with the base station.

## III. ENERGY MODELS

To calculate the energy consumption of sensor nodes, we use two different energy models: LEACH-C and MICAz.

The first energy model (LEACH-C) builds on the one specified in [12]. Up to a distance $d$ of 75 meters, the free space model is used to calculate a sender's energy consumption ($lE_{elec} + l\epsilon_{fs}d^2$, where $l$ gives the packet size, and $E_{elec}$ and $\epsilon_{fs}$ are set to $820nJ/bit$ and $164.14pJ/bit/m^2$, respectively). For larger distances, the multipath model is used ($lE_{elec} + l\epsilon_{mp}d^4$, where $\epsilon_{mp}$ is $0.021pJ/bit/m^4$). The energy consumption for receiving packets is computed as $lE_{elec}$.

The second energy model (MICAz) uses the energy parameters of the Crossbow MICAz sensor node [13]. The data sheet specifies a maximum range of 75 meters. Thus, we assume this range for a transmission level of $0dBm$ and determined the ranges of lower transmission levels experimentally in the WSN simulator SIDnet-SWANS [14]. As the distance to the base station may be higher than 75 meters, we introduced additional transmission levels (thus, reflecting a sensor node with multiple wireless interfaces) and used the multi-path model to estimate the corresponding transmission power. We extracted the required parameters $E_{elec}$ ($1269nJ/bit$) and $\epsilon_{mp}$ ($0.0417pJ/bit/m^4$) based on the information given in the data sheet. In packet receive mode, we assume a current draw of $19.7mA$, in accordance with the data sheet [13].

For both models, the transmission power is variable and is adjusted for each packet to one of 13 discrete levels based on the required range (see Table I). For the MICAz model, the use of the highest five levels (index 8-12) is only allowed to communicate with the base station. Each sensor is battery-powered with an initial capacity of $5mAh$ (3V).

## IV. Experiment

We conducted experiments for four different scenarios, which are summarized in Table III and Figure 1. The main function of a WSN is to provide information about the environmental phenomena occurring within a field on earth. Thus, we deploy the sensor nodes in a two-dimensional field and let the nodes periodically report measurements to the base station. The first three scenarios use the energy model LEACH-C, but differ with respect to the position of the base station. Scenario 1 simulates networks in which the base station is located at the edge of the sensor field. In Scenario 2 and 3, the base station is far away from the sensors. Consequently, direct transmissions to the base station consume considerably more energy. Scenario 4 uses the energy model MICAz and the same position of the base station as Scenario 2, i.e., Scenario 4 investigates the impact of the energy model. For each scenario, we used the WSN simulator SIDnet-SWANS [14] and used the IEEE 802.15.4 protocol for message transmissions.

We simulated a squared field with a side length of 100 meters. We varied the number of nodes from 50 to 350 to control the node density. The nodes are stationary, and know their own position coordinates and the position coordinates of the base station. Based on the minimum calculated distance, the nodes decide to transmit their messages either to the cluster head or directly to the base station.

We conduct simulation rounds until the batteries of all nodes are empty. Each round consists of a clustering phase and a sensing phase. In the clustering phase, clusters are formed as specified by the clustering protocol. The sensing phase has a duration of 30 minutes, and each node periodically reports data every 30 seconds. For the reporting, each sensor node measures 5 bytes of data, and transmits it to the cluster head or directly to the base station. The cluster head aggregates the individual measurements and transmits the aggregated data in a single message of at most 100 bytes to the base station. If a cluster head dies during a round, this does not trigger immediate re-clustering, i.e., the sensors behave as if the cluster head would still be alive until the end of the round.

To compare the different clustering algorithms, we investigate the number of rounds until the nodes run out of energy. We also determine the average number of cluster heads and the energy consumption during the clustering and sensing phase. The results are based on the averages of 50 runs per scenario. The nodes are placed randomly in each of the 50 runs. Error bars in the plots denote the 95% confidence intervals.

TABLE II
Average energy consumption (mJ) in Scenario 1

| | 50 nodes | | | 250 nodes | | |
|---|---|---|---|---|---|---|
| | Cluster | Sensing | Overall | Cluster | Sensing | Overall |
| LEACH | 13 | 131.16 | 144.16 | 42.81 | 126.06 | 168.89 |
| HEED | 19.6 | 131.15 | 150.75 | 21.06 | 145.2 | 166.26 |
| EECS | 16.64 | 126.76 | 143.4 | 40.26 | 132.94 | 173.2 |
| DHAC | 1.12 | 132.05 | 133.17 | 2.45 | 130.94 | 132.85 |
| AP | 14.59 | 161.2 | 175.79 | 82.59 | 207.16 | 289.75 |
| DS | 0 | 126.25 | 126.25 | 0 | 126.85 | 126.85 |

TABLE III
The parameters of the four scenarios. MinD (MaxD) gives the minimum (maximum) distance from the base station.

| Scenario | Energy model | Base station position | MinD | MaxD |
|---|---|---|---|---|
| Scenario 1 | LEACH-C | (50,110) | 10 m | 120 m |
| Scenario 2 | LEACH-C | (50,175) | 75 m | 182 m |
| Scenario 3 | LEACH-C | (50,200) | 100 m | 206 m |
| Scenario 4 | MICAz | (50,175) | 75 m | 182 m |

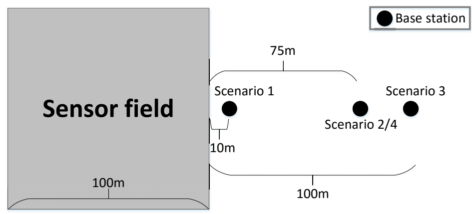



Fig. 1. Base station placement for the different scenarios

## V. Results

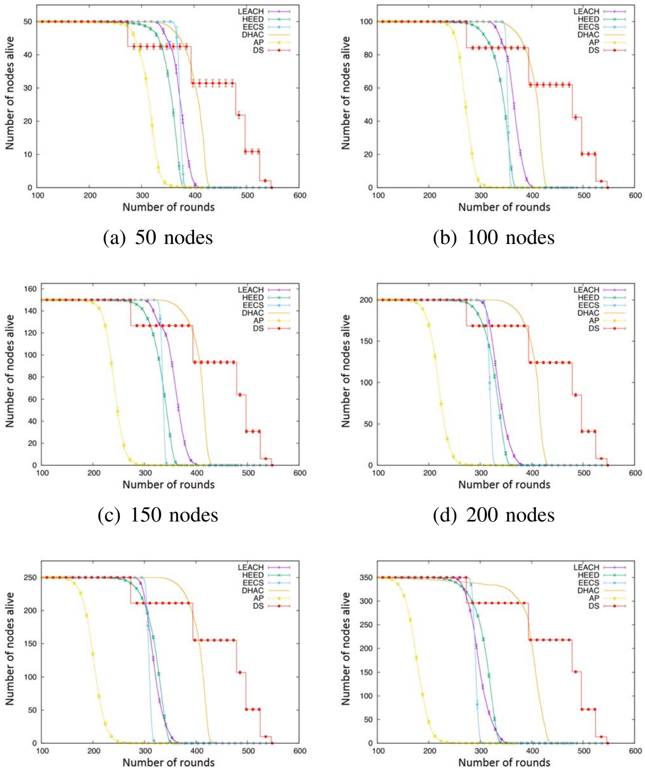
### A. Scenario 1: the base station located at the edge of the field (10 meters), LEACH-C energy model

Figure 2 shows the number of sensor nodes that are alive over the number of rounds. For AP, the batteries of the sensor nodes are exhausted much earlier than for the other algorithms. Increasing the node density results in a further performance degradation, causing the nodes to fail even earlier. There are two main reasons for this circumstance. First, AP introduces a high message overhead and uses a lot of energy during the clustering phase. Second, AP does not change the optimum transmission radius in this scenario. Thus, increasing the node density increases the number of neighbors of a sensor.

The curves of HEED, LEACH and EECS are close to each other. The first sensor node to fail is in the network clustered by HEED, which is due to the unfavorable selection of cluster heads. Nevertheless, the performance of HEED improves compared to LEACH and EECS with increasing node density because the cluster heads chosen in HEED send messages in a restricted radius, which becomes smaller as the node density increases. In contrast, the cluster heads in both LEACH and EECS send messages at higher transmission ranges. Thus, LEACH and EECS consume more energy than HEED when the node density is increased.

Table II summarizes the energy consumption for the clustering and the sensing phase for 50 and 250 nodes. DHAC benefits from the avoidance of re-clustering and from the fact that each node may directly reach any other node. This allows a single large cluster to be formed in which only one cluster head sends the data to the base station. The energy consumption of DHAC is almost the same for both 50 and 250 nodes, while the energy consumption of LEACH, HEED EECS and AP increase dramatically with the node density.

DS performs remarkably well in Scenario 1. Although the first nodes die earlier than for the clustering algorithms with exception of AP, most nodes survive longer with DS than for all clustering algorithms. Moreover, DS showed the best

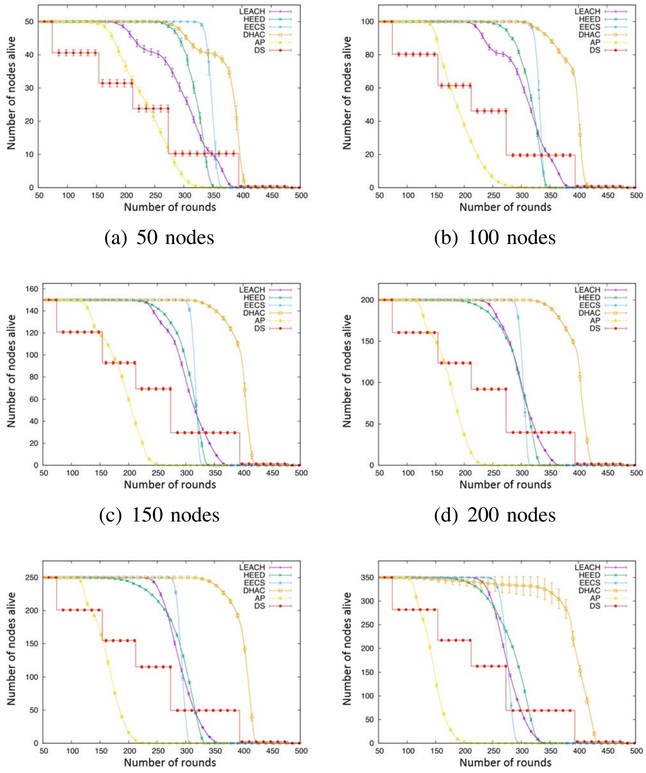



(a) 50 nodes (b) 100 nodes

(c) 150 nodes (d) 200 nodes

(e) 250 nodes (f) 350 nodes

Fig. 2. Scenario 1: Alive sensor nodes over the number of rounds




(a) 50 nodes (b) 100 nodes

(c) 150 nodes (d) 200 nodes

(e) 250 nodes (f) 350 nodes

Fig. 3. Scenario 2: Alive sensor nodes over the number of rounds

average energy consumption within the network compared to the clustering algorithms. This result is due to the fact that the base station is close to the sensor nodes.

### *B. Scenario 2: the base station located far away from the sensor nodes (75 meters), LEACH-C energy model*

Figure 3 shows the outage times of the sensor nodes when the base station is 75 meters away from the field. The sensor nodes of AP were the earliest to fail. In addition to the high energy consumption of the algorithm during the clustering phase, the relatively large number of cluster heads (see Table V), which must send the data to a remote base station, causes a huge performance degradation. Thus, in this scenario, the performance gap between AP and the other algorithms is even more apparent than in Scenario 1.

LEACH performs badly because it selects cluster heads purely by chance. Distant sensor nodes are often elected to be cluster heads and consume a lot of energy. This issue loses in significance when the node density increases. In Table IV, we see that the energy consumption during the sensing phase is reduced with higher density by approximately 27mJ.

The lifetimes of the sensor nodes for HEED and EECS show a similar trend as in Scenario 1, but the difference between these two algorithms is slightly larger in Scenario 2. EECS takes the cluster head's expected energy consumption into consideration when choosing a cluster head. In contrast, the HEED cluster members try only to minimize their own

TABLE IV
AVERAGE ENERGY CONSUMPTION (MJ) IN SCENARIO 2

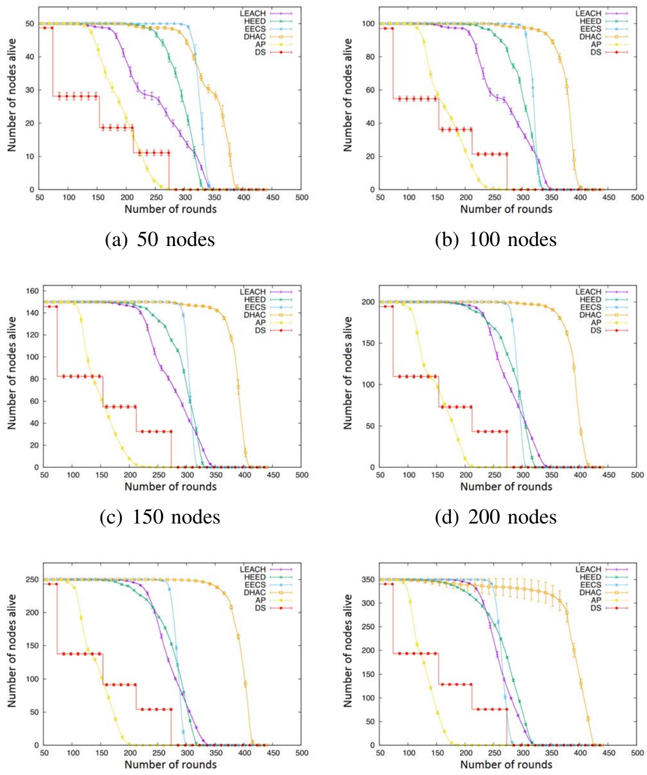
| | 50 nodes | | | 250 nodes | | |
|---|---|---|---|---|---|---|
| | Cluster | Sensing | Overall | Cluster | Sensing | Overall |
| LEACH | 13.66 | 172.73 | 186.39 | 40.91 | 145.92 | 186.83 |
| HEED | 21.68 | 145.78 | 167.46 | 24.37 | 163.41 | 187.78 |
| EECS | 17.26 | 136.51 | 153.77 | 50.62 | 132.92 | 183.54 |
| DHAC | 1.22 | 145.57 | 146.79 | 2.54 | 133.12 | 135.66 |
| AP | 41.07 | 196.12 | 237.19 | 82.67 | 272.88 | 355.55 |
| DS | 0 | 318.19 | 318.19 | 0 | 323.62 | 731.99 |

consumption. Since in this scenario the transmission to the base station uses more energy than in Scenario 1, the smart selection of cluster heads of EECS turns out to be beneficial.

With increasing node density, a WSN clustered by DHAC results in the smallest energy consumption in average (Table IV), while the energy consumption of LEACH, HEED and EECS increases with the node density. In this scenario, all the clustering algorithms except AP showed a huge improvement in energy consumption compared to DS.

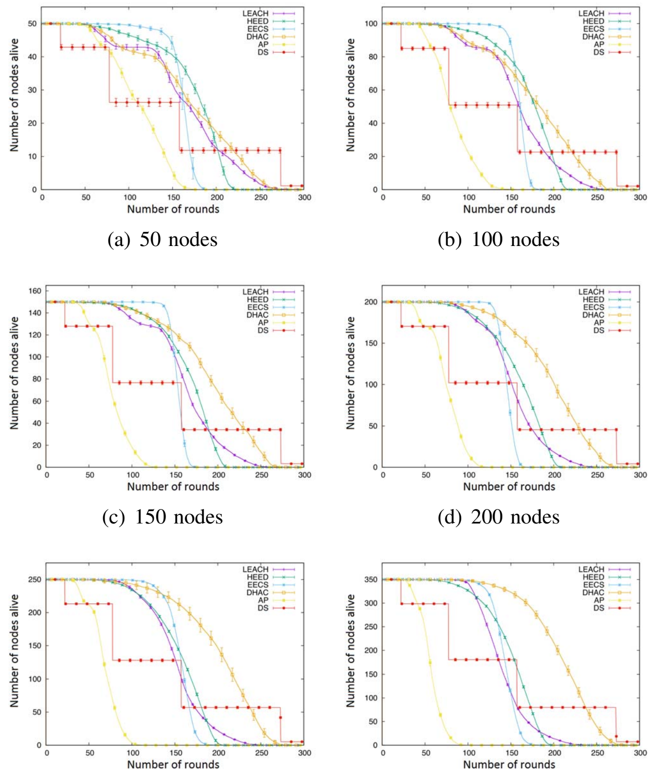
### *C. Scenario 3: the base station located far away from the sensor nodes (100 meters), LEACH-C energy model*

Scenario 3 shifts the base station even further away from the sensor field than Scenario 2. The further the base station is moved away, the more improvement in terms of network lifetime is achieved via clustering. This is particularly evident in Scenario 3 (see Figure 4). Sensor nodes performing direct




(a) 50 nodes (b) 100 nodes (c) 150 nodes (d) 200 nodes (e) 250 nodes (f) 350 nodes

Fig. 4. Scenario 3: Alive sensor nodes over the number of rounds




(a) 50 nodes (b) 100 nodes (c) 150 nodes (d) 200 nodes (e) 250 nodes (f) 350 nodes

Fig. 5. Scenario 4: Alive sensor nodes over the number of rounds

TABLE V
AVERAGE NUMBER OF CLUSTER HEADS IN SCENARIO 2

| Nodes | LEACH | HEED | EECS | DHAC | AP |
|---|---|---|---|---|---|
| 50 | 3.49 | 1.2 | 1.04 | 1 | 6.62 |
| 100 | 2.99 | 1.25 | 1.19 | 1 | 10.26 |
| 150 | 4.5 | 1.27 | 2.45 | 1 | 18.19 |
| 200 | 3.99 | 1.49 | 2.53 | 1 | 20.76 |
| 250 | 4.99 | 2.3 | 2.54 | 1 | 23.36 |
| 350 | 3.51 | 5.9 | 6.16 | 2.78 | 31.07 |

TABLE VI
AVERAGE ENERGY CONSUMPTION (MJ) IN SCENARIO 4

| | 50 nodes | | | 250 nodes | | |
|---|---|---|---|---|---|---|
| | Cluster | Sensing | Overall | Cluster | Sensing | Overall |
| LEACH | 30.89 | 350.41 | 381.3 | 111.37 | 262.89 | 374.26 |
| HEED | 42.24 | 290.42 | 332.66 | 44.01 | 302.08 | 346.09 |
| EECS | 38.27 | 300.7 | 338.34 | 118.06 | 232.96 | 351.02 |
| DHAC | 5.63 | 366.16 | 371.79 | 17.25 | 281.06 | 298.31 |
| AP | 116.29 | 427.77 | 544.06 | 246.51 | 729.89 | 976.4 |
| DS | 0 | 717.59 | 717.59 | 0 | 731.99 | 731.99 |

communication with the base station fail early. If DS is executed in the WSN with 50 nodes, the first node fails in round 32, while the last node fails in round 273. On the other hand, using a clustering algorithm such as EECS, the first node in Figure 4(a) fails in round 291. In WSNs with few nodes, EECS promises the best performance. In contrast, in the case of LEACH and DHAC, the very distant sensor nodes fail early. With increasing node density, DHAC shows better performance than the other algorithms. AP performs still worse than the other clustering algorithms. However, AP has identical performance as LEACH in some cases (see Figure 4(a)). Considering the residual energy of nodes, AP may achieve a better load distribution than LEACH.

### *D. Scenario 4: the base station located far away from the sensor nodes (75 meters), MICAz energy model*

Scenario 4 differs from Scenario 2 only with respect to the used energy model. As Scenario 4 uses the MICAz model, the achievable transmission range in Scenario 4 is much lower. Some algorithms, for example DHAC, perform best when the number of nodes in their transmission range is high. Hence, for lower number of nodes, we expect these algorithms to perform worse compared to Scenario 2.

Figure 5 shows the lifetime of the sensor nodes. Again, in this scenario, AP results in the worst performance. The high number of cluster heads (see Table VII) and many messages in the clustering phase result in that bad performance.

As can be seen in Table VI, in networks with low node density, LEACH and DHAC consume more energy in the sensing phase than HEED and EECS. The reason is that LEACH and DHAC select more cluster heads (see Table VII) than the other two algorithms and, thus, need to conduct more transmissions to the base station in each round. DHAC can no longer form a single large cluster, but now produces several small clusters. In fact, DHAC chooses more cluster heads than

HEED and EECS. This has a negative impact on the energy consumption during the sensing phase. In denser networks, on the other hand, expensive transmissions are replaced with more favorable intra-cluster communication. Thus, the lifetime of distant nodes increases. DHAC particularly benefits from avoiding re-clustering of the network.

The issue of unfavorable cluster head selection of HEED can be observed again. Furthermore, it can be seen that EECS has a similar total energy consumption as HEED.

TABLE VII
AVERAGE NUMBER OF CLUSTER HEADS IN SCENARIO 4

| Nodes | LEACH | HEED | EECS | DHAC | AP |
|---|---|---|---|---|---|
| 50 | 3.51 | 2.4 | 3.67 | 4.68 | 6.68 |
| 100 | 5.02 | 2.57 | 3.92 | 5.27 | 10.07 |
| 150 | 4.51 | 2.59 | 3.97 | 5.12 | 14.55 |
| 200 | 6.02 | 4.63 | 3.98 | 5.6 | 20.37 |
| 250 | 4.98 | 4.68 | 3.95 | 6.01 | 24.46 |
| 350 | 6.98 | 6.11 | 4.01 | 5.9 | 27.75 |

## VI. DISCUSSION

In all scenarios, AP resulted in the worst performance compared to the other clustering algorithms. In WSNs, it is important to have a low messaging overhead to build the clusters and to select the best possible cluster heads. However, AP has a high messaging overhead, resulting in higher energy consumption and early sensor node failures. Moreover, the random selection of cluster heads in WSNs is unfavorable because the sensor nodes require different amounts of energy.

LEACH significantly improves the lifetime of the network compared with direct transmission to the base station. The residual energy is an important criterion in selecting cluster heads. However, this is not considered by LEACH. Still, the LEACH algorithm is easy to implement, and LEACH might come as a quick fix for networks where the nodes have a similar distance to the base station.

HEED shows advantages over LEACH especially in networks with low node density. HEED does not only consider the residual energy but also the expected energy consumption in the selection of cluster heads, which results in a better performance. However, the iterative sequence required during the clustering phase is inconvenient.

EECS distributes the energy consumption most uniformly over the sensor nodes. However, it is important to state that the cost function takes the distance of each sensor node to the base station as input. The presence of this information can not always be assumed. Nevertheless, the cost function of EECS is expected to be able to ensure a long lifetime of the sensor nodes compared to the other algorithms.

In most cases, DHAC ensures the longest network lifetime through the rotation of cluster heads and the avoidance of re-clustering. Also, the low average number of cluster heads results in good performance compared to the other algorithms.

## VII. CONCLUSION

In this study, five clustering algorithms for WSNs were investigated, implemented and evaluated for four different scenarios. The comparison between the clustering algorithms provides important general insights into the design of clustering algorithms. Less messages exchanged between the nodes during the clustering phase and a shorter transmission radius promise a long lifetime for the WSN. Furthermore, the residual energy should be considered during the cluster head selection to further increase the network lifetime and improve the fairness of load distribution. Consideration of these factors results in a lower energy consumption within the network. Overall, DHAC and EECS show the best performance as both consider the factors just mentioned. EECS performs better if the node density is low. If the density is high enough, DHAC outperforms all other algorithms consistently.

Given the current results, we plan to extend the study by also considering multi-hop clustering algorithms. Here, the cluster head can be more than one hop away from any cluster member. We will also consider different routing schemes for the backbone, i.e., the interconnection of the cluster heads.


## ACKNOWLEDGMENT

This work has been funded by the DFG as part of projects A01 and B02 within the CRC 1053 – MAKI.


## REFERENCES


[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] A. Kumar Somappa, K. Øvsthus, and L. Kristensen, "An Industrial Perspective on Wireless Sensor Networks – A Survey of Requirements, Protocols, and Challenges," *IEEE Commununications Surveys & Tutorials*, vol. 16, no. 3, pp. 1391–1412, 2014.

[3] G. Werner-Allen, K. Lorincz, J. Johnson, J. Lees, and M. Welsh, "Fidelity and Yield in a Volcano Monitoring Sensor Network," in *Proceedings of USENIX OSDI*, 2006, pp. 381–396.

[4] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14–15, pp. 2826–2841, 2007.

[5] G. P. Picco, D. Molteni, A. L. Murphy, F. Ossi, F. Cagnacci, M. Corrà, and S. Nicoloso, "Geo-referenced Proximity Detection of Wildlife with WildScope: Design and Characterization," in *Proceedings of ACM/IEEE IPSN*, 2015, pp. 238–249.

[6] K. Martinez, R. Ong, J. Hart, and J. Stefanov, "GLACSWEB: A Sensor Web for Glaciers," in *Adjunct Proceedings of EWSN*, 2004, pp. 1–4.

[7] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," in *Proceedings of HICSS*, 2000, pp. 1–10.

[8] M. Ye, C. Li, G. Chen, and J. Wu, "EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks," in *Proceedings of IEEE IPCCC*, 2005, pp. 535–540.

[9] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[10] C.-H. Lung and C. Zhou, "Using Hierarchical Agglomerative Clustering in Wireless Sensor Networks: An Energy-efficient and Flexible Approach," *Ad Hoc Networks*, vol. 8, no. 3, pp. 328–344, 2010.

[11] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.

[12] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[13] "MICAz datasheet," http://www.openautomation.net/uploadsproductos/micazdatasheet.pdf, accessed: 2016-02-11.

[14] O. C. Ghica, G. Trajcevski, P. Scheuermann, Z. Bischof, and N. Valtchanov, "SIDnet-SWANS: A Simulator and Integrated Development Platform for Sensor Networks Applications," in *Proceedings of ACM SenSys*, 2008, pp. 385–386.