HetEng: An Improved Distributed Energy Efficient Clustering Scheme for Heterogeneous IoT Networks

Abstract—Network lifetime is always a challenging issue in battery-powered networks due to the difficulty of recharging or replacing nodes in some scenarios. Clustering methods are a promising approach to tackle this challenge and prolong lifetime by efficiently distributing tasks among nodes in the cluster. The present study aimed to improve energy consumption in heterogeneous IoT devices using an energy-aware clustering method. In a heterogeneous IoT network, nodes (i.e., battery-powered IoT devices) can have a variety of energy profiles and communication capabilities. Most of the existing clustering algorithms have neglected the heterogeneity of energy capacity among nodes and assumed that they are of the same energy level. In this work, we present HetEng, a Cluster Head (CH) selection process that extended an existing clustering algorithm, named Smart-BEEM. To this end, we proposed a statistical approach that distributes energy consumption among highly energetic nodes in the network topology by constantly changing the CH role between the nodes based on their real energy levels (in joules). Experimental results showed that HetEng resulted in a 6.6% increase of alive nodes and 3% improvement in residual energy among the nodes in comparison with Smart-BEEM. Moreover, our method reduced the total number of iterations by 1% on average.

Index Terms—IoT networks, Clustering, Low power networks, Energy efficiency, Energy distribution, CH selection.

I. INTRODUCTION

Internet of Things is considered as a networking paradigm that connects a wide range of devices to the internet, From wearable gadgets which capture data from body [1] to sensors and devices that interact with the environment in a smart home and also surveillance systems in smart cities [2]. Optimal energy consumption at battery-powered nodes is always been a challenging topic for researchers in this era. The rapid growth of hardware technologies has made it possible for devices to become smaller, which poses new energy challenges that need to be addressed [3]. This single constraint imposes several others in regard to choices of routing protocol, network coverage, and longevity [4]. A well-known network topology management method used to tackle this challenge is called the clustering method [5], where the nodes are grouped into several clusters and one or many cluster heads (CHs) are elected in the network [2]. Many clustering algorithms have been proposed in the context of homogeneous IoT and Wireless Sensor Networks (WSN) [2]. [6]. However, most of these algorithms have neglected the diversity of energy profiles, assumed that nodes had been supplied with the same energy level and this will cause fast energy depletion in nodes with weak power. To be utilized in the new paradigm of IoT networks, the available clustering methods, need to be modified to consider the heterogeneity of nodes in IoT environments [4].

In this paper we propose a distributed energy efficient clustering method named HetEng to detect nodes with high power and distribute energy consumption by changing the CH role using a statistical way continually. We evaluated our approach using MATLAB platform and our method shows a significant improvement in alive nodes and residual energy.

The remainder of this paper is organised as follow: The related works is presented in section 2, The network model for formulating the problem in section 3, based on Smart-BEEM our detailed modification and improvement introduced in section 4, Performance evaluation and numerical results are indicated in section 5. Finally, in section 6 we discussed the conclusion and future works.

II. RELATED WORKS

Various energy-efficient approaches have been proposed in clustering context, each using specific methods for selection of the CH and for routing between the CH and the nodes. The reviewed works are categorized in Table 1.

| TABLE I | ENERGY-EFFICIENT CLUSTERING METHODS |
|-----------------------------|---------------------------|
| Clustering Approach         | Notable Methods |
| Daily Cycle Ratio           | EnergyIoT [7] |
| Data compression & fusion   | LIDAR [8] |
| Meta-Heuristic              | SCE, PSO [9], [10], FAMACROW [11] |
| Mobility                    | 12 |
| Routing & Hierarchical      | LEACH [13], Modified-LEACH [14] |
| Statistical/Mathematical    | HEED [15], BEEM/AM [16], Smart-BEEM [17] |

According to Table 1, duty cycle methods are proposed for reduction of energy consumption in the Internet of Things, where the nodes switch between the active and sleep modes. Many heuristic algorithms have been proposed for that purpose [7]. Compression and data fusion methods are also used to reduce energy consumption, which are mainly of heuristic nature [17]. Clustering methods are also used for transmission at shorter distances, where data are transferred to the cluster, at a shorter distance, rather than to the sink,
at a longer distance. Selection of the optimal cluster head using heuristic and meta-heuristic algorithms is effective in reduction of energy consumption. This can be based on parameters such as temperature, residual energy, load, and number of remaining nodes. Another challenge that can be added to clustering algorithms is mobility, which is evaluated using other parameters, such as maximum lane speed and traffic flow rate. In \cite{10}, a heuristic similar to the existing mathematical methods is presented for optimal selection of cluster heads, thereby reducing energy consumption in IoT. In \cite{13} the authors divide nodes into a number of clusters according to the LEACH algorithm. Cluster heads then generate a minimum spanning tree according to Prim’s algorithm. The tree is balanced constantly using the AVL algorithm.

In \cite{4} the authors propose a smart clustering algorithm (Smart- BEE(M), based on one in their earlier work, referred to as BEE(M), to achieve energy efficiency and support for the Quality of user Experience (QoE) in communication in cluster-based IoT networks. It is a context- and user-behavior-aware approach, aiming to simplify the selection of beneficial communication interfaces and cluster headers for data transmission on the part of IoT devices.

In another paper, Hy-IoT provides an efficient hybrid energy-aware clustering communication protocol for green IoT network computing. It also provides a real IoT network architecture for testing the proposed protocol against the existing ones. Efficient cluster head selection boosts the utilization of the node’s energy contents, thereby increasing network lifetime and the rate of packet transmission to the base station. Hy-IoT uses different weighted election probabilities for selection of a cluster head based on the heterogeneity level of the region. Besides prolonging network lifetime, it increases throughput with respect to the amounts obtained by SEP, LEACH, and Z-SEP \cite{19}.

III. NETWORK MODEL

Hereon, We define our considered clustering approach adopted in the research, and formalized the problem. An IoT network contains various IoT devices (e.g., smartwatches and sensors with different battery capacities). Therefore, we define $N = \{N_1, ..., N_i\}$ as heterogeneous set battery-powered IoT nodes in a network field with the dimension of $Z \times Y$. As an example, Figure 1 illustrates a network field of $100m \times 100m$. Due to the heterogeneous energy level of the nodes, they have various values of energy capacity. We denote nodes energy values as $E = \{E_1, ..., E_i\}$ (in joules). On that basis, it is more challenging here than in a homogeneous network scenario to calculate and select a high-energy node to take the CH role in the network, since some nodes have higher or infinite initial energy and some have lower. Therefore, previous works conducted in the context of traditional clustering may not detect optimum CHs efficiently in IoT scenarios. Moreover, increasing number of iterations in the CH selection procedure imposes higher overload on the entire network due to the large number of broadcast packets containing information about residual energy and the density of nodes around them. In each iteration, each node calculates a grade which indicated the probability of being elected as CH based on its own energy level. Since the grade of every node should reach the point of 1.0 before it can be declared as a final CH, the packets are sent to the CHs or gateways according to the positions. In a homogeneous network, the initial energy for all the cluster nodes (in the first iteration) is the fixed value of 100% for the entire time denoted as T, representing the first iteration, and the value changes to 80% in $T + 1$ . . . n (the following iterations), continuing to decrease until it reaches 0%, where the node becomes inactive.

This affects network coverage and density, which results in unbalanced energy consumption in different areas which inevitably results in electing low-energy nodes as CH. In contrast, in heterogeneous IoT networks it is possible to elect high energy nodes to play the CH role. Figure 1 illustrates the network and node distribution in the initial phase.

IV. PROPOSED APPROACH

Clustering should be done so that highly energetic nodes that are denser in terms of position receive more packets from the neighboring nodes. Therefore, it is necessary to calculate the residual energy values of the nodes so that the re-clustering operation is performed according to the states and energy statuses of the nodes. The proposed statistically-based method is a re-modification of CH selection methods based on a change in the role of the CH through calculation of the amount of real energy in joules divided by the average of the surroundings and distribution of the energy consumption value among the initial highly-energetic nodes. In the first round, nodes with high real energy values rather than terms of percentage are selected. Then, as mean neighbor energy status is used, the energy consumption value is distributed so that high-energy nodes play a role in the sending of packets to the gateways. Therefore, energy consumption in the cluster is reduced. With these parameters taken into account, the proposed model for CH selection is shown as in Eq. (1).

$$CH_{prob} = \frac{E_{rest current node in j} \times M_{max neighbor nodes is j}}{\sqrt{N - 1} \sum_{i=1}^{N} (E_i - M)^2}$$

In Eq. (1) $CH_{prob}$ is the probability of the CH role, which is based on the following factors. $C_{prob}$ has a fixed probability.
value of 5%. $E_{rest}$ is the amount of node energy that is considered in joules. $M$ is mean neighboring node energy, which appears in the denominator of the average residual energy function of the target node, also calculated in joules. In fact, a comparison is made in this part of our algorithm between the residual energy of the node and the residual energy state of the surrounding nodes in a cluster. The amount of node energy that is considered in joule. In the next term of the equation, the residual energy of the desired node is again considered in joules, but sample deviation is used this time for calculation of the variance of the desired node. In the deviation section, the criterion $N$ is the number of surrounding nodes, $E_i$ is the residual energy of each node, which is subtracted from the mean for the neighboring nodes.

To calculate the relative variance of residual energy in the network, we used sample deviation for each cluster; that is, a comparison is made between the residual energy of the node and the sample deviation of the average energy of the remaining neighbour nodes around. In fact, the variance of residual energy of the node is calculated relative to its neighboring nodes. A positive value for the standard deviation means that the examined node has more energy than the average energy of its neighbour nodes, and therefore, will get elected as the CH. In Eq (2), another term will be added, which will serve as a criterion for selecting the CH based on the conditions. On that basis, the number of repetitions in the competition round with other nodes will also be considered.

$$CH_{prob} = C_{prob} \times \frac{E_{rest} \text{ current node in } j}{\sum_{i=1}^{M_{\text{neighbour nodes in } j}} E_{i}} \times \frac{E_{rest}}{\sum_{i=1}^{N} (E_i - M)^2} \times \min\left(\frac{\text{node degree}}{D_{\text{avg}}}, 1\right)$$

In Eq. (2), the three parts of the previous formula are repeated, except that in this step, it is necessary for specification of the protocol of choice at the node and its selection as a CH to consider its position and the density of the surrounding nodes. In the third part of the formula, nodes with more sparse surroundings are selected as CHs, and their data need to be transferred. $D_{\text{avg}}$ is the average density of the surrounding nodes, calculated as in Eq. (3).

$$D_{\text{avg}} = \pi R^2 \times \frac{\text{NumDevices}}{\text{Area}}$$

In Eq. (3), $R$ is the communication radius of the node, $\text{NumDevices}$ is the number of battery-powered devices in network, and $\text{Area}$ is the operation environment of sensors and devices. The value is 1 when the node can be elected as the final CH. Where the node is subjected to calculation, there are three possible cases, as given in Eq. (4).

$$C_1 : \text{Random}(0, 1) \leq C_{prob}$$

$$C_2 : \sum_{\text{neighbour node in } j} E_{i} \geq 1$$

$$C_3 : \frac{\text{Node Degree}}{D_{\text{avg}}} \geq 1$$

If two of the three conditions in Eq. (4) are met, the node will be selected as the final node. If the node in question is unable to communicate with any of the nodes, it prepares itself for the transfer as the final CH without having to participate in the competition. Table II summarizes represents a truth table of possible scenarios.

| Type | Parameters | Values |
|------|------------|--------|
| Area | Network Area | From (0,0) - (100, 100) |
| Number of nodes | N | 300 |
| w7 Initial energy | Energy | Heterogeneous |
| Transmission range | R | various |
| Gateway | Sink | At (50, 175) |
| Default cluster radius | 25m | |
| Data packet size | 100 bytes | |
| Broadcast packet size | 25 bytes | |
| Data header size | 25 bytes | |
| Each round | CH data compress rate | 5 TDMA frames |
| Duration | 1000 rounds | |
| Interfaces | NAI1-NAI5 | |

V. PERFORMANCE EVALUATION

In recent years, several works aim to simulate various layers of cloud computing stack: IoT [20], [21] edge [22], [23], fog [24], and cloud-based scenarios [25] by considering different levels of system details and complexities. However, in the context of clustering algorithms, the majority of papers used MATLAB as their base simulation platform due to the statistical analysis nature of their evaluations. For the same reason, to evaluate our method we also used MATLAB as the base simulation platform and all conditions, including network size, simulation area, and number of nodes, as well as network distribution are assumed to be the same in all compared algorithms. The average of ten runs in the simulation process was calculated to obtain the final results.

In this study, we named the result of the proposed algorithm HetEng. It was compared statistically with different algorithms, including HEED, LEACH, BEE, BEEM, and Smart-EEEM. As initial energy levels were generated randomly in the simulation environment, the running phase was assumed to be the same in all the algorithms. For large areas, the proposed method exhibited an increase in...
network lifetime as compared to the other algorithms. It was also observed that network lifespan was generally longer in networks with more nodes than in ones with fewer nodes. The results indicated an increase in the number of alive nodes and residual energy in the proposed algorithm as compared to the others. In the large networks, however, network coverage was much closer to network lifetime, because the energy consumption in the network was significantly higher.

In Fig. 2, the superiority of the proposed algorithm over all the compared algorithms can be observed. On that basis, an average 3% improvement over the Smart-BEEM algorithm has occurred in the residual energy of the nodes, and the slope is closer to linear than those of the other algorithms, which has resulted from the closer-to-normal distribution of energy consumption. Moreover, the HEED and LEACH algorithms exhibited steeper slopes, in that order, which indicated lower amounts of energy remaining after 1000 rounds.

As indicated in Fig. 4 in term of network coverage, the proposed algorithm exhibited a slight improvement in some scenarios as it depends on different factors e.g. nodes distribution and their initial energy state in the scenario, however, its performance was equal to that of Smart-BEEM and SelfCon on average, resulting from the use of five communication protocols in all the compared algorithms, which greatly affects network coverage. The diagram exhibited more normal slopes for Smart-BEEM, SelfCon, and HetEng than other compared algorithms.

Fig. 5 demonstrates the probability of selecting a CH among the nodes. As can be seen, the HEED algorithm exhibited the largest number of jumps and changes at the beginning of the competition, and the probability constantly decreased over time with the loss of alive nodes. The five algorithms BEE, BEEM, Smart-BEEM, SelfCon, and HetEng
were closer to linear, and exhibited almost equal values of probability of the role.

Fig. 6 shows the importance of reducing the number of iterations in the selection of the CH in the network and its impact on reduction of network overhead. As indicated, the HEED algorithm exhibited the largest number of rounds and of fluctuations due to the greedy mechanism of selecting the cluster head and the focus on distribution of the role among all the nodes (including those with low energy). HetEng reduced the iteration number by 1% on average in comparison with Smart-BEEM and SelfCon on average. It should be noted that the Leach algorithm exhibited the smallest number of rounds in the selection of the CH in the network and its impact on reduction of network overhead. As indicated, the LEACH algorithm exhibited the smallest number of rounds in the selection of the CH in the network and its impact on reduction of network overhead.

**VI. CONCLUSION AND FUTURE WORKS**

In this paper, we proposed a distributed energy-efficient cluster head (CH) selection scheme in low-power clustered IoT networks to support heterogeneity and detect the most optimum nodes with high energy to play the CH role. Despite the many advancements that have been made in the field of hardware and downsizing technologies, energy consumption, network longevity, and, most importantly, network coverage are considered as the main challenges in such networks. Many of the challenges that exist in all low-power networks, such as wireless sensor networks, are still there in IoT networks, where they are even severer due to the much more complex scenarios. In the proposed algorithm, real energy variance was calculated using sample deviation, and the nodes were compared to those surrounding them in the cluster, indicating a high level of intelligence, which makes it possible to manage the network in terms of longevity, energy consumption, and network coverage under different conditions. We reduced the number of iterations by 1%, and used average values to distribute energy consumption among high-energy nodes, which resulted in: 1- reduction of energy consumption by 6.6% over the whole network rather than at high-energy nodes, which resulted in: 2- reduction of energy consumption by 6.6% over the whole network rather than at

of rapid depletion of energy at nodes with low energy. Finally, there are issues that need to be addressed for improvement of the present study in more realistic environments: (1) mobility, which is an important factor in IoT networks, and (2) Quality of Service (QoS).

**ACKNOWLEDGMENT**

The authors would like to express their highest gratitude to Mr. Michel Gokan Khan from Karlstad University in Sweden for his genuine support and constructive feedback.

**REFERENCES**

[1] D. Amiri, A. Anzanpour, I. Azimi, M. Levorato, P. Liljeberg, N. Dutt, and A. M. Rahmani, “Context-aware sensing via dynamic programming for edge-assisted wearable systems,” *ACM Trans. Comput. Healthcare*, vol. 1, no. 2, Mar. 2020. [Online]. Available: https://doi.org/10.1145/3351286

[2] L. Xu, R. Collier, and G. M. P. O’Hare, “A survey of clustering techniques in wsns and consideration of the challenges of applying such to 5g iot scenarios,” *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1229–1249, 2017.

[3] A. Shahraai, A. Taherkordi, Haugen, and F. Eliassen, “A survey and future directions on clustering: From wsns to iot and modern networking paradigms,” *IEEE Transactions on Network and Service Management*, pp. 1–1, 2020.

[4] L. Xu, G. M. P. O’Hare, and R. Collier, “A smart and balanced energy-efficient multihop clustering algorithm (smart-beem) for mimo iot systems in future networks,” *Sensors*, vol. 17, no. 7, 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/7/1574

[5] M. S. Batta, H. Mäbel, Z. Alikout, and S. Harous, “A distributed multihop intra-clustering approach based on neighbors two-hop connectivity for iot networks,” *Sensors*, vol. 21, no. 3, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/3/873

[6] P. Saini and A. K. Sharma, “E-deec- enhanced distributed energy efficient clustering scheme for heterogeneous wsns,” in *2010 First International Conference On Parallel, Distributed and Grid Computing (PDGC 2010)*, 2010, pp. 205–210.

[7] Q. Li, S. P. Gochhayat, M. Conti, and F. Liu, “Energiot: A solution to improve network lifetime of iot devices,” *Pervasive and Mobile Computing*, vol. 42, pp. 124–133, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574119217301299

[8] L. Malambo, S. Popescu, D. Horne, N. Pugh, and W. Rooney, “Automated detection and measurement of individual sorghum panicles using density-based clustering of terrestrial lidar data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 149, pp. 1–13, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271618303502

[9] D. R. Edla, M. C. Kongara, and R. Cherukuv, “See-psb based clustering approach for load balancing of gateways in wireless sensor networks,” *Wirel. Netw.*, vol. 25, no. 3, p. 1067–1081, Apr. 2019. [Online]. Available: https://doi.org/10.1007/s11276-018-1638-2

[10] C. Iwendi, P. K. R. Maddikunta, T. R. Gadekallu, K. Lakshmanana, A. K. Bashir, and M. J. Piran, “A metaheuristic optimization approach for energy efficiency in the iot networks,” *Software: Practice and Experience*, vol. n/a, no. n/a. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2797

[11] S. Gajjar, M. Sarkar, and K. Dasgupta, “Famcrowor: Fuzzy and ant colony optimization based combined mac, routing, and unequal clustering cross-layer protocol for wireless sensor networks,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 149, pp. 1–13, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271618303502

[12] S. Gajjar, M. Sarkar, and K. Dasgupta, “Famcrowor: Fuzzy and ant colony optimization based combined mac, routing, and unequal clustering cross-layer protocol for wireless sensor networks,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 149, pp. 1–13, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271618303502

[13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2000, pp. 10 vol. pp.-

[14] T. M. Behora, U. C. Samal, and S. K. Mohapatra, “Energy-efficient modified leach protocol for iot application,” *IET Wireless Sensor Systems*, vol. 8, no. 5, pp. 223–228, 2018. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-wss.2017.0099
