

The Effect of the Position of BS and the Size of Network on the Decision of Implementing a Centralized or a Distributed Clustering in WSNs

İpek Abasikeleş-Turgut

Abstract—Due to the resource limitations of the sensor nodes, the most important issue in Wireless Sensor Networks (WSNs) is prolonging overall network lifetime by using energy-efficient routing algorithms. Cluster-based routing protocols have proven to be effective in network topology management, energy minimization and data aggregation. Clustering process can be conducted by a distributed or a centralized manner. In this paper, a distributed and a centralized clustering algorithms are scrutinized under different network parameters, including the size of the network and the location of the base station. The results are evaluated by using the death of the first node and the death of the last node since they are the indicators of the quality and the lifetime of the network. The results show that the distributed approach performs better than the centralized equivalent up to 24.5% for small scale networks when the BS is inside the network. However, as the size of the network increases or as the BS moves away from the network, the centralized approach gets the advantage.

Index Terms—Base station, centralized clustering, distributed clustering, network size, wireless sensor networks.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of small sensor nodes that act as information collectors and one or more processing centers connected via wireless links. A sensor node has resource limitations such as the limited transmission capacity, energy/power usage and the memory or storage capability. Among these constraints, the energy/power constraints of sensor nodes are vital problems as in most of the cases the batteries of the sensor nodes are not rechargeable or replaceable. Hence prolonging the lifetime of the network by efficient utilization of sensor nodes has become a major issue in WSNs and the routing protocols in WSNs need to consider this problem [1], [2].

Since traditional routing protocols for WSNs are not optimal in terms of energy consumption, clustering approach takes the place of them as an energy-efficient communication protocol. The objectives of clustering are minimizing the total transmission power aggregated over the nodes in the selected path and balancing the load among the nodes for prolonging the network lifetime [3]. For this purpose, clustering divides larger networks into small manageable clusters and each node is assigned or belongs to a unique cluster, which can

communicate with the base station (BS) via cluster heads (CH) [4].

The formation of clusters can either be distributed or centralized. In a centralized clustering approach, an authority (such as the BS) makes the decision of both the clusters and the cluster heads while in a distributed clustering, all the nodes in the clusters can take the decision of becoming a CH for the current round [5]. There are plenty of studies in literature that use the distributed clustering [6]-[8] or the centralized clustering [9]-[11] approaches. In centralized clustering, the BS utilizes global knowledge of the network to produce better clusters that require less energy for data transmission. Besides, the number of the CHs in each round of a centralized clustering equals a predetermined optimal value, whereas the number of the CHs varies from round to round due to the lack of global coordination among the sensor nodes under distributed clustering [9]. However, if the BS fails in centralized clustering, the entire network will collapse. Hence, there is no guarantee for reliability in centralized clustering mechanism. On the other hand, self-organized sensor nodes cope with this problem and better data collection with minimum redundant information is performed in distributed clustering [12].

Apart from their advantages and disadvantages, the network parameters, including the position of the BS, the number of the sensor nodes and the size of the network area, considerably affect the performance of both the centralized and the distributed clustering approaches. Hence, for making a decision on implementing either a centralized or a distributed clustering, these parameters should also be taken into consideration.

As far as is known, there are four studies in literature, which compare the performance of the centralized and the distributed clustering algorithms under various network parameters. All of them investigate the effect of the position of the BS by using LEACH [6] and LEACH-C [9] algorithms. Twelve different BS positions is used in [13] and [14] for a 100 node network. Another study [15] uses five different BS positions on a 100 node network and briefly analyses the performance of the algorithms. Only one study [16] utilizes additional parameters, including different numbers of the nodes and the clusters. All of these studies use a 100×100 network and none of them researches the effect of the size of the network.

Therefore, in this paper, a detailed comparison between the distributed and the centralized clustering algorithms is performed under different network sizes and BS locations. LEACH [6] routing algorithm, which is the headstone of the

Manuscript received August 10, 2015; revised January 18, 2016. This work was supported by the Scientific and Technological Research Council of Turkey under the project of TÜBİTAK-EEEAG - 115E211.

İpek Abasikeleş-Turgut is with İskenderun Technical University, Hatay, Turkey (e-mail: abasip@gmail.com).

distributed clustering approaches in literature, is used as a distributed clustering algorithm. However, LEACH-C [9] is not utilized as a centralized clustering approach in this paper as opposed to the corresponding studies in literature. The reason is that LEACH uses a randomized CH selection methodology, while LEACH-C elects the CHs according to their energy level. Therefore, for a fair comparison between the centralized and the distributed approaches, the conditions must be formed equally. The centralized approach used in this paper is created by modifying the CH election phase of the LEACH algorithm. Distinctly from the studies in literature, instead of selecting the position of the BS randomly, four basic locations is chosen, including the center of the network, the corner of the network, the nearby place of the network and the far-away from the network. Similarly, the effect of the size of the network, which varies as from the smaller to larger sizes, is also considered differently from the studies in literature. After simulating the system, the performance of the distributed and the centralized clustering approaches according to the number of the living nodes is investigated in an all-inclusive manner.

The rest of this paper is organized as follows. Section II describes the simulation parameters and environment with the clustering network models. In Section III, the simulation results are presented and discussed, while concluding remarks and future extensions are given in Section IV.

II. EVALUATION METHODOLOGY

In this paper, OMNeT++, which is an extensible, modular, component-based C++ simulation library and framework, primarily for building network simulators [17], is used to perform and compare the distributed and the centralized clustering algorithms.

A clustering based network (Fig. 1) has three basic components called the BS (sink), the CH and the member sensor node (MN) [18]. The sink is responsible from evaluating the data, which is gathered from the CHs. The MNs transmit their data to their CHs, while the CH receives data from all the cluster members, performs signal processing functions on the data, i.e. data aggregation, and transmits the data to the remote sink.

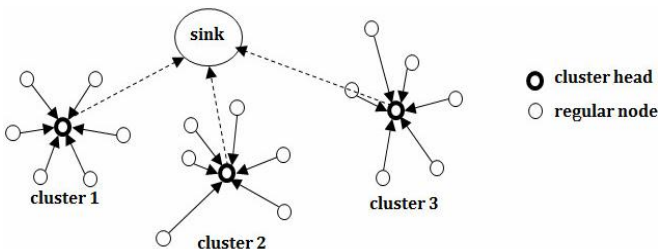


Fig. 1. A cluster-based network model.

Both for the centralized and the distributed clustering approaches used in this paper, the network operation cycle is based upon rounds as LEACH. Each round contains two phases called the set-up and the steady-state. The set-up phase (i.e. first stage) carries out cluster construction, while data transmission is performed in the steady-state phase (i.e. second stage). As is seen in Fig. 2, randomly generated CHs

broadcasts its information into the network for constructing the clusters. Based on the strength of the received signal, a node decides which cluster to join in and sends the message back to the corresponding CH. The CH allocates the communication time slot for each MN in its cluster based on TDMA (Time Division Multiple Access). In steady-state phase, the MNs send their data to their CHs according to the TDMA schedule. After receiving all of the data from their MNs, the CHs aggregate the data and send them to the BS in a single hop manner. In order to minimize the power consumption, the execution time period of the steady state phase is far greater than the set-up phase. The difference between a centralized and a distributed clustering is the decision methodology of the CHs. In the distributed clustering approach, the nodes elect themselves as CHs in a distributed manner independently from the other nodes in the network. However in the centralized clustering, the BS informs the CHs after processing the CH election algorithm in a centralized manner.

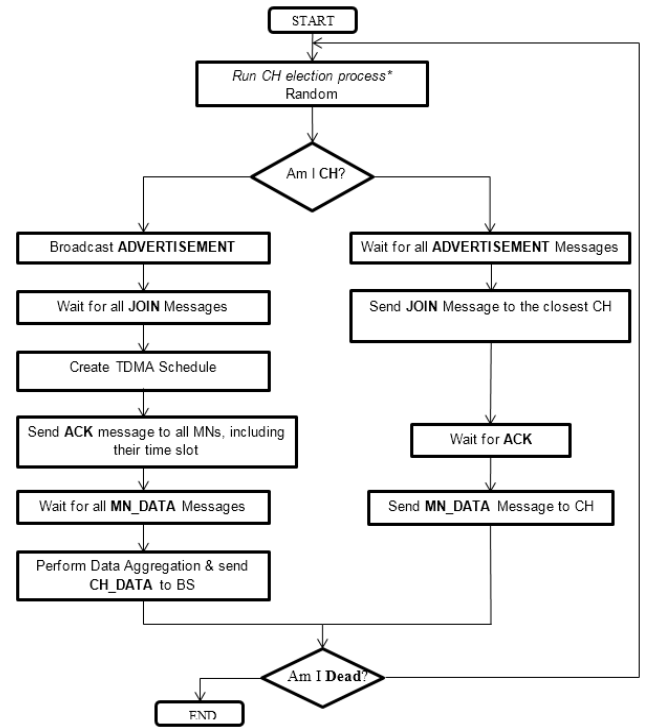


Fig. 2. The flow chart of a clustering routing protocol. * This process is conducted centralized at the BS for the centralized approach, while each sensor node independently decides to become a CH in distributed approach.

The parameters used in the simulations can be seen in Table I. As is seen in Table I, a 100 node with a uniform distribution is spread on a 100 m × 100 m, 200m × 200m and 400m × 400m network. The position of the BS is either within (center and corner) or out of the network (near-by and far-away), as is seen in Fig. 3. “Center” means that the BS is located at the center of the network with a position of $(x/2, y/2)$, where x is the width and y is the height of the network. “Corner” denotes that the BS is at the left corner of the network with a position of $(0, 0)$. Both for “near-by” and “far-away”, the BS is located out of the network with a position of $(0, -(y + 100))$ and $(0, -(y + 200))$, respectively. The energy model used in this paper is the same with LEACH-C [9].

TABLE I: THE SIMULATION PARAMETERS

Parameter	Value
Network Area	100m × 100m, 200m × 200m, 400m × 400m
Number of Nodes	100
Desired percentage of Clusters	5%
Position of BS	Center, Corner, Nearby, Far-away
Initial Energy	0.5 Joule
Node Distribution	Uniform
Broadcast / Control	25 byte
Packet Size	
Data Packet Size	100 byte
E_{elec}	50 e^{-12}
E_{fs}	10 e^{-12}
E_{mp}	0.0013 e^{-12}
Threshold Distance (d_0)	75m

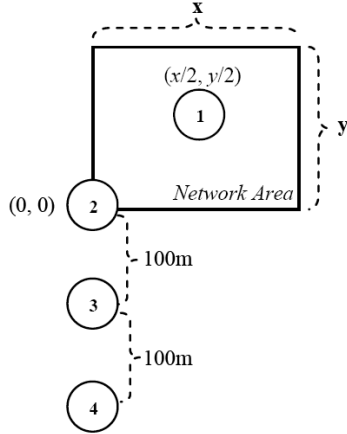


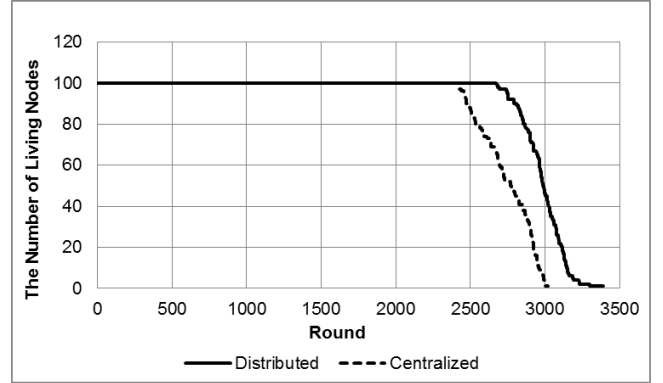
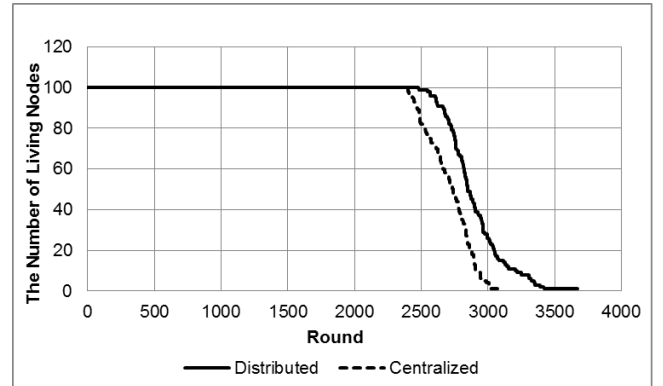
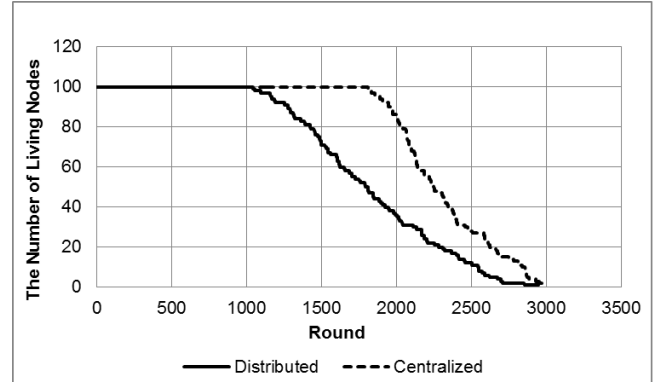
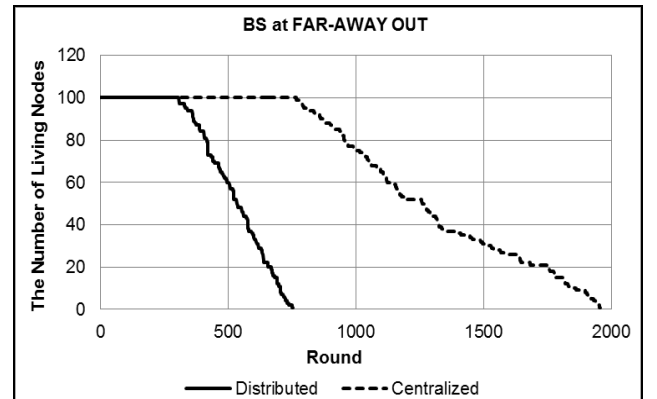
Fig. 3. The position of the BS is within (center (1) and corner (2)) or out of the network (near-by (3) and far-away (4)).

III. RESULTS AND DISCUSSION

The evaluation parameters used to compare the distributed and the centralized approaches are the number of the living nodes in each round, the death of the first and the death of the last node. The number of the living nodes states the number of the nodes, which have enough energy to continue sensing and communication on the following round, i.e. alive nodes. The death of the first node reflects the quality of the sensed data. Until that time, all MNs send their data to BS. Hence, it can be said that sufficient data rate is gained. The death of the last node represents the lifetime of the network. The higher round for the death of the first node and the death of the last node means the higher data rates and the longer network lifetime. Therefore, these parameters show the performance of the algorithms.

The number of the living nodes in each round under a distributed and a centralized clustering algorithms with a BS located at the center of, the corner of, near-by place of and far-away from the network for a 100×100 , 200×200 and 400×400 networks can be seen in Fig. 4 through Fig. 15, respectively.

Both the death of the first node and the death of the last node move to an earlier time as the network size increases regardless of the position of the BS. Communication distance among the system nodes increases as the network size increases. Hence, the nodes spend more transmission energy to communicate with each other and consume their battery earlier. Accordingly, both the network lifetime and the sensed data rate decreases as the network size increases.

Fig. 4. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the center of a 100×100 network.Fig. 5. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the corner of a 100×100 network.Fig. 6. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at near-by place of a 100×100 network.Fig. 7. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at far-away from a 100×100 network.

Since the number of the CHs selected by the BS is on equal with the desired percentage of CHs in the centralized approach, lesser CH than the expected number to be appeared in the network in any round is an impossible situation. Therefore, the load balance provided by the number of the CHs is guaranteed in this approach. However, the BS needs to hear of an extra information about the nodes in the network on the purpose of conducting an accurate decision. Accordingly, extra control messages are needed to carry out this task. The control messages bring about additional communication between the nodes, hence presume on battery and adversely affect the lifetime of the nodes.

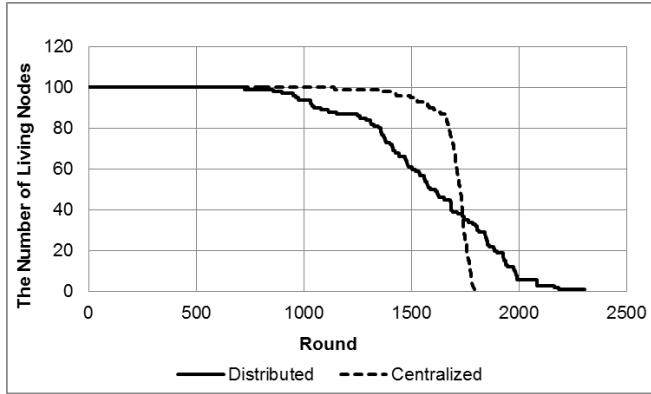


Fig. 8. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the center of a 200×200 network.

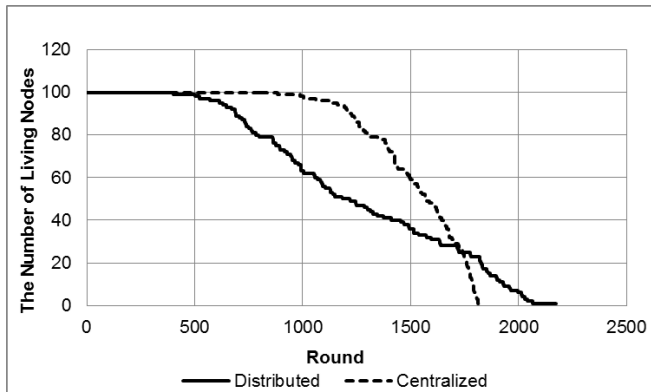


Fig. 9. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the corner of a 200×200 network.

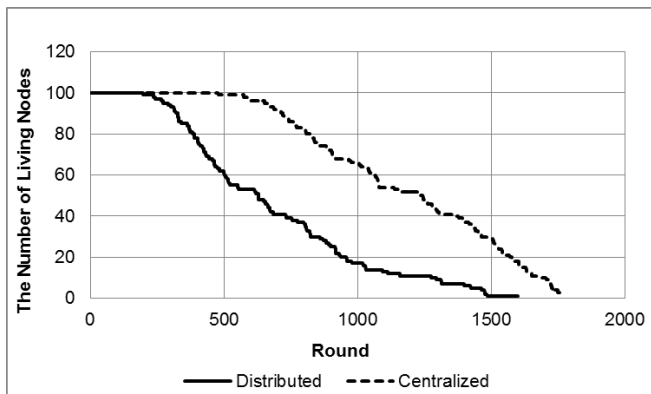


Fig. 10. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at near-by place of a 200×200 network.

The advantage of the distributed approach is that no extra

control message is required. The nodes can make their decisions about being a CH for any round regardless of the other nodes in the network. Although the independency of the nodes provides self-organization and minimum communication cost in the distributed approach, due to the lack of global information, there is no guarantee that the desired percentage of CHs presents in the network. For instance, neither of the nodes can elect themselves as a CH in a round and accordingly all of them have to communicate with the BS directly, which wastes much more energy in the network.

As the BS moves far away from the network, since the communication cost increases, the importance of providing the load balance in the network takes precedence over the additional cost caused by the extra control messages. Accordingly, the centralized approach performs better than the distributed one both for the death of the first and the death of the last node regardless of the network size. The round that the centralized approach commences to surpass the distributed equivalent differs by the size of the network.

As is seen in Fig. 4 through Fig. 7, the position of the BS plays an important role on selecting the best routing approach on a 100×100 network. If the BS locates inside the network boundaries (i.e. either at the center or at the corner), the distributed clustering approach should be chosen for a longer network lifetime with higher data rates.

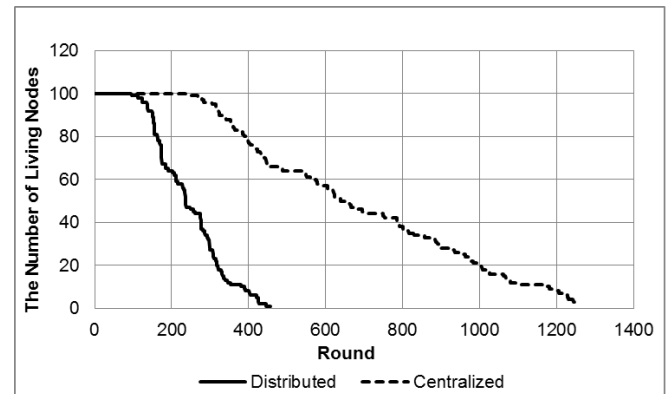


Fig. 11. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at far-away from a 200×200 network.

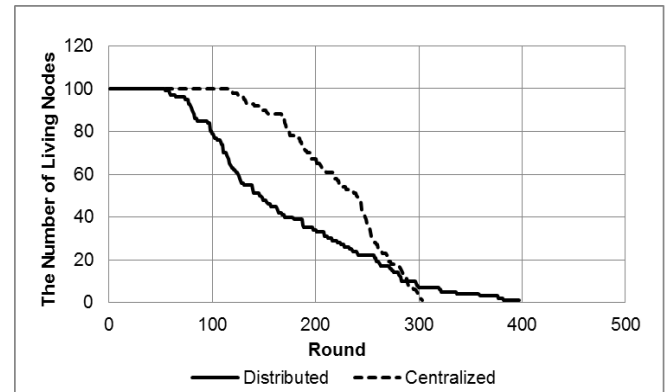


Fig. 12. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the center of a 400×400 network.

The distributed approach performs 24.5% better than the centralized equivalent when the BS is at the center of the 100

$\times 100$ network, while this rate decreases to 14.4% when the BS locates at the corner. As the BS steers away from the network, the centralized approach becomes more favorable and the performance difference between the centralized and the distributed approaches increases as the distance between the BS and the network increases. The centralized routing algorithm yields 24.5% and 49.9% higher performance than the distributed one if the BS is at near-by place of the network and far-away from the network, respectively.

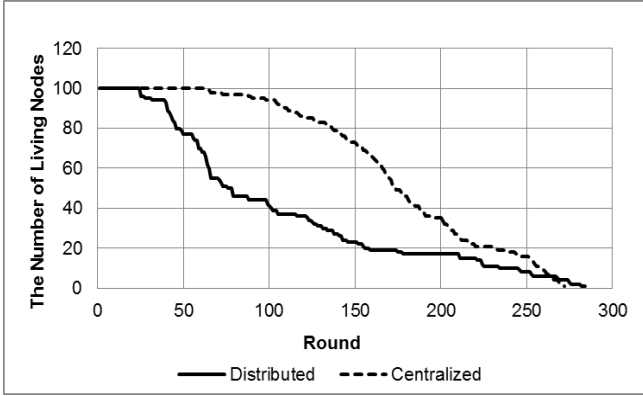


Fig. 13. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at the corner of a 400×400 network.

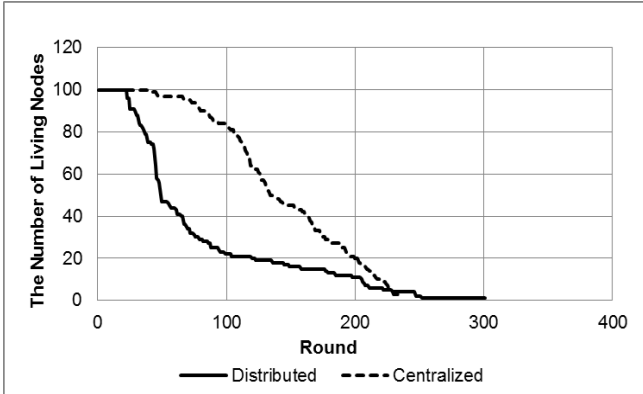


Fig. 14. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at near-by place of a 400×400 network.

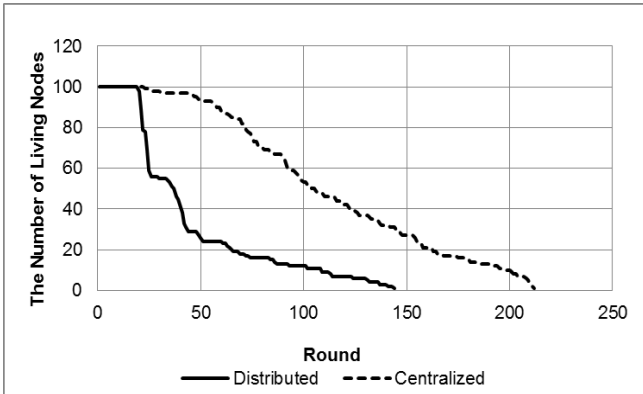


Fig. 15. The number of the living nodes in each round under the distributed and the centralized clustering algorithms with the BS located at far-away from a 400×400 network.

When the BS locates out of the network on a 200×200 network area (Fig. 10 and Fig. 11), the centralized approach performs better as is on a 100×100 network. However, the difference of the performance between the centralized and the

distributed approach increases as the network size increases. For instance, the centralized algorithm yields 32% and 38% better than the distributed algorithm on a 200×200 network where the BS stands near-by place of and far-away from the network, respectively. As is seen in Fig. 8 and Fig. 9, if the BS locates into the network on a 200×200 area, the network displays a different behavior than it is on a 100×100 network. Centralized approach leads the field from the beginning of the simulation until a certain round, which is close by the end of the network lifetime, while the distributed approach comes up from behind for the rest of the rounds. When the BS is at the center (corner) of the network, the centralized approach antecedes the distributed one up to 16.85% (21.51%) until the round of 1737 (1741). However, after these rounds, the distributed approach forges ahead with a rate of 10.6% (7.86%) for the rest of 570 (435) rounds when the BS is at the center (corner). The results for 200×200 network shows that as the size of the networks increases, the centralized approach gets the edge on for all BS locations.

Finally, although the tendency of the algorithms on a 400×400 network are in the same ballpark to that of on a 200×200 network, the results trend in favor of the centralized approach as is seen in Fig. 12 through Fig. 15. If the BS is at the center (corner) of the network, the centralized approach precedes the distributed one 24.4% (29.9%) until the round of 289 (264). However, after these rounds, the distributed approach forges ahead with a rate of 3.81% (2%) for the rest of 108 (20) rounds when the BS is at the center (corner). The results show that as the network size increases the distributed approach loses its advantage and the centralized approach holds the high ground for all situations.

IV. CONCLUSION

In this paper, a distributed and a centralized clustering algorithm are compared under different network parameters and an answer for which approach is better in which situation is searched. For this purpose, a 100 node is uniformly distributed on a network with three different sizes and four different BS locations. After conducting detailed simulations, the performance of the system is examined by using the death of the first and the death of the last node. Since the death of the first node reflects the quality of the sensed data and the death of the last node signifies the lifetime of the network, these parameters exactly show the performance of the algorithms.

If the BS is located inside a small scale network, both the lifetime and the quality of the network is higher under the distributed approach than that of the centralized one due to providing self-organization between the nodes. Regardless from the size of the network, if the BS is detracted from the network, the centralized approach gets the edges on due to balancing the load of the nodes. As the size of the network increases, the centralized approach performs better than the equivalent one for the death of the first node (i.e. the quality of the network), while the distributed approach still holds the high ground for the death of the last node (i.e. the lifetime of the network).

The future work includes the comparison of these approaches when the CH election process is not only based on

stochastic but also different parameters, including the energy level and the location of the nodes.

REFERENCES

- [1] R. N. Enam, S. Misbahuddin, and M. Imam, "Energy efficient round rotation method for a random cluster based WSN," in *Proc. 2012 International Conference on Collaboration Technologies and Systems*, 2012.
- [2] B. J. Lakshmi and M. Neelima, "Maximizing wireless sensor network life time through cluster head selection using Hit sets," *International Journal of Computer Science Issues*, vol. 9, no. 2, 2012.
- [3] S. K. Singh, M. P. Singh, and D. K. Singh, "A survey of energy-efficient hierarchical cluster-based routing in wireless sensor networks," *International Journal of Advanced Networking and Application*, vol. 2, no. 2, pp. 570-580, 2010.
- [4] M. Aslam *et al.*, "HADCC: Hybrid advanced distributed and centralized clustering path planning algorithm for WSNs," in *Proc. 2014 IEEE 28th International Conference on Advanced Information Networking and Applications*, 2014.
- [5] K. Wankhede and S. Sirsikar, "Review of clustering algorithms in wireless sensor networks," *International Journal of Advance Foundation and Research in Computer*, vol. 1, issue 11, November 2014.
- [6] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. the 33rd Annual Hawaii International Conference on System Sciences*, 2000.
- [7] 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.
- [8] P. Ding, J. Holliday, and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks," *Distributed Computing in Sensor Systems*, pp. 322-339, 2005.
- [9] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wirel. Commun.*, vol. 1, pp. 660-670, 2002.
- [10] S. D. Murugunathan *et al.*, "A centralized energy — efficient routing protocol for wireless sensor networks," *Communications Magazine*, vol. 43, no. 3, pp. 8-13, 2005.
- [11] T. Kang *et al.*, "A clustering method for energy efficient routing in wireless sensor networks," in *Proc. the 6th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications*, 2007.
- [12] S. R. B. Prabhu and S. Sophia, "A survey of adaptive distributed clustering algorithms for wireless sensor networks," *International Journal of Computer Science and Engineering Survey*, vol. 2, no. 4, pp. 165-176, 2011.
- [13] X. H. Wu and S. Wang, "Performance comparison of LEACH and LEACH-C protocols by NS2," in *Proc. 9th International Symposium on Distributed Computing and Applications to Business, Engineering and Science*, 2010, pp. 254-258.
- [14] K. N. A. Kumar, B. H. C. Kumar, B. A. Mohan, and H. Sarojadevi, "The study of distributed and centralized cluster formation protocol in WSNs," *International Journal of Science and Applied Information Technology*, vol. 3, no. 3, pp. 8-12, 2014.
- [15] P. Nayak, "Comparison of routing protocols in WSN using NetSim simulator: LEACH vs. LEACH-C," *International Journal of Computer Applications*, vol. 106, no. 11, 2014.
- [16] V. A. Geetha. V. Pranesh, and S. Kallapur, "Clustering in wireless sensor networks: Performance comparison of LEACH & LEACH-C protocols using NS2," *Procedia Technology*, vol. 4, pp. 163-170, 2012.
- [17] A. Varga, "The OMNeT++ discrete event simulation system," in *Proc. European Simulation Multi-conference (ESM'2001)*, Prague, Czech Republic, 2001.
- [18] I. Abasikeleş-Turgut, M. N. Aydın, and K. Tohma, "A realistic modelling of the sinkhole and the blackhole attacks in cluster-based WSNs," in *Proc. 2nd International Conference on Electrical and Electronics Engineering*, 2015.



İpek Abasikeleş-Turgut has completed her university education at the Department of Computer Engineering of Istanbul Technical University (ITU) with a degree of *magna cum laude* in 2007. She has received her MSc in computer engineering and her PhD in electrical and electronics engineering from Çukurova University in 2009 and 2013, respectively.

She currently has published 2 books, 4 SCI journal papers, 4 national and 5 international conference papers. Her interest areas are routing and security issues in wireless sensor networks, parallel processing, simulation and modeling of multiprocessors with distributed shared memory.