

Adaptive Simulated Evolution based Approach for Cluster Optimization in Wireless Sensor Networks

Abdulaziz Alsayyari

Computer Engineering Department
Shaqra University
Dawadmi, Ar Riyadh, Saudi Arabia

Abstract—Energy consumption minimization is crucial for the constrained sensors in wireless sensor networks (WSNs). Partitioning WSNs into optimal set of clusters is a promising technique utilized to minimize energy consumption and to increase the lifetime of the network. However, optimizing the network into optimal set of clusters is a non-polynomial (NP) hard problem, and the time needed to solve such problem increases exponentially as the number of sensors increases. In this paper, simulated evolution (SimE) algorithm is engineered to tackle the problem of cluster optimization in WSNs. A goodness measure is developed to measure the accuracy of assigning nodes to clusters and to evaluate the clustering quality of the overall network. SimE was developed such that the number of clusters and cluster heads are adaptive to number of alive nodes in the network. In fact, extensive simulation results demonstrate that SimE provides near optimal clustering and improves the lifetime of the network by about 21% compared to the traditional LEACH-C protocol.

Keywords—Clustering algorithm; cluster optimization; network lifetime; simulated evolution; wireless sensor networks

I. INTRODUCTION

Wireless sensor networks (WSNs) are formed using small sensor nodes to monitor certain phenomena of environments where human presence may be impossible or not preferred. After wireless nodes are deployed and connected together, data about sensed events is typically gathered and reported to a centralized location for further processing [1], [2]. Nevertheless, the applications of WSNs are wide and vary from one application to another [3]. The application often customizes the details of designing wireless sensor nodes and WSNs' planning; including node architecture, communication protocols, network topology, and deployment schemes [4]. In large-scale deployment scenarios of WSN such as battlefields and forest habitat monitoring, sensor nodes often have limited resources. This is because batteries in such deployment scenarios are mostly neither changeable nor chargeable. As a result, batteries of sensor nodes are considered a sacred resource [5]. Therefore, minimizing energy consumption is necessary to increase the life time of the WSN.

In WSNs, data is exchanged between sensor nodes in an ad hoc fashion. This technique allows the network to cover larger geographical areas, extend the reach of the network, and help sensor node in saving energy by lowering transmission power of the node and allowing neighboring nodes to perform certain network duties alternately [6]. In fact, clustering is a popular method commonly utilized in WSN to prolong network

lifetime [7]. Furthermore, efficient clustering directly leads to energy saving and, hence, results in extending network lifetime [8], [9]. Clustering is achieved by grouping a specific set of sensor nodes in one cluster and, then, assigning a cluster head (CH) to handle certain tasks in the cluster. Typically, nodes are selected in one cluster according to criteria such as cluster size and nodes' locations. In such scenario, nodes in the cluster communicate with the cluster head instead of directly communicating with the base station (BS). Later, the cluster head aggregates packets received from cluster nodes and sends them to a BS.

In this work, a simulated evolution (SimE) algorithm for cluster optimization in WSNs to provide near optimal solutions is presented. More specifically,

- A simulated evolution (SimE) algorithm is developed to cluster the WSN and increase its lifetime. The results show that the proposed SimE algorithm minimizes energy consumption in the network. This is achieved by minimizing the total sum of squared distances between cluster nodes and the CHs.
- A goodness measure, which is the core part of the SimE algorithm, is proposed to tackle the WSN clustering problem and to evaluate the quality of the produced clusters.
- Unlike previous methods, the adaptivity of number of clusters (or CHs) to the network size is also addressed, where it is shown that number of clusters is adaptive to number of alive nodes in the network and a clustering algorithm should be adaptive to number of alive nodes per round instead of assuming a fixed number of clusters. The adaptivity of the proposed SimE approach eliminates the need to develop a multi-objective optimization function to account for load balancing of the clusters.
- This paper investigates the effect of BS location and deployment area on the network lifetime and addresses the change in number of clusters as deployment area changes.
- The simulation results show that the proposed SimE approach enhances the network lifetime by about 21% compared to the LEACH-C protocol.

The remainder of this paper is organized in the following order. While Section II provides a background study on the

research literature in relevance to cluster optimization in WSNs, Section III presents an overview of the proposed SimE method, including the algorithm and the goodness measure details. Meanwhile, Section IV provides the performance results and finding of the proposed SimE approach. Finally, Section V concludes the work.

II. RELATED WORK

K-means algorithm is a popular approach utilized in WSNs among many other applications to produce clusters [10], [11]. In fact, various approaches that are based on such algorithm have been developed to ensure a more efficient clustering [12], [13]. Such development efforts are continuously attempted as a result of inherent challenges that exist in the WSN clustering scheme. The objective of this problem is to find k optimal clusters such that the total energy is minimized and the lifetime of the WSN is increased. The nodes of the network are grouped (clustered), where they are either member nodes or CH node. Member nodes send to the CH instead of sending directly to the BS. This allows for a reduction in the communication distance and an increase of the lifetime of the network [8]. In general, the number of clusters and CHs are not previously known. Therefore, this number might change over time due to the complete energy loss in some nodes in the network, which further complicates the problem of finding optimal clustering using k-means algorithm. In fact, the problem of k optimal cluster optimization in WSNs was proven to be non-deterministic polynomial-time hard (NP-hard) problem [14].

Many evolutionary approaches and protocols targeting cluster optimization were proposed for WSNs. One of the most well-known approaches is the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol, which is a distributed clustering algorithm [15]. In the LEACH protocol, CHs are randomly selected. Then, they advertise their presence by utilizing the Carrier Sense Multiple Access (CSMA), which is a Medium Access Control (MAC) protocol. Cluster members (CMs) that have not been selected as CHs choose the corresponding CH based on the received signal strength (RSS). In fact, they send their packets to the corresponding CH instead of sending them to the base station (BS) to reduce the energy consumption of the CMs. The CH, however, aggregates the received packets into a single message and forwards it to the BS using spreading codes and CSMA/MAC protocol.

It was shown in [8] that random selection of CHs using decentralized approaches as in LEACH is not efficient in terms of energy consumption. It was also shown that using centralized approach increase the lifetime of the network since it is possible to rotate the selection of the CHs in each round. Furthermore, number of CHs is proportional to the network energy consumption, which directly affects the network lifetime [16]. In fact, the study in [8] proposed a revised version of LEACH called centralized LEACH (LEACH-C) protocol. Generally, the BS centrally configures the clusters according to the communication distance and the energy levels obtained from network's nodes. In LEACH-C protocol, however, the simulated annealing (SA) algorithm is utilized to configure clusters [17]. To balance the energy consumption in each node, only nodes that have energy levels greater than the

average energy of the nodes are nominated to be CHs. The BS runs SA to form the clusters by utilizing the nominated CHs.

A genetic based approach was proposed in [18] and several factors affecting the optimization of the clusters such as the BS location were discussed. In fact, the number of clusters is not adaptive, which may cause uneven number of nodes in the clusters, and it thus was assumed that number of CHs is 10%. The results in [18], however, showed that as number of nodes doubles, the population size needs to be doubled as well for the purpose of maintaining comparable performance. 80% reduction in the distance, on average, was achieved compared to the distance obtained by direct transmission. Several studies discussed QoS routing in WSNs including [19], [20], [21]. The study in [19] presented a multi-objective genetic algorithm for efficient QoS routing in two tiered WSNs. Three fitness functions were introduced to form the multi-objective function of the genetic algorithm, which are energy consumption, delay and reliability. Additionally, it was shown that genetic algorithm is reliable in optimizing these functions including QoS in WSNs. However, performance results of the proposed genetic algorithm in terms of number of alive nodes per round were not reported.

Nevertheless, several studies utilized the particle swarm optimization (PSO) algorithm for cluster optimization in WSNs; for example, see [16] and [22]. In [22], however, the objective function is formed from the Euclidean distance of nodes and the energy consumption of nodes in each round. A constant was utilized to weigh these functions to form the multi-objective function. The proposed PSO approach was compared to LEACH-C to demonstrate the effectiveness of PSO algorithm. Though, unequal initial energy for the nodes and a fixed (not adaptive) number of CHs were assumed. Furthermore, the study in [16] utilized PSO for cluster optimization in WSNs. In this study, the PSO algorithm aims to minimize energy consumption by minimizing number of active CHs. However, minimizing CHs is not always the most appropriate strategy for minimizing the energy consumption as was explained in [8].

Furthermore, the study presented in [23] provided WSNs clustering algorithms based on simulated annealing (SA) and PSO algorithms. Their approach was presented to provide better clustering when compared to LEACH protocol. However, their objective function is actually a multi-objective, which allows less flexibility for energy load distribution among the clusters and that number of clusters in their approaches is fixed. Meanwhile, Tabu search based centralized approach was proposed in [24] for cluster optimization in WSNs. The nodes and the connection between them were represented as a hypergraph, which is a graph with edges having multiple nodes. This approach initially represents the cluster nodes and their CH as a Clique and apply Tabu search to optimize the Clique problem. Although the authors showed that their proposed Tabu search outperforms SA algorithm, the runtime of their proposed approach is higher than SA and it requires addressing many complicated Tabu search related structures such as short, medium, and long-term memories. Recently, the new nature inspired Cuckoo search algorithm was applied for cluster optimization in WSNs, for example, see [25]. In this study, the approaches aims to optimize randomly created

clusters. However, no description of how the Cuckoo search was applied to the clustering optimization, which is discrete in nature, giving that cuckoo search is mainly developed for continuous objective functions.

Nevertheless, many other proposals attempted to solve the problem of clustering the WSN as a pure clustering problem, for example, see [14], [26] and [27]. For this to work, fixed cluster need to be assumed, which is not suitable for the problem of clustering WSNs due to the change in number of nodes in the network over time. In addition, pure clustering might perform worst when the number of nodes decreases in the network due to the complete loss of energy.

III. DESCRIPTION OF THE PROPOSED APPROACH

In this section, a full description of the proposed SimE approach is provided, which is being utilized to optimize number of clusters in the network. This includes describing the SimE algorithm, the proposed goodness measure, and how SimE is engineered for cluster optimization in WSNs.

A. Assumptions

In this work, the following are assumed (no assumption about network density is made):

- The BS has unconstrained power source.
- Each sensor node belongs to exactly one cluster.
- The sensor nodes are static given that in the majority of applications sensor nodes have no mobility.
- Initially all sensor nodes are charged with the same amount of energy.
- Communication links are bidirectional.
- The computation and communication capabilities are the same for all network nodes.
- The only source of energy in sensor nodes is the battery.
- The sensor nodes are unaware of their location. Most of the contributions found in the literature assumed that the sensors can determine their location by means of the Global Positioning System (GPS), which is an unrealistic assumption. This paper, however, adopts the approach described in [16], which assumes that each sensor maintains a list of its neighbors. In that work, the flooding method is utilized to send the list to the BS, where it can decide which nodes will be CHs based on the information received.

B. SimE Description

SimE is a very attractive and elegant evolutionary iterative algorithm that is being utilized over the years to solve various types of optimization problems. By employing SimE, the search space is traversed in a smarter way using smart moves, which makes it outperform other iterative algorithms for most different problems. The evolution of the SimE is as follows: first ill-assigned nodes are determined, and they become candidates for moving them to a better cluster. With iterations, the quality of the solution is improved as the ill-assigned nodes either decrease in number or placed in the best possible cluster.

Therefore, unlike other iterative algorithms such as genetic algorithms, SA and PSO, the evolution of SimE with iterations is smarter and more efficient.

Typically, the SimE algorithm consists of three main steps that are executed in sequence; the evaluation, selection and allocation steps as described in Fig. 1 [28]. In the evaluation step, the nodes are evaluated based on the goodness of each node in the cluster solution and ill-assigned nodes are marked to be considered for movement. Note that in order for the SimE to escape local minima, some nodes that are good might be chosen based on some random parameter. The selection step performs this and also puts nodes to be moved in a selection list PS as in Fig. 1. The allocation step allocates the selected nodes to the clusters based on checking best cluster of the current solution.

In [28], the SimE was selected for cluster optimization in WSNs because it is believed that it is naturally more suitable for cluster optimization in WSNs. This is believed because of the following reasons. Firstly, the nature of WSNs clustering depends on choosing CHs, which are not necessarily the same each round. And, there are certain nodes that can join/leave certain cluster at certain round. Secondly, nodes might die or completely lose energy including CHs, which is not a problem for SimE, given that these nodes simply can be discarded from the clusters without affecting other clusters. This also suggests that the SimE is adaptive in this regard. For example, number of clusters could be reduced at certain round due to the complete loss of energy of all its nodes. Unlike other protocols and heuristics, the SimE itself determines, given an upper bound, the best number of clusters at each round.

```
ALGORITHM Simulated Evolution(N,  $\emptyset_{initial}$ );  
/* B: Selection bias; */  
/*  $n_i$ : node i; */  
/*  $g_i$ : goodness of node  $n_i$ ; */  
/*  $\emptyset$ : Complete Solution; */  
INITIALIZATION;  
Repeat  
EVALUATION:  
ForEach  $n_i \in \emptyset$  Do Evaluate  $g_i$ , EndForEach;  
SELECTION:  
ForEach  $n_i \in \emptyset$  Do  
If Selection( $n_i, B$ ) Then PS = PS  $\cup$  { $n_i$ }  
EndIf;  
EndForEach;  
Sort the elements of PS;  
ALLOCATION:  
ForEach  $n_i \in PS$  Do Allocation( $n_i, \emptyset$ ) EndForEach;  
Until Stopping-criteria are met;  
Return (BestSolution);  
End Simulated Evolution.
```

Fig. 1. Simulated evolution structure [25].

In general, the main component of the algorithm is the goodness measure. Such measure requires to be carefully developed in order to get a good quality final solution produced by the SimE. The goodness value indicates how well a certain cluster node is currently assigned. In such case, the

higher the value of the goodness provided, the lower the probability of the node being selected for reallocation is. In fact, allocation is the most important step in SimE algorithm and has the most impact on the quality of the produced solution. The selection set PS and the partial solution \emptyset_i are the inputs of the allocation operator. A new complete solution \emptyset_N is generated according to an allocation function, which depends on the optimization problem being solved and generally allocates the elements in the PS. The importance of the allocation step comes from the iterative improvement, where previous solution is improved as PS elements are being assigned to a better cluster, without being too greedy.

C. Goodness Evaluation

The idea of the presented goodness measure is to utilize the fact that a node is considered for moving it from the current cluster to another cluster if its goodness value in the current cluster is low. To determine the goodness of a node in a cluster, one must find the total cost of the cluster nodes when a direct communication is made between them and BS. Then, a calculation should be performed of the total cost when one of the cluster nodes is randomly selected as a CH. This represents the goodness of the cluster. To find the goodness of a node in the cluster, the distance from the node to the nodes in the cluster is divided by the total cost from the cluster nodes to the BS. The lower the goodness value of the node is, the higher probability it is to move the node from the cluster to another cluster. To illustrate this, consider the example in Fig. 2 for one cluster having four nodes and a BS, assuming node c is the CH.

In the above case, the goodness of node a (gda) will be:

$$gda = 1 - ((d0 + d1 + d2 + d8) / (d6 + d7 + d8 + d9)).$$

$$gdb = 1 - ((d1 + d3 + d5 + d8) / (d6 + d7 + d8 + d9))$$

$$gdc = 1 - ((d2 + d3 + d4 + d8) / (d6 + d7 + d8 + d9))$$

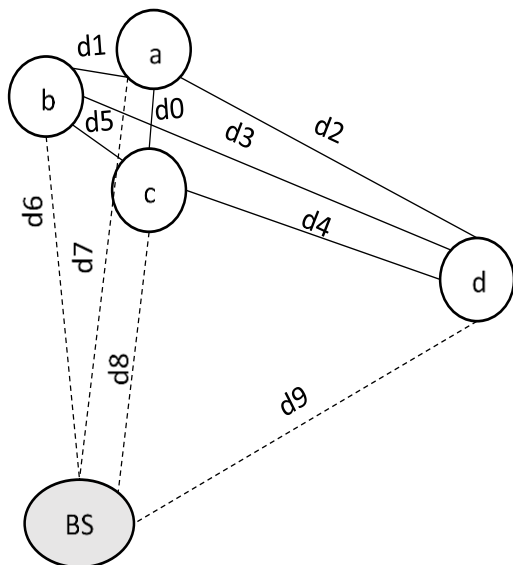


Fig. 2. Illustration of the goodness measure.

The goodness gdc is less than gda and gdb and, hence, node d is less likely to be part of the cluster and will be considered for movement to another cluster. However, node d

might not move into another cluster if gdc is the best possible goodness for all clusters. This illustrates the need for making number of clusters adaptive to maximize the lifetime of the network. The goodness of the assumed CH c is calculated in the same way and might move to another cluster, where another cluster node will act as a CH. Therefore, CH selection is not important in the presented algorithm as any one of the cluster nodes can act as a CH with some very little increased energy, which will be discussed in more details in the results section.

IV. SIMULATION RESULTS AND ANALYSIS

This section explains the simulation implementation, illustrates the results, and provides an analysis of the obtained results.

A. SimE Performance

SimE was implemented in C++ and a random initial network was created. It is worthwhile to mention that the BS is responsible about running the SimE and producing the clusters. Hence, no extra work is needed by the CHs except the aggregation of the data collected from the CMs. Nodes were also deployed randomly. For the simulation, a laptop with Intel i5 processor, 8G memory, and 750G of physical memory was utilized. To demonstrate the output of the presented SimE implementation, Fig. 3 shows an example of clustering 100 nodes in an area of 50X50 m². The BS was placed at (50,175). The upper bound of number of clusters was set to 100 and number in the figure represents the cluster that the node in that location is assigned to.

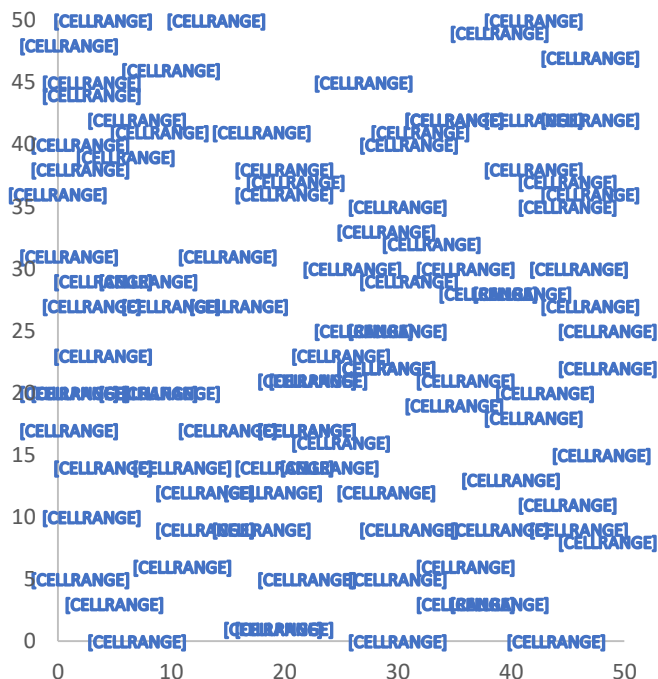


Fig. 3. The resulting clustering of SimE for 100 nodes in a 50X50 m² area.

After many experiments, B value was determined to be 0.1 and a balance parameter was introduced to balance the energy load among the clusters. If a node is considered for allocation

in a cluster x and number of nodes in x exceeded the limit, i.e., $(\# \text{ of nodes in the network} / \# \text{ of clusters}) + \text{balance}$, the allocation of the node to the cluster x will be discarded and the node will be allocated to another cluster.

To evaluate the quality of the solution produced by the SimE and to test the goodness measure being proposed, many experiments were carried out. Considering a network deployed in a $100 \times 100 \text{ m}^2$ area and a BS located at $(50, 175)$, Fig. 4 depicts the overall average goodness over 1000 iterations. The simulation is repeated 100 times and the average goodness was taken. As can be seen from the figure, the goodness initially starts at low value which indicates that the initially random clustering is not good and start improving till it reaches above 0.9 and starts the hill climbing process trying to escape local minima. The maximum average goodness is 0.97. As expected, selection list PS is behaving the opposite of the goodness and is generally decreasing through iterations as in Fig. 5.

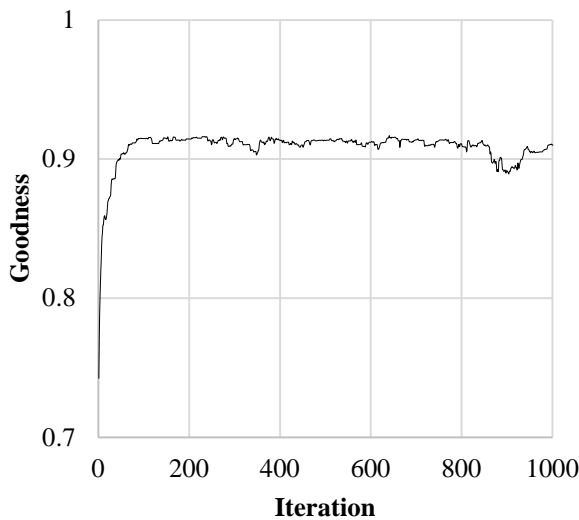


Fig. 4. The behavior of the average goodness over iterations.

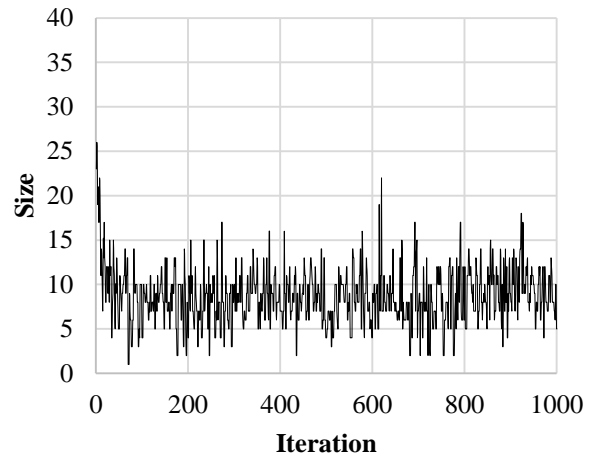


Fig. 5. Selection list size over iterations.

Table I summarizes the results obtained by the SimE for an area of $100 \times 100 \text{ m}^2$. The table shows the goodness, number of iterations taken till no improvement in the solution, the percentage of reduction in the distance compared to direct communication between the nodes and the BS, the runtime and number of clusters produced. The table was produced based on the average of 100 runs. The goodness increases when number of nodes increases and also the reduction in the distance tends to increase as number of nodes increases. These results suggest that the upper bound of number of clusters should be higher when number of nodes increases because the cluster coefficient will be higher in smaller area. The results also demonstrate that the location of the BS plays an important role in the optimization of the clusters. This is due to the fact that the distance of some nodes from the BS is less than the distance to any other node in the network and, therefore, it is concluded that when optimizing and forming clusters, the BS should be far from the area of the deployment to obtain good clustering.

Further, to study the relationship between the cluster optimization and the area of deployment, simulations for an area of $200 \times 200 \text{ m}^2$ were also carried out. The simulation results of this experiment are shown in Table II.

TABLE I. SUMMARY OF PERFORMANCE FOR AREA OF $100 \times 100 \text{ M}^2$

Area	100X100											
	500,500				200,200				100,100			
Node size	100	200	400	800	100	200	400	800	100	200	400	800
Goodness	.98	.98	.98	.98	.95	.95	.95	.94	.79	.74	.85	.87
Iterations	625	766	771	952	330	506	847	960	22	26	33	99
Distance reduction	81.3	88.2	93	95.3	78.5	84.5	88.4	90.8	71.2	73.1	77.9	80
Time (sec)	.39	.71	2.6	7.74	.34	0.64	3.86	5.85	.07	.15	.2	.45
# clusters	17	20	20	20	16	20	20	20	14	18	40	77

TABLE II. SUMMARY OF PERFORMANCE FOR AREA OF 200X200 M²

Area	200X200											
BS location	500,500				400,400				300,300			
Node size	100	200	400	800	100	200	400	800	100	200	400	800
Goodness	.93	.94	.96	0.97	.93	.95	.96	.97	.92	.92	.94	.96
Iterations	174	226	314	400	148	217	321	409	166	169	296	366
Distance reduction	86.4	88.1	90.4	92	82.98	84.9	86.7	88.1	75.8	82.9	85.2	86.4
Time (sec)	.15	1.55	.94	3.35	.12	.61	2.71	6.14	.11	.32	1.5	7.81
# clusters	5	10	20	40	10	18	38	75	18	19	38	78

Since the area size is 200X200 m², the BS was placed outside the deployment area by selecting 300X300 m² and 400X400 m² locations. There is a fluctuation in the distance reduction and the goodness since the network is deployed randomly. Furthermore, no assumptions were made regarding the clustering coefficient, in which the nodes tend to be groped to some extent while randomly deploying the network. For this reason, clustering coefficient will be lower for larger deployment area. Therefore, it is better to increase the upper bound of number of clusters for larger area.

Fig. 6, however, depicts the relationship between the deployment area, the distance reduction, and number of clusters for 100 nodes and BS located at (250,250). As the deployment area increases, number of clusters decreases and the distance reduction increases until some point. The number of clusters decreases as the area increases. This finding reflects the fact that the number of clusters should not be assumed fixed at each round as most of contributions in the literature assumed, i.e., 5%, while it needs to be adaptive in order to account for the deployment area size. Moreover, as the deployment area increases, the parentage of distance reduction tends to decrease. The reason of this is that the distances between the CHS and the BS increases, which does not mean a low quality clustering. On the other hand, number of clusters is also affected by the location of the BS, which also demonstrates the need for making number of clusters adaptive for better clustering results.

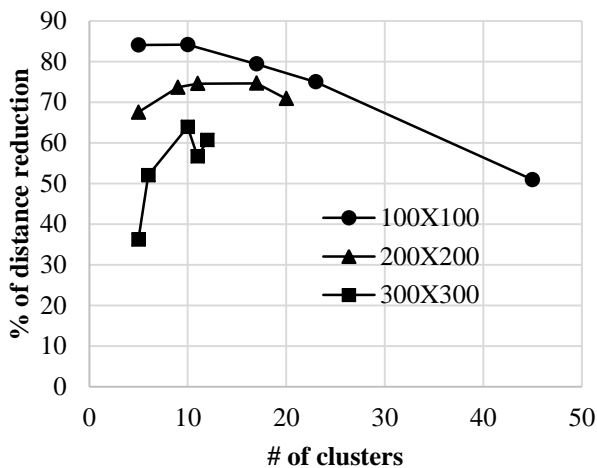


Fig. 6. The effect of the area size on the percentage of the distance reduction and number of clusters.

B. Network Lifetime

To investigate the network lifetime using the proposed SimE approach, experiments were also conducted in C++. The energy consumption model is assumed to be the same as in [8] and [29]. The energy consumed to transmit (ETx) and receive (ERx) 1 bits of packet over a distance d in radio hardware can be written as in (1) and (2), respectively.

$$ETx(l, d) = \begin{cases} 1 Eelc + l efs d^2, & d < d_0 \\ 1 Eelc + l emp d^4, & d \geq d_0 \end{cases} \quad (1)$$

$$ERx(l) = 1 Eelc \quad (2)$$

$$d_0 = \sqrt{\frac{efs}{emp}} \quad (3)$$

Where efs and emp are factors of energy dissipation rate in the power amplifier and Eelc is the per bit energy dissipation in the radio electronics.

In this experiment, 100 sensor nodes randomly deployed in 100X100 m². The base station was positioned at (50, 175) m and the upper bound for number of clusters is set to 5% of nodes. The initial energy (d₀) of each sensor node was set to 2 J while the parameters utilized in the radio model are Eelc = 50 nJ/bit, efs = 10 pJ/bit/m² and emp = 0.0013 pJ/bit/m⁴. The microcontroller energy consumption for data aggregation (Eda) is assumed to be 5 nJ/bit/signal. The following assumptions are made throughout the experiment: error free communication channel, ideal MAC layer, and nodes are in range of each other and BS. Control packet size was set to 25 bytes, data packet size was set to 500 bytes, and 6 TDMA frames per each data gathering period was assumed.

Fig. 7 shows number of alive nodes per round for SimE and LEACH-C. The simulation was ended when number of dead nodes is greater than or equal to 90%. For LEACH-C, the first node died at round 512 and 90% of nodes died in round 1050. For SimE, the first node died at round 643 and 90% of nodes died in round 1151. Considering the network when 50% of the nodes died, SimE improves the lifetime by about 21% compared to LEACH-C. Energy consumption of the network over time is an important factor for measuring the efficiency of clustering the wireless sensor networks. In addition, the total energy consumption of network over time is shown in Fig. 8. The figure shows that SimE algorithm reduced the energy usage more than LEACH-C.

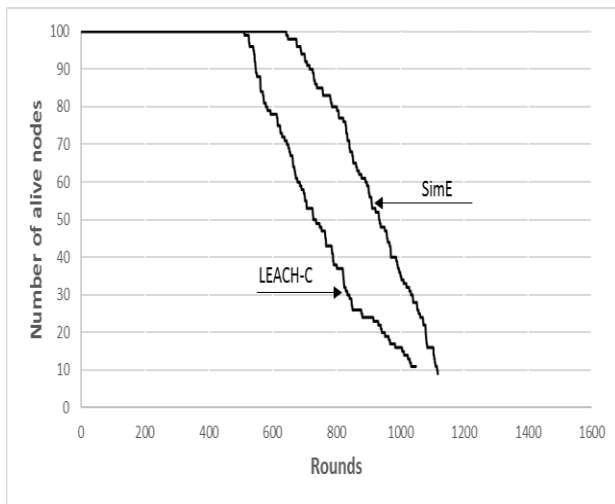


Fig. 7. Alive nodes vs. Simulation rounds.

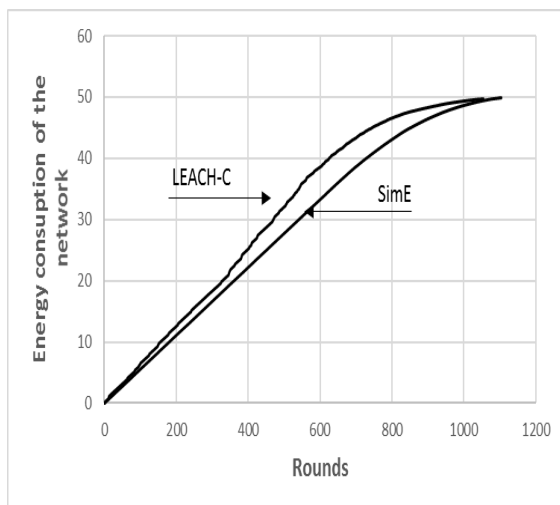


Fig. 8. Energy consumption of the network per round.

In conclusion, it is difficult to run a comprehensive comparison between the findings and performance results of this paper and other approaches proposed in the literature since the parameters utilized and the assumption made are different among the approaches. For instance, many proposals assumed a location for the BS inside or outside the network deployment area. As was illustrated in previous sections, the BS location is greatly influencing the simulations of the network. Also, other proposals assumed a fixed number of clusters (fixed CHs) in their simulations. Though, this assumption is completely avoided in our adaptive SimE approach and only an upper limit of number of clusters is used. However, some contributions assumed nearly the same assumptions made in LEACH/LEACH-C protocol. Looking at the distance reduction, however, the provided approach produced higher average distance reduction compared to [18] in most cases and CHs percentage is less. Comparing the presented SimE approach to other metaheuristics, SimE is about 8% better than PSO and SA algorithms presented in [23], which utilized mostly the same configuration and assumptions that were presented in this work.

V. CONCLUSION

In this paper, cluster optimization in wireless sensor networks was presented using simulated evolution iterative algorithm. A goodness measure was proposed to evaluate the produced clusters. The proposed SimE approach and its goodness measure had the advantage of adaptively varying the clusters and their nodes when number of nodes in the network is decreased due to the complete loss of energy. This adaptivity is important for network lifetime as the nodes in a cluster might completely lose their energy in some round; causing unbalance in the produced clusters and re-clustering the whole network. Using adaptive SimE approach, the other clusters will remain unchanged and the whole network needs not to be re-clustered.

The results showed that SimE can produce a very high quality clusters for WSNs. In addition, the results showed that there is a relationship between the size of the deployment area, the number of clusters, and the reduction in the total distance. This suggests that number of clusters should be adaptive to number of alive nodes for better clustering in WSNs. Furthermore, the results demonstrated that the base station location is crucial for effectively clustering the WSN. Finally, the results depicted that the presented SimE approach increased the network lifetime by about 21% compared to LEACH-C protocol, which utilized SA algorithm as the base for selecting CHs.

REFERENCES

- [1] E. Otero, R. Haber, A. Peter, A. AlSayyari and I. Kostanic, "A Wireless Sensor Networks' Analytics System for Predicting Performance in On-demand Deployments," *IEEE Systems Journal*, vol. 9, no. 4, pp. 1344-1353, 2015.
- [2] A. AlSayyari, I. Kostanic and C. E. Otero, "An empirical path loss model for Wireless Sensor Network deployment in an artificial turf environment," in *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control*, Miami, FL, USA, 2014.
- [3] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393-422, 2002.
- [4] A. Alsayyari, I. Kostanic and C. E. Otero, "An empirical path loss model for Wireless Sensor Network deployment in a concrete surface environment," in *2015 IEEE 16th Annual Wireless and Microwave Technology Conference (WAMICON)*, Cocoa Beach, FL, USA, 2015.
- [5] A. Alsayyari, I. Kostanic, C. E. Otero and A. Aldasary, "An empirical path loss model for wireless sensor network deployment in a dense tree environment," in *IEEE Sensors Applications Symposium (SAS)*, Glassboro, New Jersey, USA., 2017.
- [6] A. Alsayyari, I. Kostanic, C. Otero, M. Almeer and K. Rukieh, "An empirical path loss model for wireless sensor network deployment in a sand terrain environment," in *IEEE World Forum on Internet of Things (WF-IoT)*, Seoul, 2014.
- [7] P. Azad and V. Sharma, "Cluster Head Selection in Wireless Sensor Networks under Fuzzy Environment," *ISRN Sensor Networks*, vol. 2013, p. 8 pages, 2013.
- [8] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communication*, vol. 1, no. 4, pp. 660-670, 2002.
- [9] V. Kumar, A. Kumar, Gaurav and M. Singh, "Improving network lifetime & reporting delay in wireless sensor networks using multiple mobile sinks," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, 2016.
- [10] A. Jain and A. K. Goel, "Energy Efficient Algorithm for Wireless Sensor Network using Fuzzy C-Means Clustering," *International*

- Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 4, 2018.
- [11] A. Ray and D. De, "Energy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network," *IET Wireless Sensor Systems*, vol. 6, no. 181-191, p. 6, 2016.
- [12] D. Napoleon and P. G. Lakshmi, "An efficient K-Means clustering algorithm for reducing time complexity using uniform distribution data points," in *Trendz in Information Sciences & Computing (TISC2010)*, Chennai, India, 2010.
- [13] F. Cao, J. Liang and G. Jiang, "An initialization method for the K-Means algorithm using neighborhood model," *Computers & Mathematics with Applications*, vol. 58, no. 3, pp. 474-483, 2009.
- [14] P. k. P. C. M. Agrawal, "Exact and Approximation Algorithms for Clustering," *Algorithmica*, vol. 33, no. 2, pp. 201-226, 2002.
- [15] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *the 33rd Annual Hawaii International Conference on System Sciences*, Maui, HI, USA, 2000.
- [16] R. Elhabyan and M. Yagoub, "PSO-HC: Particle swarm optimization protocol for hierarchical clustering in Wireless Sensor Networks," in *10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, Miami, FL, USA, 2014.
- [17] S. Kirkpatrick, C. D. Gelatt Jr. and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671-680, 1983.
- [18] S. Jin, M. Zhou and A. S. Wu, "Sensor network optimization using a genetic algorithm," in *7th World Multiconference*, 2003.
- [19] G. EkbataniFard, R. Monsefi, M.-R. Akbarzadeh-T and M. H. Yaghmaee, "A multi-objective genetic algorithm based approach for energy efficient QoS-routing in two-tiered Wireless Sensor Networks," in *IEEE 5th International Symposium on Wireless Pervasive Computing 2010*, Modena, Italy, 2010.
- [20] Yang, Zhengyu, et al. "Intermediate Data Caching Optimization for Multi-Stage and Parallel Big Data Frameworks." arXiv preprint arXiv:1804.10563, 2018.
- [21] Yang, Zhengyu, et al. "AutoTiering: automatic data placement manager in multi-tier all-flash datacenter." Performance Computing and Communications Conference (IPCCC), 2017 IEEE 36th International. IEEE, 2017.
- [22] N. M. Abdul Latiff, C. C. Tsimenidis and B. S. Sharif, "Energy-Aware Clustering for Wireless Sensor Networks Using Particle Swarm Optimization," in *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, Athens, Greece, 2007.
- [23] M. T. Mekonnen and K. N. Rao, "Cluster Optimization Based on Metaheuristic Algorithms in Wireless Sensor Networks," *Wireless Personal Communications*, vol. 97, no. 2, p. 1-15, 2017.
- [24] A. El Rhazi and S. Pierre, "A Tabu Search Algorithm for Cluster Building in Wireless Sensor Networks," *IEEE Transactions on Mobile Computing*, vol. 8, no. 4, 2009.
- [25] C. Boucetta, H. Idoudi and L. A. Saidane, "Hierarchical Cuckoo Search-based routing in Wireless Sensor Networks," in *IEEE Symposium on Computers and Communication (ISCC)*, Messina, Italy, 2016.
- [26] Yang, Zhengyu, et al. "A fresh perspective on total cost of ownership models for flash storage in datacenters." Cloud computing technology and science (CloudCom), 2016 IEEE International Conference on. IEEE, 2016.
- [27] R. Ostrovsky and Y. Rabani, "Polynomial time approximation schemes for geometric k-clustering," in *Proceedings 41st Annual Symposium on Foundations of Computer Science*, Redondo Beach, CA, USA, 2000.
- [28] R. M. Kling and P. Banerjee, "ESp: Placement by simulated evolution," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 8, no. 3, pp. 245 - 256, 1989.
- [29] M. N. Halgamuge, M. Zukerman, K. Ramamohanarao and H. L. Vu, "An Estimation of Sensor Energy Consumption," *Progress in Electromagnetic Research*, vol. 12, pp. 259-295, 2009.