

Dynamic Clustering based on Evolutionary Algorithms Asexual Reproduction Optimization (ARO)

Safoura Akhlaghi^{1*}, Mohammad Bagher menhaj²

1*- MSc. computer engineering . Department of computer engineering, Faculty of Electrical, Computer and Information Technology, Qazvin Branch, Islamic Azad University ,Qazvin, Iran

2- Faculty member. Department of computer engineering ,Faculty of Electrical ,Amirkabir University of Technology, Tehran,Iran.

Abstract

Wireless sensor networks (WSNs) in the dynamic system have various limitations, of which the limitation of using and supplying of energy, bandwidth limitation of wireless communication links among the nodes can be cited. In spite of the various applications of these networks and the efforts made to eliminate the raised constraints, the challenges still survive, The ARO optimization method which is responsible for the dynamic clustering of the system is used in this study. Then, according to fuzzy algorithm, it performs optimal routing between the clusters. In the new proposed method here, F_WU_ARO wireless sensor network is defined as 200 sensors, distributed uniformly and randomly in a square-shaped area with an area of 100 square meters. Moreover, each cluster-head (CH) has the address or the location of its neighboring clusters and can send the relevant data if needed. In the proposed hybrid method of F_WU_ARO, with its used parameters, the residual energy based on the death of the first, middle, and last network sensors showed an improvement in the level of system sustainability in comparison to F_WU_PSO algorithm at the values of 1.5, 3, and 5 percent, respectively. In addition, it showed a reduction in the time spent with the amount of 0.14 (based on the minute).

Keywords: ARO Optimization Method, WSN, PSO Algorithm, Fuzzy Algorithms, Energy Aware, Buffer

Introduction

Meta-heuristic optimization techniques have attracted great attention in solving complex optimization problems in recent years. They are stronger than traditional methods relative to the traditional method based on formal logic or math planning for many real-world applications. These techniques have been obviously developed and new ideas have been developed in optimization algorithms over the past decades.

Since many of the real-world optimization problems are becoming more complex, better optimization algorithms are required to select the best solution among a variety of candidate solutions for each problem. The purpose of using WSNs is to collect data from the surrounding environment and to send them to the well. The lack of delay of sent and a received packet, which is one of the most critical problems in the networks, have caused proposing several methods and algorithms for transferring information from sensor nodes to the well.

The data collected in WSNs are usually sent by node-to-node propagation and or multi-step propagation to a fixed control center called well [1, 2]. With the release of the collected data, significant energy is consumed from the sensor nodes. Thus, a distributed routing protocol is needed to disseminate information to the well in each sensor node. However, the timely and efficient mobility of the well in the network survey and the absence of problems in the routing and location of the wells are of great importance. In case of a delay or meeting of wells with sensor nodes in some areas before the completion of data transfer, loss of information occurs. [3].

Since examining the parameters of efficient routing algorithms are examined, e.g., parameters such as battery life, computing capability, storage and processing of sensor data are the challenges in this regard. Thus, one of the challenges is how to reduce energy consumption while increasing the network life span.

The study used a new individual-based optimization algorithm, inspired by a famous asexual reproduction optimization as a significant biological phenomenon called ARO. ARO can be considered essentially as an evolution-based algorithm that mathematically models the germination mechanism of ARO.

This study takes advantage of ARO algorithm to propose a kind of clustering suitable and applicable for dynamic and variable environments, To achieve the goal it is tried to cluster dynamic data using the methods of dealing with dynamic environments and to update the clustering at any given time, and to apply local changes to clustering.

In the proposed new approach in this study, F_WU_ARO, CH has its own neighbors' address and location and can send the relevant data if needed. Moreover, it is assumed that the clusters do not overlap, meaning that when the well is in the radio range of a cluster (CH), it is not in another cluster's radio range. The purpose of the study is to provide a new and effective approach to overcome the energy consumption challenge, which is very inexpensive and simple.

Research background

Data clustering is a kind of NP problems. Finding a solution to NP problems is very difficult. Algorithms such as swarm intelligence algorithms have partly solved this problem. By these types of algorithms, some solutions are found that are approximate to the answer.

PSO was proposed by Kennedy and Eberhart in 1995 [9, 10], presented for clustering problems and unlike K-means, they had a better performance.

Omran and El-Hajj have used PSO for the image-clustering problem. Wong et al. presented an improved version of the objective function, firstly proposed by Omran et al. [11]. In addition to Hancer, PSO et al. developed an artificial bee colony based on developed MRI images with a predetermined objective function based on the method of brain tumor segmentation [12] and [13].

Ozturk et al. have improved a new and satisfying objective function and clustered well separated and interconnected. Moreover, ant colony optimization (ACO) was used for the clustering problem. It is clear that the number of clusters cannot be easily identified in many real-world applications and datasets [14].

Thus, the mentioned algorithms that need the number of clusters as a parameter cannot be effectively used. To understand this, the optimal number of clusters in a dataset and converting to a research area is important. It was proposed by Ball and Hall [15] that ISODATA splits or integrates clusters throughout the program based on the specific criteria to increase or decrease the number of clusters.

However, ISODATA asks the user to specify the values of different parameters (e.g., merging and splitting thresholds) and can merge only two clusters under the specific threshold of a user [16]. Dynamic optimal cluster-merge dynamic search (DYNOC) is similar to ISODATA, based on maximizing the minimum distance from the cluster to the maximum distance within the cluster, but it needs to be specified by the user.

Snob [17] and Wallace's program is for unobserved classification that decided the best data classification to assign things to a cluster using the minimum principles [message or description] and length [encoding] (MMD or MML).

Evolution-based algorithms have been used for the dynamic clustering problem, especially in the past decade [18]. Omran et al. proposed a PSO based dynamic clustering PSO (Dc PSO) inspired by the ideas of Kuncheva and Bezdek [18]. In Dc PSO, a cluster set (s) was created for the first time, and then a binary PSO was used to select s cluster centers. Then the cluster centers of s were obtained in the best solution defined by K-means.

Das et al. [19] and [20] proposed differential-evolution-based algorithms (ACDE, AFDE) where the parameters of the F scale and the intersection rate are adaptively determined.

In ACDE, each solution is shown by cluster centers and associated activation values [1 and 0].

Cluster centers and their activators are updated simultaneously through evaluation. Thus, it is not necessary to use K-means to reduce the effects of initial conditions, such as in Dc PSO.

Kuo et al. [21] improved a hybrid PSO & GA algorithm to overcome the convergence problem of the PSO algorithm. However, based on the only Euclidean distance that is not very suitable for the dynamic clustering problem, the objective matching function is used.

Maulik and Saha [22] proposed a modified differential evolutionary clustering algorithm based on the best local and global situation data (MODEAFC) that automatically extract information from remote sensing images.

Rui et al., DE and PSO, have been used for pair and duplicate, respectively, and a comparative study on clustering validity indices has been presented [23].

Proposed method

Figure 1 shows the proposed method. In the proposed method, based on the combination of three powerful fuzzy algorithms, it is energy-aware and buffer with clustering based on the ARO algorithm. It is to find the least distance traversed in the algorithm is by measuring the distance of CHs and mobile nodes and satellites as well as the energy level of each node in the trajectory.

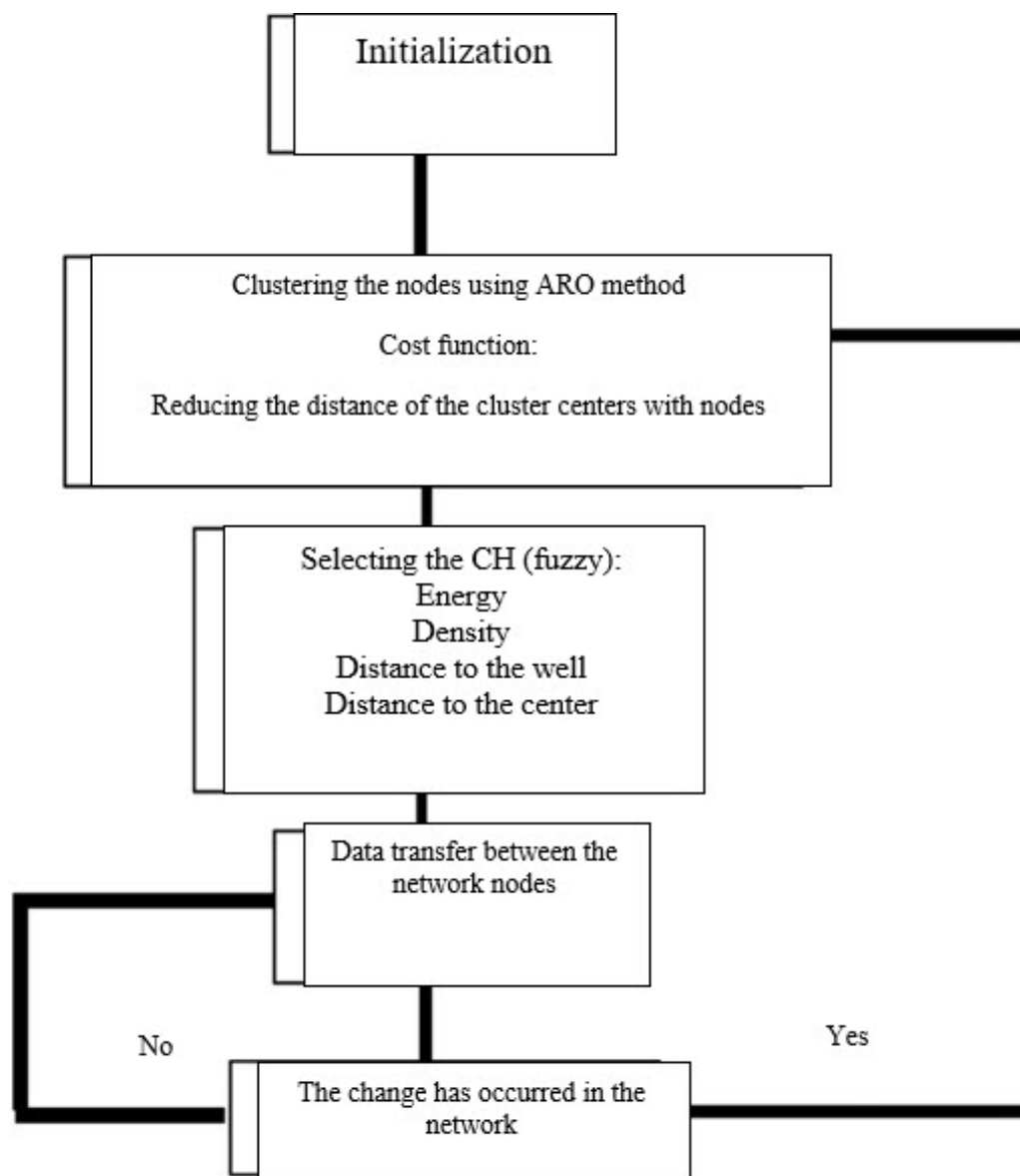


Figure 1: The proposed method algorithm F_WU_ARO

In the proposed method, it is assumed that sensor nodes are constant and equipped with a GPS device that locates its geographical location. We have a central well and the network is clustered and each cluster has a CH and mobile nodes.

Sensors with data ready to be sent will send it to the CH and the cluster will send the data to its neighbors.

The method is based on the fuzzy algorithm, so that the well route and the nodes information and packet are transmitted by creating a learned rule from their previous mobility of the packets to implement the above method using the distance traveled (the number of nodes seen as well as CH energy and the number of empty buffers), using a fuzzy method to get the best node among the routes selected above.

In the fuzzy system designed for the proposed method, there are nine rules for the fuzzy inference. The rules are given in Table (1).

Table 1: The rules used in fuzzy section

	Energy	Distance	Center	Density	Chan
1	far	Low	Low	High	Very_large
2	Far	Low	Low	Med	Large
3	Far	Low	Low	Low	Large_low
4	Low	High	High	Low	Very low
5	Med	Med	Med	Med	Little_med
6	Med	High	Med	Med	Little_low
7	Low	Low	Low	Low	Low
8	far	Med	Low	High	Large
9	far	High	Med	Med	Large_low

A WSN has been defined with 200 sensors randomly distributed in a square-shaped with a 100-meter wide area. All of the initial parameters and their values are given in Table (2) for the network definition.

Table 2: The parameters used

Parameter	Value
Cluster number	4
Number of nodes	200
Number of satellites	3
Network range	[100 100]
Well interval	[50 50]
Node awaking energy	0.1%
Number of mobile nodes	5
Message length	8
Duration of the move	20
Landa queue (landa_que)	2^5
Miou Queue (mue_que)	5
Buffer	10

Results

The results of the comparison of the basic methods and the proposed method, including nodes distribution diagram, node death drawing diagram, the first node death in the fuzzy algorithm, half death, and the end of death are studied. As is seen from the figures, diagrams and Table 3, one can understand that in the proposed method F_WU_ARO, with the parameters used, the residual energy based on the death of the first, middle, and last network sensors showed an improvement in the level of system sustainability in comparison to F_WU_PSO algorithm at the values of 1.5, 3, and 5 percent, respectively. In addition, it showed a reduction in the time spent to the amount of 0.14 (based on the minute).

However, these improvements have occurred while the total number of packets has been at its peak, which means that the network has been tested at its worst state and has reached the results and improvements reported. Although there is no change in the number of interference packets, the time elapsed (based on the minute) has decreased significantly.

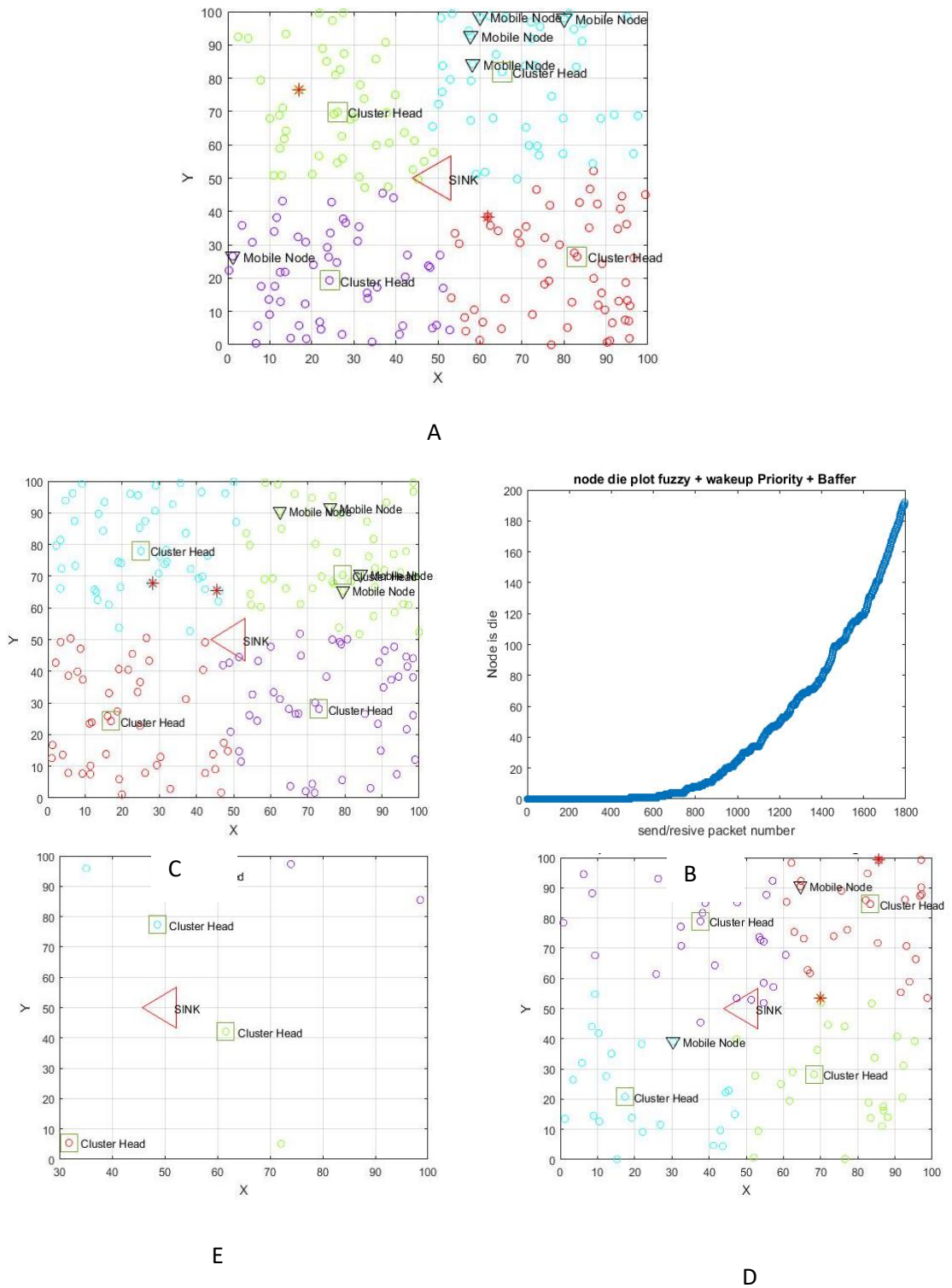


Figure 2: Results of the PSO method: A) Node distribution diagram, B) Diagram of the first node death, C) The first node death the in the fuzzy algorithm, D) Half death, E) The end of death

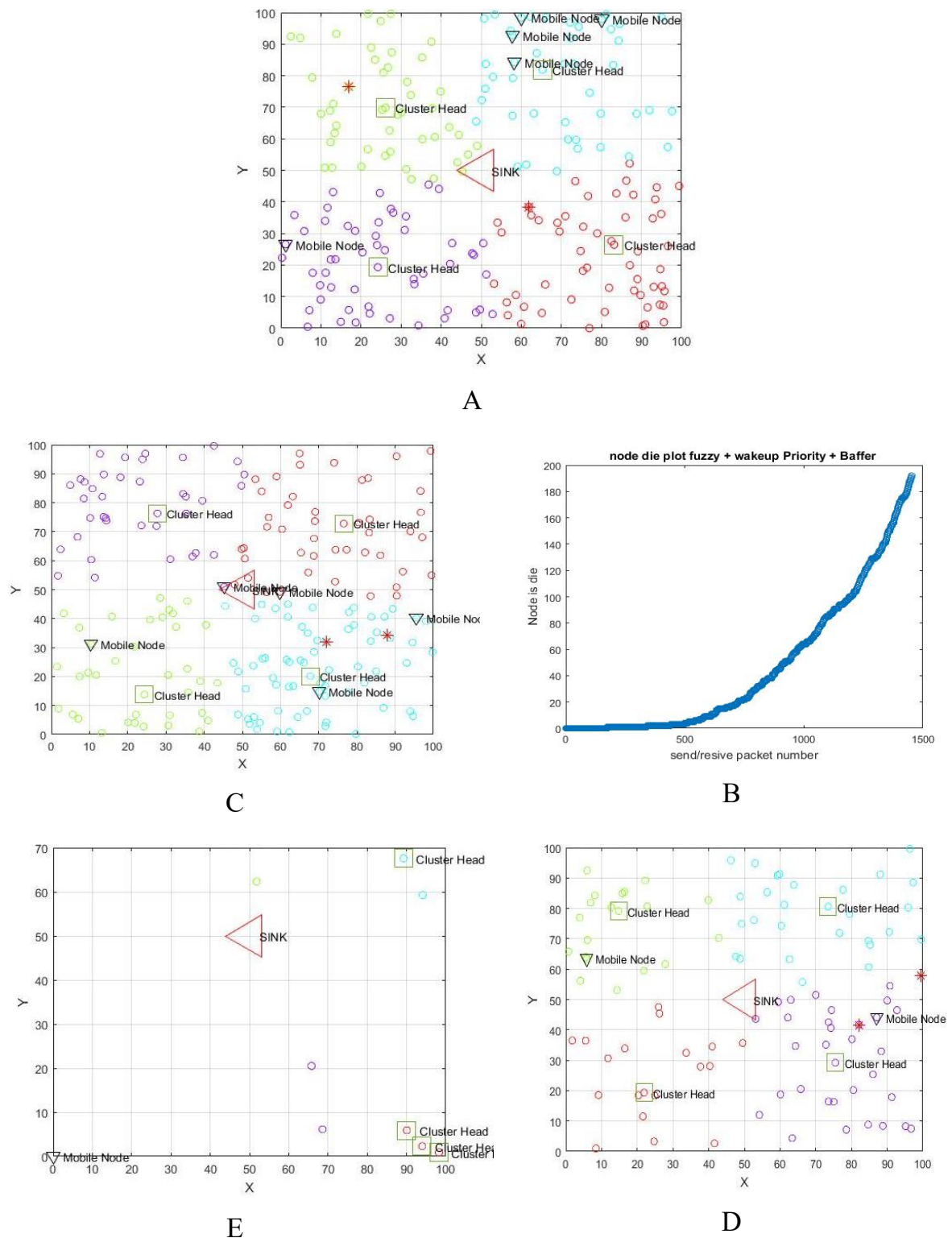


Figure 3: Results of ARO method: A) Nodes distribution diagram, B) Diagram of the first node death, C) The first node death in the fuzzy algorithm, D) Half death, E) The end of death

Table 3: Comparison of Simulated Models

Model Parameter	fuzzy + wakeup Priority + Buffer + PSO	fuzzy + wakeup Priority +Buffer +ARO
Total number of clusters	46	35
Residual energy based on the death of the first network sensor	950 Energy Remained %(5.5)	960Energy Remained %(3.8)
Residual energy based on the death of the middle network sensor	973 Energy Remained %(3.8)	993Energy Remained %(2.7)
Residual energy based on the death of the last network sensor	996 Energy Remained %(0.35)	996 Energy Remained %(0.3)
Number of he sent packets	890	996
The number of received packets	865	992
Time elapsed (in minutes)	0.219497	0.189497

Conclusion

One of the most significant parameters is the use of these networks for monitoring and controlling sensor nodes, so that it reduces energy consumption and at the same time sends packets without losing information and at the fastest time. In traditional routing, the constant well node and certain sensor nodes near the wells have the role of relay the data of most network segments. Thus, heavy traffic volume and high-energy consumption and battery drain lead to a quick disconnection on the network in a short time.

In this study, the ARO optimization method is used, which is responsible for the dynamic clustering of the system. Then, it performed optimal routing between the clusters according to the fuzzy algorithm. Among the advantages of the proposed method are eliminating any additional costs on the network and that in the proposed network, the failure of a part of the network does not lead to the failure of any part of the network other than that part besides it is considerably inexpensive and simple to implement due to its software-based method.

The study combined three methods based on the combination of fuzzy, energy aware, and buffer algorithms - F_WU_ARO - not requiring any additional costs in the network.

In the proposed hybrid method, F_WU_ARO, with the used parameters, the residual energy based on the death of the first, middle, and last network sensors showed improvement relative to the F_WU_PSO algorithm and the time spent (based on the minute) reduced drastically. However, these improvements have occurred when the total number of sent packets has been at its peak, meaning the network has been tested at its worst state and has reached the results and improvements reported. Although there is no change in the number of interference packets, the time elapsed (based on the minute) has significantly reduced. Moreover, there is no additional network cost in the proposed method. Additionally, the failure of a part of the network does not lead to the failure of any part of the network other than that part. In addition, given the software nature of the proposed new method, the implementation cost is very low and simple in terms of computational complexity.

References

- [1] Anastasi. G, Conti. M, Di Francesco. M, Passarella. A, "Energy Conservation in Wireless Sensor Networks: A Survey", Ad Hoc Networks, vol. 7, no. 3, PP. 537–568, 2009.
- [2] Akyildiz. I.F, Su. W, Sankarsabramaniam. Y, Cayirci. E, "Wireless Sensor Networks: A Survey", Computer Networks, vol. 38, PP. 393-422, 2002.
- [3] Estrin. D, Girod. L, Pottie. G, Srivastava. M, 2001,"Instrumenting the World with Wireless Sensor Networks", International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001), Salt Lake city, Utah.
- [4] Akyildiz. I.F, Su. W, Sankarsubramaniam. Y, Cayirci. E, 2002, "Wireless Sensor Networks: a Survey", Computer Networks 38, PP. 393-442.
- [5] J. Wang, Y. Yin, J.U. Kim, S. Lee, and C.F. Lai, "An Mobile-Sink Based Energy-Efficient Clustering Algorithm For Wireless Sensor Networks", IEEE 12th International Conference on Computer and Information Technology, pp. 678–683, 2012.

- [6] Hancer E, Xue B, Zhang M, et al. Pareto front feature selection based on artificial bee colony optimization. *Information Sciences* 2018; 422: 462-79
- [7] Das, K., and Sengupta, S., "A Hierarchical Process Industry Production–Distribution Planning Model", *International Journal of Production Economics*, Vol.117, pp. 402-419, 2009.
- [8] Saha, I., Maulik, U., Plewczynski, D., —Multi objective differential crisp clustering for evaluation of clusters dynamically, *Advances in Intelligent and Soft Computing*, AISC 103 , pp. 307-313, 2011.
- [9] J. Kennedy, R. Eberhart, Particle swarm optimization, in: *Proceedings of the IEEE International Joint Conference on Neural Networks*, Australia, 1995, pp.1942–194
- [10] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: *6th International Symposium on Micro Machine and Human Science*, 1995.
- [11] W. Man To, H. Xiangjian, Y. Wei-Chang, Image clustering using particle swarm optimization, in: *IEEE Congress on Evolutionary Computation (CEC)*, 2011, 2011, pp. 262–268.
- [12] E. Hancer, C. Ozturk, D. Karaboga, Extraction of brain tumors from MRI images with artificial bee colony based segmentation methodology, in: *8th International Conference on Electrical and Electronics Engineering (ELECO)*, 2013, 2013, pp. 516–520.
- [13] E. Hancer, C. Ozturk, D. Karaboga, Artificial Bee Colony based image clustering, in: *IEEE Congress on Evolutionary Computation (CEC)*, 2012, Brisbane, Australia, 2012.
- [14] C. Ozturk, E. Hancer, D. Karaboga, Improved clustering criterion for image clustering with artificial bee colony algorithm, *Pattern Anal. Appl.* (2014), <http://dx.doi.org/10.1007/s10044-014-0365-y> (in press).
- [15] G. Ball, D. Hall, A clustering technique for summarizing multivariate data, *Behav. Sci.* 12 (1967) 153–155.
- [16] J. Tou, R. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley, Massachusetts, USA, 1979.
- [17] C. Wallace, D. Dowe, Intrinsic classification by MML – the snob program, in: *Seventh Australian Joint Conference on Artificial Intelligence*, UNE, Armidale, NSW, Australia, 1994, pp. 37–44. [30] L.I. Kuncheva, J.C. Bezdek, Nearest.
- [18] M.G.H. Omran, A. Salman, A.P. Engelbrecht, Dynamic clustering using particle swarm optimization with application in image segmentation, *Pattern Anal. Appl.* 8 (2006) 332–344.
- [19] L.I. Kuncheva, J.C. Bezdek, Nearest prototype classification: clustering, genetic algorithms, or random search? *IEEE Trans. Syst. Man Cybern. C: Appl. Rev.* 28(1998) 160–164.
- [20] S. Das, A. Abraham, A. Konar, Automatic clustering using an improved differential evolution algorithm, *IEEE Trans. Syst. Man Cybern. C: Syst. Hum.* 38 (2008) 218–237.
- [21] S. Das, A. Konar, Automatic image pixel clustering with an improved differential evolution, *Appl. Soft Comput.* 9 (2009) 226–236
- [22] R.J. Kuo, Y.J. Syu, Z.-Y. Chen, F.C. Tien, Integration of particle swarm optimization and genetic algorithm for dynamic clustering, *Inf. Sci.* 195 (2012) 124–140.
- [23] U. Maulik, I. Saha, Automatic fuzzy clustering using modified differential evolution for image classification, *IEEE Trans. Geosci. Remote Sens.* 48 (2010) 3503–3510.