Enhancement of LEACH Based on K-means Algorithm and Stochastic Optimization

Tuyen Nguyen Viet, Trang Pham Thi Quynh, and Hang Duong Thi
Faculty of Electronics, Hanoi University of Industry, Hanoi, 100000, Vietnam
Email: nvtuyen@haui.edu.vn, pham.trang@haui.edu.vn, hangdt@haui.edu.vn

Abstract —In Wireless Sensor Networks (WSNs), maximizing the life of the Sensor Nodes (SNs), and energy conservation measures are essential to enhance the performance of the WSNs. A Low-Energy Adaptive Clustering Hierarchy (LEACH) routing protocol has been proposed specifically for WSNs to increase the network lifetime. However, in LEACH protocol the criteria for clustering and selecting Cluster Heads (CHs) nodes were not mentioned. Accordingly, researchers have been focusing on ways to strengthen the LEACH algorithm to make it more efficient. In this paper, we propose to improve the LEACH protocol by combining the use of K-means algorithm for clustering and Slime Mould Algorithm (SMA), a new stochastic optimization to select nodes as CHs. The proposed routing algorithm, called SMA-LEACH, is superior to other algorithms, namely PSO-LEACH, BA-LEACH, which using Particle Swarm Optimization (PSO), Bat Algorithm (BA) to improve LEACH, respectively. Simulation analysis shows that the SMA-LEACH obviously reduces network energy consumption and extends the lifetime of WSNs.

Index Terms—Wireless sensor networks, low-energy adaptive clustering hierarchy protocol, network lifetime, slime mould algorithm, energy efficiency

I. Introduction

Wireless sensor networks include SNs with limited energy which collect parameters environment and transmit information to the Base Station (BS) to monitor and detect parameters according to different applications [1]. Due to the limited energy of the SN, it is essential to expand their use as much as possible when studying WSNs. Routing protocol technology has a great influence on the network life cycle. LEACH was the first and most commonly used hierarchical routing protocol [2]. In the LEACH protocol, SNs are clustered, the clusters perform the function of collecting and transmitting data to BS through the CH. With this principle, LEACH can extend the life of the network, reduce the energy consumption of each node, and gather data to reduce message transmission in the network. However, the LEACH protocol does not consider the current node energy and random selection of CHs which can easily lead to uneven energy consumption between network nodes, and shorten the network life. Much work has been done related to LEACH and a new study in the WSNs field is a good idea to adopt LEACH and its variations, many of which reduce energy consumption in the LEACH protocol; others consider the energy consumption balance [2], [3]. Most of the recent improvements are based on CHs selection using nature-inspired algorithms. Typical of which are improving LEACH use PSO (PSO-LEACH) [4], [5], and use BA (BA-LEACH) [6] which increased the global search ability of the algorithm.

In 2020, Shimin Li et al proposed a new method for stochastic optimization algorithm, which is called slime mould algorithm (SMA) [7]. This algorithm was inspired by the diffusion and foraging conduct of slime mould. Compared with the previously proposed nature-inspired optimization algorithm, it is regarded as a global optimization method with better robustness and applicability.

In this paper, SMA-LEACH routing algorithm based on SMA and K-means algorithm is first proposed, a new CH selection mode is adopted to solve the problem of uneven CH distribution in LEACH, which causes non-uniform energy consumption. The rest of this paper is structured as follows: In Section II, we review the background of the LEACH convention and SMA. Section III uses SMA for CH selection optimization. Section IV, we verify the proposed improvements through simulation experiments through comparing with some of the existing LEACH enhancements, PSO-LEACH and BA-LEACH, and the conclusions are presented at the end of the paper.

II. Overview of the LEACH Protocol

LEACH was proposed by Wendi B. Heinzelman in 2002, which was the first hierarchical wireless sensor routing protocol [8]. Fig. 1 shows the architecture of LEACH.

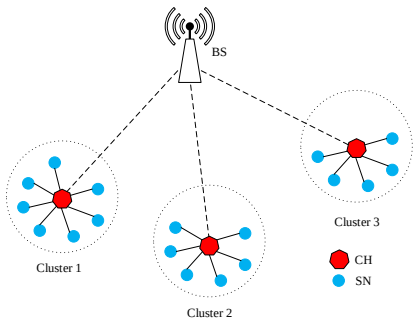

Fig. 1. Architecture of LEACH

Manuscript received February 28, 2021; revised August 17, 2021.
This work was supported by the Hanoi University of Industry (HaUI) under Grant No. 27-2020-RD/HĐ-ĐHCN.
Corresponding author email: nvtuyen@haui.edu.vn
doi:10.12720/jcm.16.9.406-410

The LEACH protocol operates in many rounds, each of which consists of two phases: set-up phase and steady-state phase. The CHs are elected in the set-up phase and CHs selection is solved at the beginning of each round. The CH for each cluster receives and aggregates the data from cluster members and then transmits the aggregated data to the BS. To become a CH, nodes generate a random number between 0 and 1 comparing with a threshold $T(n)$. The node is elected as a CH at current round if the number generated is less than a certain threshold $T(n)$. The threshold value is defined as (1)

$$T(n)=\begin{cases}\dfrac{P}{1-P\left(r \bmod \left(\dfrac{1}{P}\right)\right)} & , \mathrm{n} \in \mathrm{G} \\ 0 & , \mathrm{n} \notin \mathrm{G}\end{cases} \tag{1}$$

where, P is the denoted percentage of the node to be selected as CHs, r represents the current round , and G is the set of nodes that have not been accepted as CHs in the last $1/P$ rounds.

III. Improved LEACH Approach

In LEACH protocol, the CH selection process is completely based on the random number which node generates. Other features of the nodes, such as the current residual energy, and location are without considering which leads to following problems:

- The residual energy of nodes is not considered when selecting CH, which can cause low-energy node to be selected as a CH, causing the node energy exhausted rapidly.
- Without considering the location of the node when CH is distributed, and without guaranteeing the uniform distribution of CH may cause some CHs densely distributed or too sparsely dispersed.

To reduce these defects, we present a new stochastic optimization algorithm into CH selection through optimization of objective the function and use the K-means algorithm to cluster.

A. Built Fitness Function

In WSNs using the LEACH protocol, energy consumed when transferring from sensor nodes to CH node is determined by (2) [9]

$$E_{SN_i-CH_j}(l)=\begin{cases} l\left(E_{Tx}+\varepsilon_{mp} d^4_{(SN_i-CH_j)}\right) \\ \qquad if \quad d_{(SN_i-CH_j)}>d_0 \\ l\left(E_{Tx}+\varepsilon_{fs} d^2_{(SN_i-CH_j)}\right) \\ \qquad if \quad d_{(SN_i-CH_j)}\le d_0 \end{cases} \tag{2}$$

where, E_{Tx} is transmitter energy per node, l is number of bits per data packet, ε_{mp} is amplification energy when distance from a sensor node to CH is greater than threshold d_0, ε_{fs} is amplification energy when distance from a SN to CH is less than threshold d_0

$$d_0=\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{3}$$

$d_{(SN_i-CH_j)}$ is Euclidean distance from $i-th$ SN to $j-th$ CH.

The energy consumed when transmitting the signal from CH to BS station is calculated by the formula (4)

$$E_{CH_j-BS}(l)=\begin{cases} l(E_{Tx}+E_{DA}+\varepsilon_{mp} d^4_{(CH_j-BS)}) \\ \qquad if\ d_{(CH_j-BS)}>d_0 \\ l(E_{Tx}+E_{DA}+\varepsilon_{fs} d^2_{(CH_j-BS)}) \\ \qquad if\ d_{(CH_j-BS)}\le d_0 \end{cases} \tag{4}$$

where, E_{DA} is data aggregation energy, $d_{(CH_j-BS)}$ is Euclidean distance from $j-th$ CH to BS.

Our aim is to select the central node so that the total energy consumed during transmission and receiving data in the cluster is minimal. Furthermore, the energy transmitting and receiving data between sensor nodes were highly dependent on the distance between them. Therefore, we proposed the fitness function for selecting CH node as follows:

$$\begin{aligned} Fitness=&\xi\left(E_{\mathrm{CH_j\text{-}BS}}+\sum_{i=1}^{C_j}E_{SN_i-CH_j}\right) \\ &+(1-\xi)\left(d_{(CH_j,BS)}+\sum_{k=1}^{C_j}d_{(SN_i-CH_j)}\right)\end{aligned} \tag{5}$$

In (5), $\xi \in (0,1)$ is the weight, set at 0.9 in this experiment based on study [10], C_j is the number of nodes in $j-th$ cluster.

B. Improved LEACH based on K-means algorithm and SMA

The SMA was inspired by the propagation wave of slime mould in nature based on bio-oscillator to form the optimal path for connecting food. In this algorithm, each cluster is assumed to be a slime mould. The location of sensor node in the cluster whose coordinates are $X(x_d, y_d)$ represents the individual location in the slime mould's current state. The individual location with the highest odor concentration found corresponding to CH. The highest odor concentration was found according to the criteria of fitness function (5). The SMA is summarized as follows:

The mathematical formula for updating the location of slime mould is as (6)

$$X_{new}=\begin{cases} rand\left(U_B-L_B\right)+L_B & rand<z \\ X_{best}+v_b\left(\mathrm{W}.X_{A_{old}}-X_{B_{old}}\right) & \beta<p \\ v_c X_{old} & \beta\ge p \end{cases} \tag{6}$$

where, U_B and L_B denote the lower and upper boundaries of search range, *rand* and β denote the random value in $[0,1]$.The value of z will be discussed in the parameter setting experiment. v_b is a parameter with a range of $[-a,a]$.

$$a = \arctan h(-(\frac{t}{\max\_t})+1) \tag{7}$$

v_b decreases linearly from one to zero. X_{best} is the best position, X_{new} and X_{old} is represents the location of slime mould after and before developing. $X_{A_{old}}$ and $X_{B_{old}}$ represent two individuals randomly selected from slime mould.

$$p = \tanh|fitness(i) - DF| \tag{8}$$

DF represents the best fitness obtained in all iterations. W, the weight of slime mould is listed as follows:

$$W(i) = \begin{cases} 1+\beta log\left(\frac{bF - fitness(i)}{bF - wF}+1\right)? & Con \\ 1-\beta \log\left(\frac{bF - fitness(i)}{bF - wF}+1\right)? & Others \end{cases} \tag{9}$$

where *Con* indicates that ranks first half of the population, β denotes the random value in the interval of $[0,1]$, bF is the best fitness , wF is the worst fitness.

Input: Positions, Energy of sensor nodes, position of BS
Output: Cluster head position, state of each sensor node, Energy consumption in each sensor node.
Set parameters: N (number of individuals) U_B, L_B, Max_round (loop number)
While not all node die
    Compute the number of clusters by (1)
    Clustering nodes according to K-means algorithm
    Set parameters: N, U_B, L_B, Max_round, w, X(i).
    Compute the Fitness (X_i), of each particle use equation (4) find the best position and set it to X_{best}.
    While (i< Max_round)
    Update the temp position by (6)
    Calculate Fitness (X_i)
        Arrange the value of the fitness function
    Compute the value of weight by (9)
    Update the X_{best}, X_{worst}
        i = i+1
    End
End.

Fig. 2. Pseudo-code of SMA-LEACH

The K-means algorithm is used for clustering, which aims to partition N sensor nodes into K clusters in which each SN belongs to the cluster with the nearest mean (cluster centers or cluster centroid). Combining K-means algorithm and SMA to improve LEACH, called SMA-LEACH. The pseudo-code of SMA-LEACH is presented as Fig. 2.

IV. Simulation Results

In this section, we evaluate the effectiveness of the proposed algorithm via implementation in MATLAB. The parameter settings of simulation as described in Table I.

TABLE I: THE PARAMETERS OF SMA-LEACH

Parameter	Value
Network field	500mx500m
Number of SN	100
Initial energy of Nodes E_0	0.5J
E_{Tx}	50 nJ/ bit
E_{Rx}	50 nJ/ bit
E_{DA}	5 nJ/ bit
Message Size	4000 bits
ε_{mp}	0.00013pJ/bit
ε_{fs}	10pJ/bit
Maximum No. of Iteration	3000
ξ	0.9
P	0.1

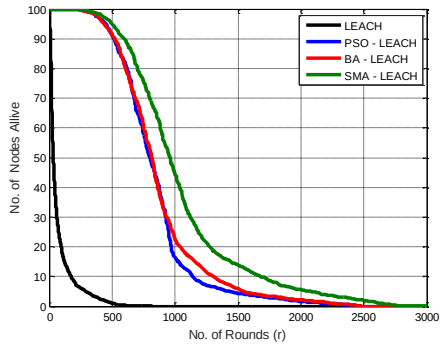

Fig. 3. The number of live nodes

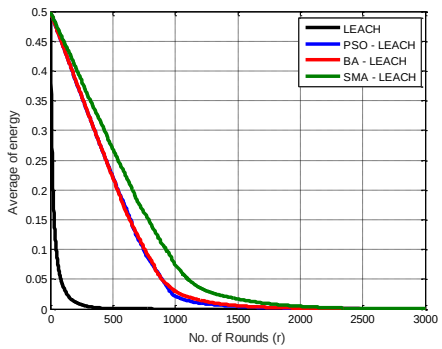

Fig. 4. The average residual energy

Fig. 3 shows the comparison of live nodes in each round for LEACH, PSO-LEACH, BA-LEACH, and SMA-LEACH algorithms. Fig. 4 shows the average of residual energy in each round of LEACH, PSO-LEACH, BA-LEACH, and SMA-LEACH protocol. The results of this simulation show that the number of live nodes and average energy corresponding rose considerably in SMA-LEACH, BA-LEACH and PSO-LEACH protocol when compared with LEACH protocol. Furthermore, the number of live nodes in SMA-LEACH was higher than that of PSO-LEACH and BA-LEACH.

Fig. 5 shows the time first node, middle node, and all dead nodes in LEACH, PSO-LEACH, BA-LEACH, SMA-LEACH protocol. It is clear that in LEACH all nodes die after 631 rounds, in PSO-LEACH is 1755 rounds, in BA-LEACH is 1846 rounds, and in SMA-LEACH stood at 2313 rounds. The lifetime of WSNs using PSO algorithm increased by 78.1% and using BA algorithm increased by 92.6% compared to LEACH. While the lifetime of WSNs using SMA rose significantly about 166.6% compared to LEACH, about 8.1% and 3.8% when compared with PSO-LEACH and BA-LEACH. The number of rounds when the middle node and first dead node of SMA-LEACH was higher than that of LEACH, PSO-LEACH, and BA-LEACH. In summary, enhancement of LEACH protocol based on SMA shows higher efficiency than the previous optimization algorithms such as PSO, BA.

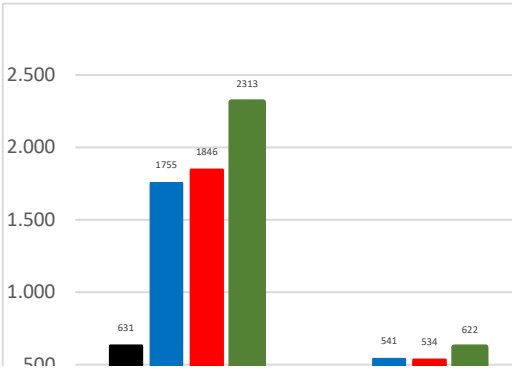

Fig. 5. Compare the time when the first, middle and all dead nodes

V. Conclusion

Saving and efficiently using energy consumption and load balancing are significant challenges of the routing algorithms in WSNs. In this paper, an efficient clustering and CH selection algorithm is proposed to upgrade the LEACH protocol. By using K-means algorithm for clustering, using SMA to select SN as CH node leads to reduced average WSN energy consumption, the lifespan of WSNs is extended compared to LEACH, PSO-LEACH, BA-LEACH. The SMA-LEACH algorithm has flexibility in choosing CH as well as clustering.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Tuyen Nguyen Viet proposed an idea, Trang Pham Thi Quynh, and Hang Duong Thi contributed to the simulation and analyzed the data. All authors had written the paper and approved the final version.

References

[1] D. Zhang, G. Li, K. Zheng, X. Ming, and Z. H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Trans. Ind. Informatics*, vol. 10, no. 1, pp. 766–773, 2014.

[2] S. K. Singh, P. Kumar, and J. P. Singh, "A Survey on Successors of LEACH Protocol," *IEEE Access*, vol. 5, pp. 4298–4328, 2017.

[3] Z. Cui, Y. Cao, X. Cai, J. Cai, and J. Chen, "Optimal LEACH protocol with modified bat algorithm for big data sensing systems in Internet of Things," *J. Parallel Distrib. Comput.*, vol. 132, pp. 217–229, 2019.

[4] G. K. Nigam and C. Dabas, "ESO-LEACH: PSO based energy efficient clustering in LEACH," *J. King Saud Univ. - Comput. Inf. Sci.*, January, 2018.

[5] D. R. Edla, M. C. Kongara, and R. Cheruku, "A PSO based routing with novel fitness function for improving lifetime of WSNs," *Wirel. Pers. Commun.*, vol. 104, no. 1, pp. 73–89, 2019.

[6] X. Cai, Y. Sun, Z. Cui, W. Zhang, and J. Chen, "Optimal leach protocol with improved bat algorithm in wireless sensor networks," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 5, pp. 2469–2490, 2019.

[7] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Futur. Gener. Comput. Syst.*, vol. 111, pp. 300–323, 2020.

[8] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wirel. Commun.*, vol. 1, no. 4, pp. 660–670, 2002.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. Hawaii Int. Conf. Syst. Sci.*, p. 223, 2002.

[10] H. Miao, X. Xiao, B. Qi, and K. Wang, "Improvement and application of LEACH Protocol based on Genetic Algorithm for WSN," in *Proc. IEEE 20th Int. Work. Comput. Aided Model. Des. Commun. Links Networks, CAMAD 2015*, 2016, pp. 242–245.

Tuyen Nguyen Viet was born in 1978 in Hanoi, Vietnam. He received the B.E degree in Electronics Telecommunications and the M.E degree in Information processing and Communication from Hanoi University of Science and Technology, in 2001 and 2006, respectively. He obtained the PhD degree in Electronic Engineering from Le Quy Don University in 2019. He has been working as a Lecturer at the Faculty of Electronics, Hanoi University of Industry. His research interests include digital signal processing and wireless communication.

Trang Pham Thi Quynh was born in 1979 in Nam Dinh, Vietnam. She has been a lecturer at Faculty of Electronic Engineering, Hanoi University of Industry since 2001. She received the B.S and M.S degree from VNU University of Engineering and Technology, in 2003 and in 2006, respectively. She is interested in wireless communication, wireless sensor network Optimization algorithm, digital signal processing, neural network, and nature–inspired algorithm application, and FPGA technology.

Hang Duong Thi was born in 1978 in Bac Giang, Vietnam. She has been a lecturer at Faculty of Electronic Engineering, Hanoi University of Industry since 2000. She received the B.S and M.S degree from VNU University of Engineering and Technology, in 2000 and in 2005, respectively. She is currently pursuing the Ph.D. degree in telecommunication engineering at VNU. Her main research interests are indoor positioning systems, machine learning, pattern classification, and nature-inspired algorithm application.