Journal of information and
communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 22(1): 1-6, Mar. 2024 Regular paper

Development of Energy-sensitive Cluster Formation and Cluster Head Selection Technique for Large and Randomly Deployed WSNs

Sagun Subedi and Sang Il Lee*

Department of Information and Electronic Engineering, Mokpo Nation University, 58554, Republic of Korea

Abstract

Energy efficiency in wireless sensor networks (WSNs) is a critical issue because batteries are used for operation and communication. In terms of scalability, energy efficiency, data integration, and resilience, WSN-cluster-based routing algorithms often outperform routing algorithms without clustering. Low-energy adaptive clustering hierarchy (LEACH) is a cluster-based routing protocol with a high transmission efficiency to the base station. In this paper, we propose an energy consumption model for LEACH and compare it with the existing LEACH, advanced LEACH (ALEACH), and power-efficient gathering in sensor information systems (PEGASIS) algorithms in terms of network lifetime. The energy consumption model comprises energy-sensitive cluster formation and a cluster head selection technique. The setup and steady-state phases of the proposed model are discussed based on the cluster head selection. The simulation results demonstrated that a low-energy-consumption network was introduced, modeled, and validated for LEACH.

Index Terms: Clustering algorithm, Cluster head, Energy efficiency, Network lifetime

I. INTRODUCTION

Wireless sensor networks (WSNs) consist of a few thousand sensor nodes that sense the physical properties of the environment, such as temperature, humidity, or vibration. This physical information contributes to several applications, including disaster management [1], RFID [2], and medical applications [3]. These sensor nodes are wirelessly connected to each other. There exists at least one node called a base or sink node that connects all other nodes to the outer world. Most studies in this regard can be classified into random and deterministic deployments [4]. WSN models have been upgraded from simple graph-based characterizations of interference to more accurate physical models, such as the signal-to-interference-plus-noise ratio model [5]. The structure of a sensor network and its topology determine whether a node can successfully receive and decode messages. Network lifetime remains an important metric in the deployment of nodes in the environment.

Among the numerous challenges faced by sensor nodes, power utilization is the most important. The energy required by the sensor nodes is provided by onboard batteries. Sensor nodes may be deployed on difficult terrains, where changing the battery of the sensor nodes would be difficult. Another significant method for saving power is communication protocols. If any sensor nodes in a functioning network run out of battery, the network would not function. Consequently, the productivity of the network system would decrease. Hence, the parts and methods that are the most energy-hungry need to be understood. We ought to do whatever it takes to not

---

Received 21 December 2022, **Revised** 8 December 2023, **Accepted** 8 January 2024

***Corresponding Author** Sang Il Lee (E-mail: leesi@mokpo.ac.kr)
Department of Information and Electronic Engineering, Mokpo Nation University, 58554, Republic of Korea

Open Access **https://doi.org/10.56977/jicce.2024.22.1.1** print ISSN: 2234-8255 online ISSN: 2234-8883

utilize them or utilize them the least when necessary [6]. Energy is consumed by the nodes during the transmitting, receiving, idle, and sleeping states. Data aggregation is the process of aggregating sensor data by using aggregation approaches. The data aggregation algorithm uses sensor data from the sensor node and then aggregates the data using aggregation algorithms, such as the centralized approach, low-energy adaptive clustering hierarchy, and tiny aggregation. Data aggregation removes redundancy and unnecessary data forwarding, and hence reduces the total energy used by the system in communication significantly. The cluster head selection process of LEACH involves setup and steady-state phases. When the CH is not chosen in the WSN, all the nodes present in the network attempt to deliver their data to the base station (BS).

The data delivery process of a WSN involves two major issues: a) data convention and b) energy loss. When several nodes attempt to communicate with a BS, a huge load is created at the BS, which is known as a data convention [7]. Data convention results in multiple data routings by the sensor nodes and is usually related to the hit-and-trial method. Owing to the limited communication capacity between the BS and sensor nodes, powerful and expensive BSs need to be installed. Thousands of sensor nodes must be handled while maintaining the cost of the entire network relatively low. The energy level of these sensor nodes depends on the BS processing power. Regardless of whether the sensor nodes transmit, receive, or remain idle, energy is consumed. Energy loss plays a vital role in a WSN and should be reduced to extend the network lifetime. The information sent by each sensor node is not accepted by the BS. Although the BS has an unlimited supply of energy, the sensor nodes do not. Sensor nodes spend energy without knowing whether the information they hold will be accepted by the BS. Moreover, restricting data transmission can lead to high latency, and hence a low quality of service [8]. If the density of sensor nodes in a network system is high, the rate of information transmission to the BS or CH is slower. This slower mechanism may lead to packet drops before packets reach their respective destinations. In addition, the retransmission of these data packets consumes more energy [9,10]. This is due to the memory limitations and processing capacity of the sensor nodes.

LEACH mitigates the data conventions and energy-loss issues of WSNs by selecting a CH. The selection of CHs among the nodes and allowing only CHs to communicate with the BS are crucial procedures in LEACH algorithms. These processes not only reduce the load in the BS, but also save valuable energy in the sensor nodes. The battery-changing process is often inconvenient because it highly affects the communication chain. When the CHs are chosen in advance and fixed throughout the system lifetime, the selected CHs do not survive throughout the network life. Thus, introducing an algorithm to select the CH in each round is an effective method of saving energy. Here, rounds refer to the communication between sensor nodes and the BS via the CH, and vice versa.

In this paper, we discuss the routing path of information within the network to save energy in the nodes. We propose a dynamic model based on energy-sensitive cluster formation and total-energy-based CH selection for LEACH. The data aggregation mechanism is crucial and can be applied to evaluate the minimum energy consumption of a network. In the setup phase of LEACH, the energy constraints are defined and the CH is selected for the first round. The second round of CH selection is based on the steady-state phase of LEACH. The results of the proposed methodologies are compared with those of the existing LEACH, ALEACH, and PEASIS algorithms. The proposed methodology shows 50% more efficient results than the existing algorithms. This study demonstrates that distributing the energy load among the nodes increases the lifetime and quality of the network.

II. BACKGROUND

Forwarding the data gathered by a sensor to a BS is the primary function of application protocols in a sensor network [11]. Data collection remains the baseline for the development of different LEACH algorithms for hierarchical routing. The energy load in the LEACH algorithm is uniformly distributed among the sensor nodes. All the nodes are randomly placed and grouped into clusters. Among these clusters, one node is selected as the CH. The LEACH algorithm transmits data to the BS or sink via the CH. The LEACH protocol runs for several rounds, depending on the energy used by the sensor nodes. Each round is performed in two phases: the setup and steady-state phases [12].

In the setup phase, for every node n, a random value x is selected and compared with the threshold function $T(n)$. To select the CH, $T(n)$ can be calculated as

$$T(n)=\begin{cases}\dfrac{p}{n-p\left(rmod\left(\dfrac{n}{p}\right)\right)} & \text{if } n\in G\\ 0 \quad else & \end{cases} \qquad (1)$$

where p is the general probability, r is the number of rounds, and G is the number of unacceptable nodes in the network. When x is less than $T(n)$, the node becomes the CH; otherwise, it becomes a member. The nodes that were previously CHs cannot be assigned to a CH again. In this phase, CHs are selected to communicate with the BS.

In the steady-state phase, the CH forms a time-division multiple access schedule for its member nodes to send data. The member nodes send data in their time slots and sleep in the remaining slots. After receiving the signals, the CH applies data aggregation schemes such that several signals are converted into a single signal that incorporates the relevant and necessary information from all the signals. This signal is transmitted to the BS by all the CHs in a single-hop transmission.

ALEACH [13] uses a distributed algorithm to form clusters. The concept of the current state probability is introduced along with the general probability. If the nodes in a cluster simultaneously have different amounts of energy, the node with the highest energy is the CH. Using the current and initial node energies, $E_{current}$ and E_{INI}, respectively, the threshold function of ALEACH $T(n)$ is computed as

$$T(n) = \begin{cases} \dfrac{p}{n - p\left(r \bmod \left(\dfrac{n}{p}\right)\right)} + \dfrac{E_{current}}{E_{INI}} \times \dfrac{n}{p} & \text{if } n \in G \\ 0 & else \end{cases} \tag{2}$$

Unlike LEACH and ALEACH, PEGASIS [14] is a popular CH-based routing protocol with only a steady-state phase. PEGASIS operates by using a greedy chain protocol. The farthest node from the BS begins the construction of the chain, and selects the node that is the second farthest from the BS. After a node dies, the chain is reconstructed similarly to avoid dead nodes. Each subsequent node in the chain combines the received data with its own data. The last node in the chain is called the leader, and is allowed to communicate only with the BS.

PEGASIS outperforms LEACH by eliminating the overhead of dynamic cluster formation, minimizing the distance over which nonleader nodes must transmit, limiting the number of transmissions and receptions among all nodes, and using only one transmission to the BS per round. However, the assumption that all nodes can transmit with sufficient power to reach the CH if needed and that each node has computational power to support different MAC protocols is a major drawback of PEGASIS. The region of interest of a WSN is not always accessible [15]. The ALEACH algorithm uses random nodes to make autonomous decisions without centralized control. Distributed clusters in ALEACH are formed without prior knowledge of the node's location, and long-distance communication between the nodes and the BS is eliminated.

LEACH, PEGASIS, and ALEACH are not dynamic models for increasing the node lifetime and network quality. We developed a dynamic model based on energy-sensitive cluster formation and CH selection methods and compared it with existing algorithms in terms of node lifetime and network quality.

III. PROPOSED MODEL

A. Energy-sensitive Cluster Formation

In this study, we use the same radio model as in [16]. This is a first-order radio model, and is used for the exchange of information between sensor nodes. An energy of 50 nJ/bit was used to operate the transmitter and receiver circuits. We used two types of amplifier values in different modes, that is, in free space or in multiple paths. The amplifier values were, respectively, set as 1 nJ/bit/m^2 and 1.3 nJ/bit/m^4. For our simulation, both free-space (d^2 power loss) and multipath fading (d^4 power loss) channel models were used, depending on the distance between the transmitter and the receiver. The transmission energy for two nodes is five times that of the receiving energy between the nodes, for a given signal-to-noise ratio. The parameters used for the energy-sensitive cluster formation are listed in Table 1. Factors, such as digital coding, modulation, filtering, and signal spreading, vary according to the electronic energy E_{elec}. The amplifier energies $E_{fs}d^2$ and $E_{mp}d^4$ depend on the distance between the transmitter and the receiver and the acceptable bit-error rate.

Table 1. Initial values for different variables used in the simulation

Parameters	Values
Initial Energy (E_{INI})	0.5 J
Transmitted Energy (E_{TX})	$5*10^{-7}$ J
Received Energy (E_{RX})	10^{-7} J
Data Aggregation Energy (E_{DA})	$5*10^{-8}$ J
Number of Rounds (R_{MAX})	1000–2000
Free-space Amplifier Value (E_{FS})	10^{-9} J/bit/m^2
Multipath Amplifier Value (E_{MP})	10^{-9} J/bit/m^4
Operating Energy (E_{TX}-elec=E_{RX}-elec=E_{elec})	50 nJ/bit

Cluster formation is initiated during the setup phase. We used simple logic to determine the locations of the nodes and the remaining energy of each node. For the first time, we developed our model without CHs. The BS saves all the information about the nodes. In the first round, the algorithm determines the amount of energy dissipated by each node to connect to the BS. The CHs for the second round are selected.

B. Total-energy-based CH Selection

The next round of CHs is selected using the probability function expressed in Equation (1). However, the total energy is the energy required for the network to run r rounds. A network is considered efficient when r is equal to

$R_{\max}$. $T(n)$ is updated using the remaining energy of the system in the LEACH network as follows:

$$T(n)=\begin{cases}\dfrac{p}{n-p\left((r+1)mod\left(\dfrac{n}{p}\right)\right)}+\displaystyle\sum_{i=1}^{r}\frac{(E_R)_i}{E_{INI}}\times\frac{E_{INI}-E_{TX}}{\sqrt{\dfrac{E_{FS}}{E_{MP}}}} & \text{if } n\in G\\ 0 & else\end{cases} \tag{1}$$

where E_R is the received energy, E_{INI} is the initial energy, E_{TX} is the transmitted energy, E_{FS} is the free-space amplifier value, and E_{MP} is the multipath amplifier value. The first term of Equation (3) calculates the threshold value depending on the number of nodes n and the number of CHs p. The value of p is 10% for group G. In the second term of Equation (3), we introduce the initial energy and remaining energy of the nodes, where the network is developed based on the threshold distance $\sqrt{\dfrac{E_{FS}}{E_{MP}}}$.

CH selection methodologies are performed in the steady-state phase, where clusters are formed according to the requirements. As the deployment of nodes is random, a given number of CHs may not be sufficient, or may be excessive in some cases. However, the location of the nodes is known in the second round, and several clusters are formed efficiently.

IV. DISCUSSION

The simulation results were obtained using MATLAB (MATLAB R2021a Academic Version). The PC used for the simulation had an i7-9700 KF processor (3.60 GHz).

Tables 2 and 3 present the input boundary conditions for the different methodologies. The number of nodes required to execute these algorithms is denoted by the dead nodes. At the end of a predefined number of information exchanges (rounds), the total energy saved by each algorithm is included.

Fig. 1 shows the number of live nodes associated with the different algorithms. The input boundary conditions for each algorithm were identical. In this case, 50 nodes each having an initial energy of 0.5 J were deployed in an area of 50*50 units. The number of rounds in which LEACH could exchange information with the BS was 600. Subsequently, the nodes did not have sufficient energy to perform the assigned tasks. The amount of energy used by LEACH exceeded the predefined energy (25 J). The energy used in LEACH was 25.8512 J over 600 rounds. Similarly, for PEGASIS and ALEACH, 39 and 31 nodes were involved in 700 and 900 rounds of information exchange using 25.0059 and 25.3551 J of energy, respectively. The proposed model had only one node with insufficient energy. The remaining 49 nodes shared an energy load of 10.1199 J, with no node spending more than 0.5 J.

Table 2. Comparison of different algorithms in different conditions

Observation	LEACH	PEGASIS	ALEACH	PROPOSED MODEL
Area	50*50	50*50	50*50	50*50
Nodes	50	50	50	50
BS	25,150	25,150	25,150	25,150
Rounds	1000	1000	1000	1000
Dead Nodes	50	39	31	1
Energy Saved	-0.8512	-0.0059	-0.3551	14.8801

Table 3. Comparison of different algorithms in different conditions

Observation	LEACH	PEGASIS	ALEACH	PROPOSED MODEL
Area	100*100	100*100	100*100	100*100
Nodes	100	100	100	100
BS	200,200	200,200	200,200	200,200
Rounds	2000	2000	2000	2000
Dead Nodes	90	78	87	55
Energy Saved	-2.4346	-0.5078	-1.5097	0.0249

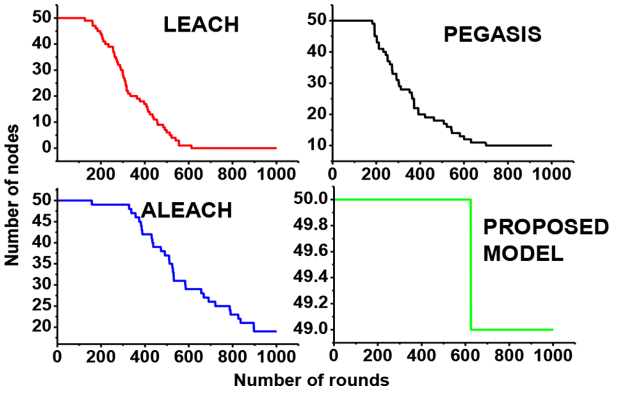

Fig. 1. Number of live nodes by the end of 1000 rounds of information exchange for 50 nodes

Fig. 2 shows the number of live nodes associated with the different algorithms. The input boundary conditions for each algorithm were identical. In this case, 100 nodes each having an initial energy of 0.5 J were randomly deployed in an area of 100*100 units. The number of rounds in which LEACH could exchange information with the BS was 609. Subsequently, the nodes did not have sufficient energy to perform the assigned tasks. Hence, the amount of energy used by LEACH exceeded the predefined energy (25 J). The energy used by 90 nodes in the case of LEACH was 52.4346 J for 609 rounds of data transmission to the BS (200,200) located outside the network. Similarly, for PEGASIS and ALEACH, 78 and 87 nodes were involved in 495 and 1315 rounds of information exchange using 50.5078 and 51.5097 J of energy, respectively. In our proposed model, only 55 nodes were out of energy.

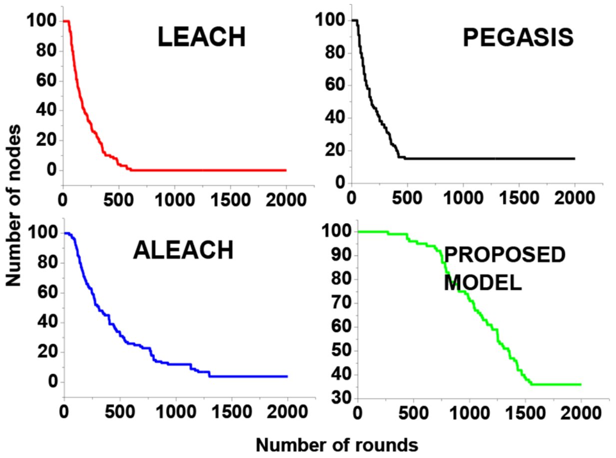

Fig. 2. Number of live nodes by the end of 2000 rounds of information exchange for 100 nodes

In Fig. 3, the results of PEGASIS are excluded because the BS is far from the nearest available node. Here, we compared our proposed model with the hierarchical methodologies. A larger number of nodes in the network indicates that more data are being transmitted to the BS. The BS was situated at the center of the network. The total energy of the system was 250 J. In our model, only 74 nodes were actively involved, as listed in Table 4.

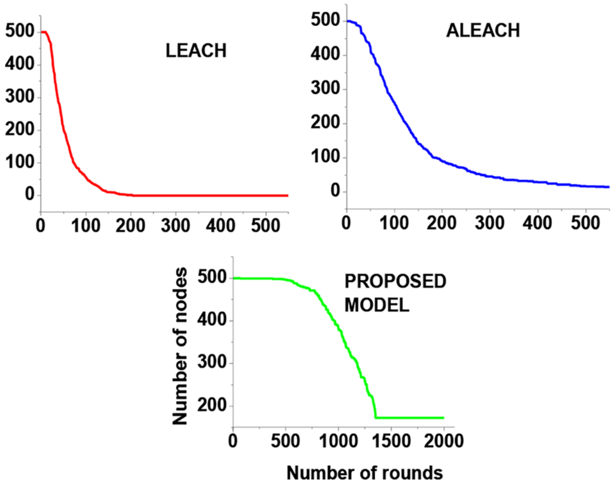

Fig. 3. Number of live nodes by the end of 2000 rounds of information exchange for 500 nodes

Based on the simulation results, the proposed LEACH was the best among the three methods. Even though PEGASIS is effective when the BS is situated far from the network, we could design a low-energy-consumption network under various conditions, such as the working area, the number of nodes, BS position, and the number of rounds of results.

Table 4. Comparison of different algorithms in different conditions

Observation	LEACH	ALEACH	PROPOSED MODEL
Area	100*100	100*100	100*100
Nodes	500	500	500
BS	50,50	50,50	50,50
Rounds	2000	2000	2000
Dead Nodes	362	404	74
Energy Saved	-42.5317	-38.2498	65.5780

V. CONCLUSIONS

LEACH is the basic concept behind the formulation of the proposed algorithm. Except PEGASIS, in which clusters are not formed, the two algorithms (ALEACH and our algorithm) discussed earlier are hierarchical methodologies derived from LEACH. Each methodology differs in terms of CH selection and data transmission rules. Our proposed model, based on energy-sensitive cluster formation and total-energy-based CH selection, demonstrated an excellent performance when compared with the existing LEACH-based methodologies. As energy is still consumed by the nodes in the idle and sleep states, our future work will incorporate different states and compute the results.

ACKNOWLEDGMENT

This Research was supported by Research Funds of Mokpo National University in 2022.

We would like to thank Editage (www.editage.co.kr) for English language editing.

REFERENCES

[1] A. S. Nandan, S. Singh, A. Malik, and R. Kumar, “A green data collection & transmission method for IoT-based WSN in disaster management,” *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25912-25921, Nov. 2021. DOI: 10.1109/JSEN.2021.3117995.

[2] A. Luntovskyy, T. Zobjack, B. Shubyn, and M. Klymash, “Energy efficiency and security for IoT scenarios via WSN, RFID and NFC: Invited Paper,” in *IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics*, Odesa, Ukraine, pp. 1-6, 2021. DOI: 10.1109/UkrMiCo52950.2021.9716591.

[3] K. Y. Bendigeri, J. D. Mallapur, and S. B. Kumbalavati, “Wireless sensor networks and its application for agriculture,” *Intelligent Data Communication Technologies and Internet of Things*, pp. 673-687, 2021. DOI: 10.1007/978-981-15-9509-7_55.

[4] G. Fan, R. Wang, H. Huang, L. Sun, and C. Sha, “Coverage-guaranteed sensor node deployment strategies for wireless sensor networks,” *Sensors,* vol. 10, no. 3, pp. 2064-2087, Mar. 2010. DOI: 10.3390/s100302064.

[5] S. K. Mohanty and S. K. Udgata, “SATPAS: SINR-based adaptive transmission power assignment with scheduling in wireless sensor

network," *Engineering Applications of Artificial Intelligence*, vol. 103, pp. 104313, Aug. 2021. DOI: 10.1016/j.engappai.2021.104313.

[6] E. Alzahrani, F. Bouabdallah, and H. Almisbahi, "State of the art in quorum-based sleep/wakeup scheduling MAC protocols for Ad Hoc and wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-33, Feb. 2022. DOI: 10.1155/2022/6625385.

[7] S. Diwakaran, A. Praghna, E. T. Reddy, and A. Sravani, "A clustering prediction model-based data collection for WSN," in *7th International Conference on Communication and Electronics Systems (ICCES),* Coimbatore, India, pp. 620-627, 2022. DOI: 10.1109/ICCES54183.2022.9835928.

[8] J. Vera-Pérez, J. Silvestre-Blanes, V. Sempere-Payá, and D. Cuesta-Frau, "Multihop latency model for industrial wireless sensor networks based on interfering nodes," *Applied Sciences*, vol. 11, no. 19, pp. 8790, Sep. 2021. DOI: 10.3390/app11198790.

[9] S. Singh and H. S. Saini, "Learning-based security technique for selective forwarding attack in clustered WSN," *Wireless Personal Communications,* vol. 118, no. 1, pp. 789-814, Jan. 2021. DOI: 10.1007/s11277-020-08044-0.

[10] S. Perumal, M. Tabassum, G. Narayana, S. Ponnan, C. Chakraborty, S. Mohanan, Z. Basit, and M. T. Quasim, "ANN based novel approach to detect node failure in wireless sensor network," *Computers, Materials and Continua,* vol. 69, no. 2, pp. 1447-1462, 2021. DOI: 10.32604/cmc.2021.014854.

[11] C. S. Fan, "Rich: Region-based intelligent cluster-head selection and node deployment strategy in concentric-based WSNs," *Advances in Electrical and Computer Engineering,* vol. 13, no. 4, pp. 3-8, 2013. DOI: 10.4316/AECE.2013.04001.

[12] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications,* vol. 1, no. 4, pp. 660-670, Oct. 2002. DOI: 10.1109/TWC.2002.804190.

[13] M. S. Ali, T. Dey, and R. Biswas, "ALEACH: Advanced LEACH routing protocol for wireless microsensor networks," in *International Conference on Electrical and Computer Engineering, IEEE,* pp. 909-914, 2002. DOI: 10.1109/ICECE.2008.4769341.

[14] K. Maraiya, K. Kant, and N. Gupta, "Efficient cluster head selection scheme for data aggregation in wireless sensor network," *International Journal of Computer Applications,* vol. 23, no. 9, pp. 10-18, Jun. 2011. DOI: 10.5120/2981-3980.

[15] S. Sumitha Pandit and B. Kalpana, "Advances in intelligent systems and computing," in Springer: *Soft Computing for Security Applications,* Singapore, vol. 1397, pp, 351-363, Oct. 2021. DOI: 10.1007/978-981-16-5301-8_27.

[16] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Maui, USA, pp. 10, 2000. DOI: 10.1109/HICSS.2000.926982.

Sagun Subedi

received his B.S. degree in electronics and communication engineering from Tribhuvan University, Pokhara, Nepal in 2015. Since 2015, he has been pursuing his M.S. and Ph.D. degrees in electronics engineering from Mokpo National University, Muan-gun, Korea (2022). His research interests include wireless sensor networks, digital communication, and optimization.

Sangil Lee

is currently working as a professor in the Department of Electronics Engineering at Mokpo National University, Korea. He was an assistant manager at LG DACOM, Korea from January, 1994 to August, 1997. He also worked as a researcher at Samsung Electronics and Telecommunications Research Institute from September, 2002 to August, 2003. He received his Ph.D. from the Department of Electrical Engineering at the University of Washington in 2002. His research interests include microwave engineering.