Supervised learning-based Lifetime Extension of Wireless Sensor Network Nodes

Jaewoong Kang, Jongmo Kim, and Mye Sohn*
Department of Industrial Enginnering, Sungkyunkwan University,
Suwan, Korea
{kjw1727, dignityc, myesohn}@skku.edu

Abstract

In this paper, we propose a new approach to increase the sustainability of the Wireless Sensor Network (WSN) nodes by extending their lifetimes. To do so, we attempt to find the optimal values of the collection interval and the transmission interval for each task that can maximize the lifetime of the WSN nodes by applying machine learning techniques. As a preprocessing for finding the optimal value of two parameters, we first determine the combination of nodes necessary to perform each task using the wrapper method. In addition, we applied Simulated Annealing (SA) to find the values of two parameter that lower power consumption without being significantly affected by the WSN's performance. To prove the superiority of the method, we perform two kinds of experiments. Finally, we prove the reduction of energy consumption using our framework.

Keywords: Sustainability, Wireless Sensor Network, Interval Determination, Machine Learning

1 Introduction

Many researchers are interested in increasing the sustainability of the Wireless Sensor Network (WSN) nodes by extending their lifetimes [19], [2], [1], [10], [15]. If the power supply from the sensor nodes goes down or becomes unstable, there is a problem with the accuracy of the sensors, leading to poor performance of the WSN. Since it causes a decrease in the performance of the WSN, extending of the battery lifetime has a direct impact on improving sustainability of the WSN. Improving sustainability issue can be tightly related to security in mobile application. Research for extending the battery lifetime are divided into two types: energy provision and energy consumption. The former attempts to extend the battery lifetime by continuously replenish the power consumed by the sensors [13]. To do so, the researcher has attempted to attach a power supply (e.g. solar charger or wind charger) to the sensors themselves or to supply power from other nodes (sensor) or external devices that have extra power [17]. The latter focuses on managing battery lifetime by adjusting the wake-up/sleep cycle of the sensors or by reducing the amount of data to be sensed [6]. In other words, it attempts to extend the battery life by reducing the load on the sensors. However, the research had no consideration for data collection intervals and transmission intervals, which are known to have a significant impact on extending the nodes lifetime [3], [11].

In this paper, we propose a method to find the optimal values of the collection interval and the transmission interval for each task that can maximize the lifetime of the WSN nodes by applying machine learning techniques. By applying machine learning methods, it is possible to use collected data without domain knowledge of the sensors and the WSN, unlike previous research. In addition, we attempt to find the optimal combination of sensors needed to perform each task to extend node lifetime. It is

Journal of Internet Services and Information Security (JISIS), volume: 9, number: 4 (November, 2019), pp. 59-67

*Corresponding author: InfoScience Laboratory, Department of Systems Management Engineering, Sungkyunkwan University, 2066 Seobu-ro, Jangan-gu, Suwon 440-746, Republic of Korea, Tel: +82-(0)31-290-7605

accomplished by supervised learning using wrapper method. As a result, it will be possible to put other nodes idle except for the ones that will control the intervals. So, it can contribute to extending the lifetime of the battery. As the WSN grows, the effectiveness of this method is expected to increase. We applied the Simulated Annealing (SA) technique to find the optimal values of the collection interval and the transmission interval of task-relevant nodes. Since the SA technique is a method that can be applied to simple problems among many meta-heuristic methods, it is the most suitable method for our research with two parameters.

This paper is organized as follow. In Section 2, we review the related research. Section 3 offers detailed descriptions about the overall framework. In Section 4, experimental results are suggested to demonstrate the effectiveness of the framework. Finally, Section 5 presents the conclusions and further research.

2 Related Works

As mentioned, state, research for extending the battery lifetime are divided into two types: energy provision and energy consumption. Energy provision improves the sustainability of the WSN by continuously charging the power consumed. To do so, it installs solar chargers or wind chargers. In this light, this method is commonly used in large scale WSNs. Energy consumption improves the WSN's sustainability by managing the remaining battery charge by either finding the optimal path for data transmission in the WSN or by forcing the use of other sensors through periodic topology changes. Table 1 summarizes the pros and cons of both methods.

Table 1: Summary of energy provision method and energy consumption method

Methods	Characteristics	Remarks	References
Energy Provision	Pros: Semi-permanent using through continuous charging, High utility in a wide range of spaces	Charger using natural energy	[20], [14], [16]
	Cons: Additional maintenance of charging equipment is required, additional cost of installation of charging equipment is required, cost increases as sensors increase	Direct energy supply from external nodes	[8], [7]
Energy Consumption	Pros: Easily extend the life of wireless sensor network, Formulated model allows definitive solution selection	Finding the Best Route for Nodes Transferring Data	[5], [4], [12]
	Cons: There may be no solution that satisfies the condition	Full power management through coercion of the overall topology	[9], [18]

However, although the value of the collection intervals and the transmission intervals of the sensor data is very important, none of these methods were concerned. In other words, in previous methods, instead of determining this parameter values, the WSN, which requires the collection of real-time data, arbitrarily shortens the two parameters, and sets them arbitrarily long in the rest. So, in order to deal with the sensor's efficiency, which is not covered in previous methods, we propose a method to adjust the sensor's optimal collection cycle and optimal transmission cycle to increase the sustainability of the WSN.

3 Overall Frameworks

In general, data collected from multimodal sensors are represented as vectors of different dimensions, depending on the sensing period of the sensors. In order to analysis of the multimodal sensor data, it should be transformed into tabular dataset consisting of vectors of the same dimension. At this time, the tabular dataset is generated by interpolation of the sensor data. In this paper, we performed the interpolation with respect to the highest dimensional data in order to minimize the data loss. Finally, we obtain the tabular data (T) as follows.

Definition 1. T is a $n \times (m+1)$ matrix and simply represented as follows.

$$T = (d^{11} \ldots d^{ab} \ldots d_{la})$$

where d^{ab} is the n dimensional column vector of b^{th} data field of a^{th} sensor $(b = 1, 2, \ldots, m)$. At this time, n is the value of the highest dimension of sensor data and m is the number of the data fields of the sensors. $d_{la}$is label vector, which is used to instance segmentation. It is augmented by human.

Since the multimodal sensor data is time-series data, the tabular data must be separated into a set of instances, which have a constant size. There are two ways to determine the size of an instance: label-based and task time-based. In this paper, we use the method of the label-based for learning the best classifier and the task time-based for determining the best task time. The instance set is defined as follows.

Definition 2. Instance set (IS) is the partitioned data set from T and composed of the instance (Ins_i) as expressed below.

$$Ins_i = (d_((s_i, f_i))^{11} \ldots d_((s_i, f_i))^{ab} \ldots), \text{where } d_((s_i, f_i))^{ab} = (d_(s_i)^{ab}, \ldots, d_(f_i)^{ab})^T$$

where $d_((s_i, f_i))^{ab}$ is the $f_i - s_i$ dimensional column vector of b^{th} data field of a^{th} sensor from s_i^{th} row to f_i^{th} row $(b = 1, 2, \ldots, m, i \in N)$.

The instances are used to the training of the classifiers after transforming to the feature vector with the labels.

3.1 Selection of Sensor Subset using Wrapper Method

In general, the WSN built to perform specific tasks contains a variety of sensors needed to achieve its purposes [20]. However, some sensors in the WSN may be redundant, and not all WSN sensors are used to perform specific tasks. In this light, if we find the WSN sensors that do not affect the execution of the tasks, and idle them, it can make a significant contribution to WSN energy saving.

This module is performed to find a subset of the sensors that can perform tasks successfully while minimizing the energy consumption of the WSN. To find the best subset of sensors, we use a wrapper method known to have the highest classification performance. But the wrapper we used is the main purpose of finding the best sensor set, not the best feature set. The classifier is then trained using all the features of all the sensors in the best subset of the sensors. The difference between the method of finding sensor subsets using the conventional wrapper method and the proposed method is shown in Figure 1.

As shown in Figure 1, the proposed method in this paper has much less computational burden than the conventional method. In general, the number of the features that a sensor can have is very large. As the

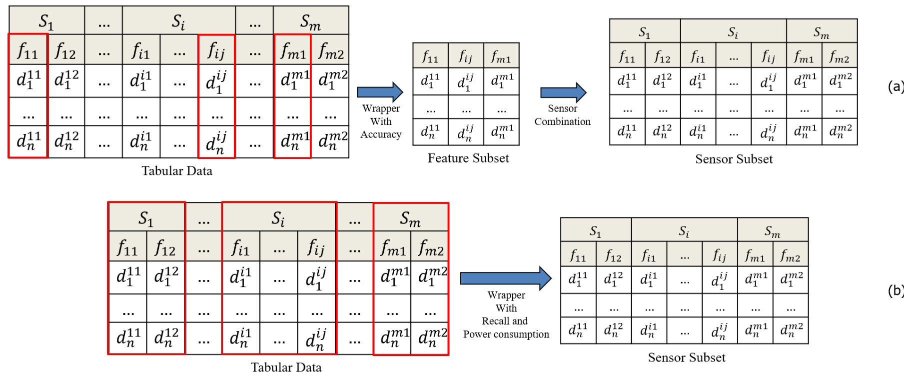

Figure 1: Comparison of sensor selection method between (a) conventional method and (b) proposed method

WSN grows, the number of feature combinations that the wrapper method must search to find the optimal feature subset increases exponentially. On the other hand, even though the size of the WSN increases, the number of sensors included therein is relatively small compared to the number of features. By reducing the search space by grouping multiple data fields in the sensors, it can contribute to lightening the burden of the wrapper to find the optimal sensor subset.

In order to develop the wrapper, which can find the best sensor subset for the tasks, we used recall and power consumption instead of accuracy, which is widely used to evaluate the superiority of the sensor itself. The reason we used recall is that the frequency of occurrences varies from task to task. In addition, we perform min-max normalization on the values of the power consumption for the scaling with the recall. As a result, we derive the energy aware score of the i^{th} as follows.

$$EAS_i = Recall - pc_i/(2(pc_{max} - pc_{min}))$$

where pc_i is power consumption of the i^{th} sensor. pc_{max} and pc_{min} are maximum and minimum power consumption of the sensors, respectively.

To find the best sensor subset using the proposed wrapper method, we first need to determine the representative task time for the tasks with various execution time. However, it is very difficult to determine by theoretically because each task has a different execution time and a slight deviation may occur even in the task time. Therefore, we determined the optimal task time (a_t) using experimental method. The method is summarized in Figure 2. In general, training a classifier has a heavy computational burden, while performing a task using a trained classifier has a very light computational burden. Therefore, a_t can be determined through repeated experiments without computational burden. Using the a_t, we generate the instances and perform the evaluation of the sensors using the instances and EAS_i. The learning process is summarized in Figure 3.

3.2 Simulated Annealing-based Interval Optimization

In the previous module, we determine the best sensor subset, which can contribute the energy saving of the WSN. In order to save the energy by minimizing the unnecessary operation of the sensor, we attempted to find and manipulate the collection interval and the transmission interval of the sensors. To do so, we propose the optimization method based on Simulated Annealing (SA), which is an efficient method to find approximate global optimum in datasets where it is difficult to collect precise values due to noise problems such as sensor data.

```
TB: Tabluar Data
IS: Instance Set
BCL: Best Classifier using IS
CT: Candidate Task Time Set
a_t: Best Task Time

Task Time Determination Process
  IS = Label-based Instance Segmentation Function (TB)
  BCL = Learning Best Classifier using IS
  CT = [for i in range(min(IS), max(IS))
  RecallList = []
  For all CT
    IS = Task time-based Instance Segmentation Function (TB, CT)
  Recall = Recall of BCL using IS
  RecallList.append(Recall)
  End For
  a_t = CT[which.max(RecallList)]
  Return a_t
End Process

Label-based Instance Segmentation Function (TB)
  IS = []
  k = 0
  For i in range(len(TB))
    if Label[i]!=Label[i+1]
       Ins = TB[k:i]
       k = i
    end if
    IS.append(Ins)
  End For
  Return IS
End Function

Task time-based Instance Segmentation Function (TB, CT)
  IS = []
  For i in range(floor(len(TB)/CT))
    Ins = TB[i×CT:(i+1)×CT]
    IS.append(Ins)
  Return IS
End Function
```

Figure 2: Task time determination algorithm

```
S: Set of Sensors
S_i: i^th sensor in S
pc_i: Power consumption of i^th sensor
pc_max, pc_min: Max, Min value of all PC_i
θ: Recall of classifier with all sensors

Wrapper-based Sensor Selection Process
  ScoreList= []
  SensorSubset = []
  For all i
    temp = Recall of Classifier(S_i)
    EAS_i = temp − pc_i / (2(pc_max − pc_min))
    ScoreList.append(EAS_i)
  End for
  For j in range(len(ScoreList))
    SensorSubset.append(S[which.max(ScoreList)])
  End for
  For k in range(len(S))
    recall = Recall of Classifer(SensorSubset[0:k])
    If recall > θ
      SensorSubset = SensorSubset[0:k]
      break
    End if
  End for
  Return SensorSubset
End Process
```

Figure 3: Wrapper-based sensor subset generation algorithm

The objective function for SA means to minimize the power consumption (P(CI,TI)) via shortened the collection interval and the transmission interval. It is as follows.

$$\min P(CI,TI) = \min \sum_{i=1}^{m} \left(\frac{\alpha_i}{CI_i} + \frac{\beta_i}{TI_i}\right)$$

where α_i is the average power consumption per a single collection of i^th sensor, β_i is the average power consumption per a single transmission of i^th sensor. CI_i is the collection interval and the transmission interval of i^th sensor to be determined, respectively. At this time, we identified the quenching conditions for performing SA as follows.

$$Pr(e,e',T) > random(0,1) \ (1)$$
$$3 \times min(CI_i) < \max(TI_i) < a_t \ (2)$$
$$Acc_{epoch} \geq \theta_acc * k \text{ and } Recall_epoch \geq \theta_rec * k \ (3)$$

where $e = P(CI,TI)$ and $e' = P(CI',TI')$. (CI',TI') is the random neighbor of (CI,TI), and $Pr(e,e',T) = exp(-(P'-P)/T)$ is acceptance probability function. In condition (2), θ_{acc} and θ_{rec} are accuracy and recall when classifying using default CI and default TI of the sensors. Acc_{epoch} and $Recall_{epoch}$ are accuracy and recall when classifying using CI_{epoch} and TI_{epoch} of the sensors for each epoch. k is an arbitrary constant for controlling the performance of (CI',TI').

The first one is the default condition of SA. Constraint (2) defines the appropriateness of the interpolation method for converting collected sensor data into the tabular data. If the performance of the WSN derived using the optimal solution of the SA is significantly lower than the existing θ_acc and θ_rec, the optimal CI and TI values obtained through the SA may not be meaningful. To solve the problem, we append the constraint (3) which can reduce power consumption by changing the CI and the TI values without affecting the WSN performance. However, unlike conventional SA, the proposed method has various variables and constraints. So, if we used the temperature drop condition $\frac{T_max}{(T+1)}$, the speed of finding the optimal value can be slowed down. To speed up finding the optimal value, we used $\frac{T_max}{(T+1+mo)}$ (mo:momentum) which was further considered as the previous direction of movement.

The overall procedure is summarized in Figure 4. It composed of two modules: Wrapper-based Sensor subset Selection Module (WSSM) and Simulated annealing-based Intervals Optimization Module (SIOM).

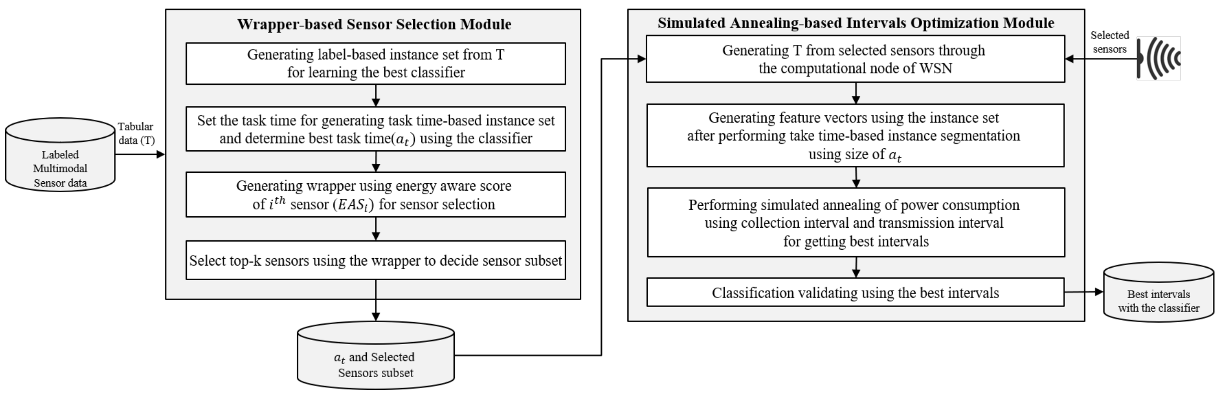

Figure 4: Overall framework of proposed method

4 Experiments

In this paper, we performed two experiments under the following environments. The sensor network consists of seven sensors, and the specifications of each sensor are shown in Table 2. These sensors were attached to a washing machine to collect data at the beginning, middle and end of dehydration.

Table 2: Specifications of sensors in the sensor network

Sensor	Hygrometer		Accelerometer- Gyrosensor							Ultrasonic	Infrared		Audio sensor
Data Field	Ambient Temp.	Ambient Humidity	AccX	AccY	AccZ	GyrX	GyrY	GyrZ	Temp	Distance	Surface Temp.	Target temp.	Volume
Default Sampling Rate (Hz)	1		1							1	1		1
Power Consumption (mA)	0.5		3.7							15	2		0.5

Experiment 1. This experiment was performed to compare the power consumption and execution time of the proposed method with the conventional method to find the best subset of sensors. As a result of the experiment, we can confirm that the proposed method not only finds the sensor with lower power consumption than the conventional method but also selects the sensor faster. The experimental results are summarized in Figure 4.

Experiment 2. This experiment was conducted to confirm the effectiveness of the momentum-constrained temperature drop condition in the proposed method. The experimental results are summarized in Figure 5.

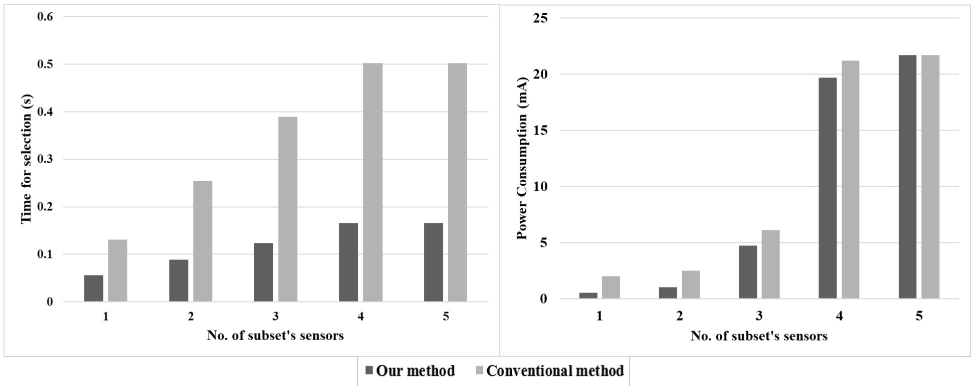

Figure 5: Comparison of execution time and power consumption

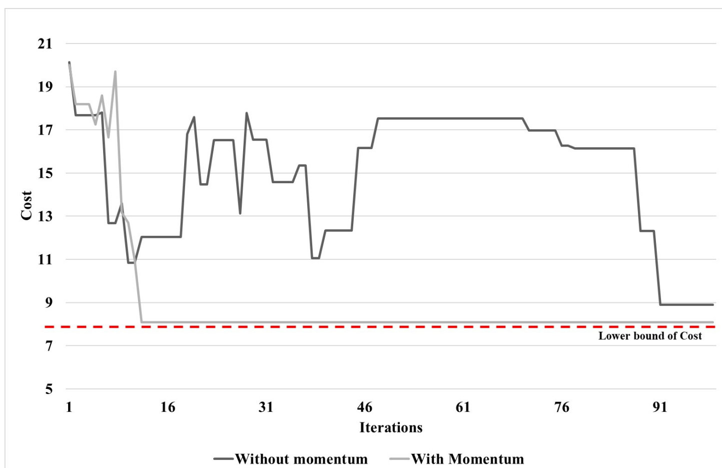

Figure 6: Results of the simulated annealing with and without momentum for 100 iterations

5 Conclusion and Further Research

In this paper, unlike the conventional approach, we propose a novel method to control the power consumption of the battery by adjusting the collection interval and transmission interval of the sensors. In this way, we increased the sustainability by reducing power consumption at the expense of WSN's task classification performance. Our contribution has been to produce good results by attempting to adjust the collection interval and transmission interval of the sensors, which were not used by conventional methods.

The limitations are as follows. First, the hardware characteristics of the sensors were not properly considered. In general, a sensor with high power consumption has a low noise, and thus a sensor with high power consumption is often used when a task that needs to be complicated by WSN needs to be classified. We attempted to solve this problem by using recall instead of accuracy of the task classification. Second, the interaction between the selection of the sensor and the determination of the period of the sensor was not considered. Our method has a limitation that it is not possible to accurately analyze the close interaction between the two by deciding one thing first and the latter one. Therefore, future research aims to conduct research that overcomes existing limitations.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2016 R1D1A1B03932110).

References

[1] M. H. Alsharif, S. Kim, and N. Kuruoğlu. Energy harvesting techniques for wireless sensor networks/radio-frequency identification: A review. *Symmetry*, 11(7):865, July 2019.

[2] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella. Energy conservation in wireless sensor networks: A survey. *Ad hoc networks*, 7(3):537–568, May 2009.

[3] K. Choi, S.-H. Lim, and J.-H. Kim. Cost-effective monitoring algorithm for cyber-physical system platform using combined spatio-temporal model. *The Journal of Supercomputing*, 74(8):3537–3548, August 2018.

[4] R. Doost, K. R. Chowdhury, and M. Di Felice. Routing and link layer protocol design for sensor networks with wireless energy transfer. In *Proc. of the 2010 IEEE Global Telecommunications Conference (GLOBECOM'10), Miami, Florida, USA*, pages 1–5. IEEE, December 2010.

[5] O. Dousse, P. Mannersalo, and P. Thiran. Latency of wireless sensor networks with uncoordinated power saving mechanisms. In *Proc. of the 5th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc'04), Tokyo, Japan*, pages 109–120. ACM, May 2004.

[6] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. In-network aggregation techniques for wireless sensor networks: a survey. *IEEE Wireless Communications*, 14(2):70–87, April 2007.

[7] J. Feng, Y.-H. Lu, B. Jung, D. Peroulis, and Y. C. Hu. Energy-efficient data dissemination using beamforming in wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 9(3):31, May 2013.

[8] B. B. Haro, S. Zazo, and D. P. Palomar. Energy efficient collaborative beamforming in wireless sensor networks. *IEEE Transactions on Signal Processing*, 62(2):496–510, January 2013.

[9] H. Karl and A. Willig. *Protocols and architectures for wireless sensor networks*. John Wiley & Sons, April 2007.

[10] J. A. Khan, H. K. Qureshi, and A. Iqbal. Energy management in wireless sensor networks: A survey. *Computers & Electrical Engineering*, 41:159–176, January 2015.

[11] S. W. Lee and S.-H. Lim. Analytic model for the power-optimal data transmission interval of wireless sensors in internet of things. *Journal of Digital Contents Society*, 19(7):1373–1379, July 2018.

[12] C. E. Perkins and E. M. Royer. Ad-hoc on-demand distance vector routing. In *Proc. of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (MCSA'99), New Orleans, Los Angeles, USA*, pages 90–100. IEEE, February 1999.

[13] M. Rahimi, H. Shah, G. S. Sukhatme, J. Heideman, and D. Estrin. Studying the feasibility of energy harvesting in a mobile sensor network. In *Proc. of the 2003 IEEE International Conference on Robotics and Automation (ROBOT'03), Taipei, Taiwan*, volume 1, pages 19–24, September 2003.

[14] S. Sudevalayam and P. Kulkarni. Energy harvesting sensor nodes: Survey and implications. *IEEE Communications Surveys & Tutorials*, 13(3):443–461, July 2010.

[15] Y. K. Tan and S. K. Panda. Review of energy harvesting technologies for sustainable wireless sensor network. *Sustainable wireless sensor networks*, pages 15–43, December 2010.

[16] J. Varley, M. Martino, S. Poshtkouhi, and O. Trescases. Battery and ultra-capacitor hybrid energy storage system and power management scheme for solar-powered wireless sensor nodes. In *Proc. of the 38th Annual Conference on IEEE Industrial Electronics Society (IECON'12), Montreal, Quebec, Canada*, pages 4806–4811. IEEE, October 2012.

[17] H. Wang, N. Agoulmine, M. Ma, and Y. Jin. Network lifetime optimization in wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 28(7):1127–1137, September 2010.

[18] Y. Xu, J. Heidemann, and D. Estrin. Geography-informed energy conservation for ad hoc routing. In *Proc. of the 7th annual international conference on Mobile computing and networking (MobiCom'01), Rome, Italy*, pages 70–84. ACM, July 2001.

[19] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. H. Hanzo. A survey of network lifetime maximization techniques in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, 19(2):828–854, January 2017.

[20] P. Zhang, G. Xiao, and H.-P. Tan. A preliminary study on lifetime maximization in clustered wireless sensor networks with energy harvesting nodes. In *Proc. of the 2011 8th International Conference on Information, Communications & Signal Processing (ICICS'11), Singapore, Singapore*, pages 1–5. IEEE, December 2011.

Author Biography

Jaewoong Kang is in Ph. D. course in the department of Industrial Engineering at Sungkyunkwan University. He received his bachelor's degree from a Sungkyunkwan University. His main interests CNN-based deep learning, pattern recognition, and machine learning.

Jongmo Kim is in Ph. D. course in the department of Industrial Engineering at Sungkyunkwan University. He received his bachelor's degree from a Sungkyunkwan University. His main interests are ontology, semantic web, Linked Open Data, web service composition, and Web-of-Things.

Mye Sohn is a professor in the Department of Systems Management Engineering at Sungkyunkwan University. She received her MS and Ph. D. from the Korea Advanced Institute of Science and Technology (KAIST). Her main interest is the machine learning, ontology, Web-of-Things, Semantic Web, and so on.