

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION



Data Fairness Transmission and Adaptive Duty Cycle through Machine Learning in wireless Sensor Networks

Junheon Jeon^a, Hyunjoo Park^{b,*}

^a Department of Electronic Engineering, Sangmyung University, Cheonan, 31066, Republic of Korea ^b Department of Smart Information and Telecommunication Engineering, Sangmyung University, Cheonan, 31066, Republic of Korea Corresponding author: *cathy2369@smu.ac.id

Abstract— In this paper, we propose the data fairness transmission and adaptive duty cycle through machine learning in wireless sensor networks. The mechanism of this paper is mainly composed of two parts. The proposed mechanism is based on the sleep-wake structure, which is one of the methods to increase the lifespan of the entire network by efficiently using the energy of the nodes. The first is a mechanism to support priority and data fairness. To this end, data input to the node is divided into priority classes according to transmission urgency and stored. Introduces the concept of cross-layer to rearrange data destined for the same destination. In addition, we propose a fair data transmission mechanism that allows even low-priority data to participate in transmission after a certain period. The second is an adaptive duty cycle mechanism through machine learning. For this purpose, public data related to forest fires are collected. The collected data is refined into data for each forest fire location and data for each forest fire time. For the refined data, an SVM (Support Vector Machine) model of supervised learning is used for machine learning, and a mechanism for adaptively adjusting the duty cycle of each node through the trained model is proposed. The computer language used for machine learning is Python language, and Google's Psychic Learn is used for the machine learning library. It was compared with the existing MAC protocol for evaluation, and it was confirmed that excellent energy efficiency results were obtained.

Keywords—Fairness transmission; duty cycle; machine learning; support vector machine; wireless sensor networks.

Manuscript received 17 Oct. 2021; revised 10 Jan. 2022; accepted 8 Mar. 2022. Date of publication 30 Jun. 2022. International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

I. INTRODUCTION

Since wireless sensor networks (WSNs) are batterypowered environments, efficient energy use is very important. Most wireless sensor networks mainly collect and transmit only basic information such as temperature, humidity, and pressure. However, in recent years, wireless sensor networks have been used in more diverse applications. Also, the requirements are different depending on the application environment. There may be an environment that requires reliability, an environment that requires accuracy, and an environment that requires real-time. Therefore, in the wireless sensor network, it is necessary to support different QoS according to the application environment [1], [2]. The development of wireless sensor networks is attracting attention in various places such as the military field, the environment field, the medical field, the transportation field, and the field of factory industrialization. Also, with the development of science and technology, the price of hardware such as CMOS cameras and microphones is getting smaller and cheaper. Therefore, it is necessary to consider the application of data with various characteristics, such as multimedia transmission, rather than staying on the application of basic information transmissions such as current temperature, humidity, pressure, and location [3], [4].

Wireless sensors are randomly distributed in a specific area to be measured, and these nodes form a network by themselves. A sensor node may also transmit basic and periodic data such as temperature and humidity. In addition to the periodic data, there may be event (image, video) data exceeding a threshold. As data that reports the temperature and humidity of the direct area, the importance of urgency is reduced, but if the transmission is not performed for a long time, the base station (BS) can determine that the sensor or network is abnormal. In addition, in the case of event data, most of the urgency data that can be generated when a temperature higher than the reference temperature is detected as data exceeding a threshold value. Such data should be transmitted without delay as much as possible because there is a requirement for the delay. However, in the case of event data, most of the event data is burst data, and once it occurs,

it is continuously and massively generated, so it can continuously occupy traffic in the wireless sensor network considering priority. Meanwhile, in the case of multimedia data, a periodic image of the surrounding environment may be received, and when event data is generated, an image may be received from a sensor adjacent to the event data generating sensor. In the field of environmental monitoring, it is difficult to expect proper transmission under the conditions required for transmission of data other than event data with the wireless sensor network MAC (Medium Access Control) protocol, which gives only priority, which has been researched in the past. In addition, various variables according to wildfires are trained on an artificial intelligence machine learning model. A suitable duty cycle can be predicted as a learning result. Through this, energy efficiency can be promoted.

Research conducted on the MAC layer of a wireless sensor network can be divided into three major categories. It can be divided into a random-access method, a slotted-access method, and a hybrid method. The first random approach is contention-based. The first random-access method is a contention-based method and typically includes a Carrier Sense Multiple Access (CSMA) method [5]. The CSMA method is a method in which a sensor node with data to transmit detects a channel in advance before transmitting data, delays transmission for a certain period of time if the channel is in use and transmits data only when the channel is idle. This method has the advantage of being easy to implement. However, the disadvantage of this transmission method is that collision may occur when access is made from two or more sensor nodes at the same time, and transmission efficiency decreases when there is a lot of data to be transmitted. Representative protocols using this method include S-MAC (Sensor-MAC) [6], [7] and T-MAC (Timeout-MAC) [8]. fig. 1 shows the S-MAC protocol.



Fig. 1 S-MAC (Sensor-MAC) protocol

S-MAC is one of the competition-based representative protocols, and the biggest feature is to save energy through periodic sleep-awake. fig. 4 shows the data transmission process of S-MAC protocol. S-MAC periodically repeats listen and sleep to reduce energy consumption when the sensor node is not a receiving node. The listen section consists

of a synchronization period and a data transmission section. As shown in the figure, node A broadcasts its own schedule information during the synchronization period. It waits some time to avoid collision and then sends a Request-To-Send (RTS). Node B sends a CTS (Clear-To-Send) packet to node A in response to the RTS packet, and data reception is performed during the sleep period. Node C knows that Node B is the receiver by the CTS packet of Node B and enters sleep mode in the sleep period to reduce energy consumption. However, since S-MAC has a fixed duty cycle as above, when data transmission does not occur, it consumes energy with a fixed listen, and when data transmission occurs, transmission is delayed with a fixed sleep.) has the disadvantage of being. Fig 2 shows the T-MAC protocol.



Fig. 2 T-MAC(Timeout-MAC) protocol

T-MAC is a competition-based protocol that supplements the problems of the fixed duty cycle of the S-MAC protocol. The figure shows the basic operation of T-MAC. T-MAC reduces the unnecessary idle listening period by actively adjusting the active listening period according to network traffic. S-MAC always maintains a wakeup state for a fixed active time, but T-MAC uses TA (Activity Time-out value) to immediately enter a sleep state when there is no transmission data for a certain period of time, thereby reducing energy consumption. However, T-MAC has the advantage of reducing energy consumption compared to S-MAC but has the same delay problem as S-MAC.

The second slot approach is a scheduling-based method, and there is a representative time division multiple access (TDMA) method [9]. In the TDMA scheme, time is divided into frames, and one frame is composed of several slots. One slot is enough time for a node to transmit one data packet.

In this method, since several sensor nodes are allotted different time slots for transmission, this method avoids collisions between nodes and transmitting data. However, the disadvantage of this method is that strict synchronization is required between the central sensor node and the sensor node, and resources may be wasted due to the possibility of empty slots for transmission in a network with low transmission. Representative protocols belonging to this include BMA (The Bit-Map-Assisted) [10] and EC-TDMA (An Energy-Efficient TDM A) [11].

Third, a hybrid method is made by combining the advantages of the above two methods. A characteristic of this method is that energy can be effectively used by responding differently to the transmission method according to transmission traffic.

The SVM (support vector machine) algorithm is one of the most used models in artificial intelligence machine learning. The SVM algorithm finds a hyperspace that can classify data composed of two classes as one of the widely used security measures for binary classification problems. The SVM model can be applied to both classification and regression and has the advantage of high accuracy. It is also not sensitive to noise. However, if there are many attribute values of the input data, it has a disadvantage in that it takes a long time to learn. In addition, it is difficult to interpret, and it is difficult to find the reason for the derivation of the result.

II. MATERIALS AND METHOD

A. QAML-MAC Protocol

There are two methods to support QoS in QAML-MAC [12]. First, it uses the cross-layer concept to reorder and transmit data destined for the same destination. Second, the priority of data is classified, stored in different classifiers, and transmitted.



Fig. 3 QAML-MAC (A QoS Aware Multi-layer MAC) protocol

fig. 3 shows the buffer allocation according to the priority of the received data. The figure shows that data received from neighboring nodes are classified according to transmission urgency and stored in a temporary buffer. Data is transmitted according to priority, and data directed in the same direction are rearranged and transmitted.

Here, for convenience, data are divided into high-priority data and low-priority data. Examples of high-priority data include multimedia application-related data or an abnormal measurement value higher than a predetermined threshold. These high-priority data have the property of being delaysensitive. Each node creates a temporary buffer, rearranges the data order according to the destination, and preferentially transmits data destined for the same destination. Each node has several buffers with a priority, and after the high-priority data transmission is completed, the low-priority data transmission is performed. In QAML-MAC, to use energy efficiently, the listening section is multi-layered without overlapping. Fig. 4 shows energy savings due to tiering of the listening section.

The QAML-MAC method will be usable in an environment where it does not matter, even if data with low priority is not transmitted. Event data exceeding a predetermined threshold has a burst property, and such burst data is continuously transmitted until the event disappears, causing a phenomenon of occupying network traffic.

However, for the QAML-MAC protocol, only the conditions of a general wireless sensor network environment were considered. In this paper, we propose a MAC supporting QoS by transmitting a portion of data that cannot be transmitted due to high-priority data in a wireless sensor network in a specific environment. This method can also be applied in a special environment that is out of the existing general wireless sensor network environment. We also propose increasing energy efficiency by having an applied duty cycle using AI machine learning.



B. Proposed MAC Protocol

As it is used in various fields, the types of data to be transmitted are also diversifying. These data may be specifically classified into multimedia data, event data exceeding a predetermined threshold, periodic data such as temperature and humidity location, and the like. However, when these various data are transmitted from the source node to the destination node, different quality of service (QoS) may be required. Among them, multimedia data is required to be transmitted without delay, even if there is some loss due to its characteristics. However, on the contrary, most measurement data, i.e., event data, exceeding a predetermined threshold value is data requiring urgency and has a request to be transmitted with a minimum delay. Finally, data that occurs periodically is often reported once at a predetermined time. In most cases, delay or loss of such data is not a big problem. However, even for periodically occurring data that does not require urgency, if it is not transmitted for a long time, it may cause a problem in which the transmission connection is disconnected or recognized as a network error. Fig. 5 shows the difference between the QAML-MAC protocol and the proposed method.

In the above figure of fig. 5, it is assumed that there is Node 1, Node 2, and Node 3 adjacent to each other among several nodes. It is assumed that the purpose of these sensors is to report temperature or humidity periodically, to report at once when event data exceeding a threshold value occurs, and to transmit an image to convey the situation from a nearby sensor when event data occurs. In the above figure, the relay node transmits the data received from the sensor node to the BS (base station) in TDMA method. That is, a case in which

a measured value exceeding a predetermined threshold is measured in sensor node 1 is considered.



Fig. 5 Comparison with QAML-MAC protocol

Consider the existing QAML-MAC protocol. The relay node will receive event data from sensor node 1, image data for a nearby area from adjacent sensor node 2, and periodic data from sensor node 3. Event data is burst data, and event data will be continuously transmitted to the relay node until the local temperature falls within the predetermined threshold range, and it will be assigned to and transmitted continuously. In this case, the surrounding sensor nodes 2 and 3 transmit data to the relay node, but the relay node stores it in a lowpriority class, and the transmission is pushed ahead of the priority, so the transmission will not be performed.

In this paper, we intend to provide quality of service (QoS) for data with lower priority by securing the transmission of some data from sensor nodes 2 and 3 even in the above situation. fig. 5 As shown in the figure below, there are adjacent sensor nodes 1, 2, and 3, and event data is transmitted from sensor node 1. Sensor node 2 takes an image and sends it as data, and sensor node 3 sends periodic data. The relay sensor node classifies the transmitted data into classes and stores them in different buffers. Different buffers have different weights. Node 1 event data is stored in a buffer with a weight of 7. Node 2 image data is stored in a buffer with

weight 2, and node 3 periodic data is stored in a buffer with weight 1. In weighted scheduling, a TDMA slot is allocated for transmission in consideration of the weight of each buffer.

At this time, when there is no data in the buffer having a weight of 20% and a weight of 10%, all TDMA slots are actively allocated with the data of the buffer having a weight of 70%. If there is no event data, only image data and periodic data from the sensor node will be transmitted to the relay node. In this case, data is stored in a buffer with a weight of 20% and a weight of 10%, and the data destined for the same destination is rearranged and transmitted with priority. fig. 6 shows that the received data is allocated to a weighted buffer through a class buffer and actively allocated to TDMA by weight and transmitted.



Fig. 6 Proposed protocol

In order to classify the data class, each sensor node gives priority according to the value measured by itself or the transmission urgency of the data received from the neighboring node [12]. Transmission urgency u is packet urgency (Cc) according to the importance of the application layer, transmission hops (Hc) indicating the retransmission cost, residual energy (Ec), and the proportional load of the queue (queue's) proportional load), and the relation is as follows.[13,14]

$$\mu = \frac{1}{4} \times \frac{E_c}{E_{max}} + \lambda + \frac{C_c}{C_{max}} + \frac{H_c}{H_{max}}$$
(1)

Here, the proportional load on the queue is

$$\lambda = \frac{1}{2} \times \left(\frac{\sum_{i=1}^{n} \omega_i Q_c(i)}{\sum_{i=1}^{n} \omega_i Q(i)} + \max_{k=1\dots n} \left(\frac{Q_c(k)}{Q(k)} \right) \right)$$
(2)

In the above equation, Emax is the initial energy, Hmax is the maximum number of vacancies, and Cmax is the urgency level of the packet. In addition, n is the number of queues, Wi is the service weight of the i-th queue, and Qi and Qc(i) are the maximum and instantaneous load values. As shown in equation (1) above, the transmission urgency u is affected by energy, queue load, packet urgency, and the number of hops. Therefore, it is possible to increase the accuracy of packet transmission by using different weights for each of the above items according to applications and environments. As the value obtained through the above equation, the equations for the priority value p of the sensor node and the collision time are as follows [15], [16].

$$\rho = \max(\lfloor (1-\mu) \times N \rfloor), N-1)$$
(3)

$$t_{CT} = \rho \times CW + rand(CW) \tag{4}$$

In these equations, the four values are the values obtained in the above equation (1), N is the priority level value, and d represents the size of the collision window. Each node puts the measured data or received data into the corresponding FIFO queue according to the priority determined by Equation (3) above. Due to the characteristics of the FIFO queue, data that arrives first is served first.

The weight fair queuing (WFQ) algorithm is used in weight schedule. WFQ is a scheduling technique that evenly allocates network traffic considering weights. Data with weights of 70%, 20%, and 10% are allocated time slots through weighted scheduling, and the method of calculating the allocated packet's transmission time and completion time is as follows [17], [18].

$$S_{i,n} = \max(v(A(t)), f_{i,n-1})$$
(5)

$$f_{i,n} = S_{i,n} + \frac{L_{i,n}}{r_i} \tag{6}$$

Combining Equations (5) and (6), we get

$$f_{i,n} = \max(\max(v(A(t)), f_{i,n-1}) + \frac{L_{i,n}}{r_i})$$
 (7)

Subscript i corresponds to buffer i, and subscript n denotes the nth packet. And S is the packet transmission start time and f is the packet transmission completion time. L is the length of the packet. v(t) is the virtual system time defined by the packet scheduling algorithm. A(t) is the time the packet actually arrived. Therefore, time v(A(t)) means the virtual time used to determine the service order of packets when they enter the weighted scheduling. r is the service rate allocated to the session and is 0.7, 0.2, and 0.1, respectively, with 70%, 20%, and 10% allocated. When a packet's start time and completion time are calculated by the above calculation method, the WFQ scheduler transmits the packet with the smallest completion time. Naturally, if there is no data in the buffer with a weight of 70%, data with a weight of 20% and a weight of 10% is allocated to TDMA and transmitted according to the above equation. Only % of the data will be transmitted via rearrangement of data destined for the same destination. Energy efficiency can be improved by actively allocating TDMA time slots using the WFQ algorithm.

It also utilizes AI machine learning to set the adaptive duty cycle. fig. 7 shows the dataset used for machine learning training. The data set in fig. 7 consists of data related to forest fires, and the data set consists of two parts. It consists of information on areas where wildfires occur most frequently and information on times when wildfires occur most frequently. For the data set, three years (2019-2021) data provided by the Fire Department and each local autonomous district were used to extract elements by fire type (hourly, regional).

Fig. 8 shows the machine learning analysis process. The data set that has undergone the preprocessing process is trained through a support vector machine (SVM) model [19] among machine learning supervised learning [20, 21]. By using the model to be learned, the duty cycle of the sensor node for each time and place is adjusted. For example, it improves reliability by having a small duty cycle in areas

where there are many wildfires and at times when there are many wildfires. Conversely, it is a method to increase energy efficiency by setting a longer duty cycle in areas where there are not many wildfires and during times when there are not many wildfires.





Fig. 8 Schematic of the analysis process

III. RESULTS AND DISCUSSION

The performance of the proposed MAC protocol and the existing QAML-MAC is compared in this paper. The difference from the existing QAML-MAC is the same listensleep structure, so in comparison to delay, it is mainly the packet transmission scheduling part. However, the proposed MAC protocol has the advantage of supporting QoS according to the urgency of data transmission, as well as supporting fairness between traffics by allowing transmission of traffic that has not been transmitted due to priorities for a certain period of time. Also, for performance comparison, it is assumed that 10% of the total data transmitted by each node has the highest priority traffic, that is, abnormal data exceeding a preset threshold. In addition, when such abnormal data occurs, it is assumed that these environments or multimedia traffic providing necessary images have the second priority, and these data are 15% of the total. Finally, it is assumed that the general data transmitted periodically is

75%. In this case, it is assumed that nodes with priority data transmit the data stored in the priority buffer first and then transmit other data. fig. 9 is a comparison of the delay of the MAC protocol proposed in this paper with that of the existing MAC. As shown in the figure, it can be seen that the MAC protocol proposed in this paper has low average delay for data with priority. Since SMAC protocol does not support priority by default, it shows the average delay of all transmitted data.



Fig. 9 Priority data average delay

Fig. 10 shows a comparison of fairness between traffic. As shown in the figure, if there is no initial abnormal data, only periodic data transmission takes place, so it occupies 100% of the total transmitted traffic. However, periodic data transmission with low priority is not performed when abnormal data occurs. After a certain period of time has elapsed after the occurrence of the abnormal state data, the necessary image information is transmitted. Even in this case, data with the lowest priority is not transmitted. However, since periodic data transmission has to be performed after the initial set 10 frame transmissions, the total ratio of periodic data increases by a certain amount. Finally, when abnormal data generation is stopped, only periodic data transmission is performed again as in the beginning.



Fig. 10 Fairness comparison between traffic

IV. CONCLUSION

In this paper, we proposed an adaptive network MAC that compensates for the problems caused by the previously studied QAML-MAC protocol's inability to transmit other low-priority data due to high-priority data. In addition, energy efficiency was strengthened by selecting and training the optimal learning model of machine learning and setting the optimal adaptive duty cycle.

In the existing QAML-MAC protocol, when data requiring urgency, that is, high priority, is continuously generated, data of relatively low priority in order to transmit such data may cause loss of information or connection problems due to transmission difficulties. There may be cases in which a waste of resources such as cutting or a wrong judgment may be made. In this paper, QoS is supported for high-priority data and low-priority data by assigning a certain weight to the lowpriority data and transmitting it. In addition, actively allocating slots during TDMA transmission can prevent wastage of unnecessary slots and increase energy efficiency. In addition, the optimal duty cycle can be set by utilizing the artificial intelligence machine learning model. This can reduce energy consumption, which is the most important issue in wireless sensor networks.

ACKNOWLEDGMENT

This research was funded by a 2021 research Grant from Sangmyung University

References

- M. Shafiq, H. Ashraf, A. Ullah, & S. Tahira, Systematic literature review on energy efficient routing schemes in WSN-a survey. Mobile Networks and Applications, 25(3), 882-895. 2020.
- [2] O.I. Khalaf, G.M. Abdulsahi. Energy efficient routing and reliable data transmission protocol in WSN. Int. J. Advance Soft Compu. Appl, 12(3), 45-53.2020.
- [3] P Maheshwari, AK Sharma, K Verma, Energy efficient cluster based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization. *Ad Hoc Networks*, 110, 102317. 2021.
- [4] C Xu, Z Xiong, G Zhao, S Yu, An energy-efficient region source routing protocol for lifetime maximization in WSN. *IEEE Access*, 7, 135277-135289.2019.
- [5] A Maatouk, M Assaad, On the age of information in a CSMA environment. *IEEE/ACM Transactions on Networking*, 28(2), 818-831.2020.
- [6] W. Ye, J. Heidemann, and D. Estrin, Medium access control with coordinated adaptive sleeping for wireless sensor networks, *IEEE/ACM Transaction on Networking*, Vol. 12, Issue3, pp. 493-506, June 2004.
- [7] Z Xu, J Luo, Z Yin, T He, F Dong, S-MAC: achieving high scalability via adaptive scheduling in LPWAN. In IEEE INFOCOM 2020-IEEE Conference on Computer Communications (pp. 506-515). IEEE.2020.
- [8] T. Dam, K. Langendoen, An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks, ACM Sensys '03, Nov 2003.
- [9] H Pan, SC Liew, Information update: TDMA or FDMA?. IEEE Wireless Communications Letters, 9(6), 856-860.2020.
- [10] M. Xie, X. Wang, "An Energy-Efficient TDMA Protocol for clustered Wireless Sensor Networks, 2008 ISECS International Colloquum on Computing, Communication, Control, and Management, Vol. 2, Issue, 3-4, pp. 547-551, Aug. 2008.
- [11] JS Lee, YS Yoo, HS Choi, T Kim, Energy-efficient TDMA scheduling for UVS tactical MANET. *IEEE Communications Letters*, 23(11), 2126-2129. 2019.
- [12] Tang Zhenzhou, Hu Qian, "An Adaptive Low Latency Cross-Layer MAC Protocol for Wireless Sensor Networks, 2009 8th IEEE International Conference on Dependable, Autonomic and Secure Computing, pp. 389-393, 2009.
- [13] A Li, W Liu, L Zeng, C Fa, Y Tan, An efficient data aggregation scheme based on differentiated threshold configuring joint optimal relay selection in WSNs. *IEEE access*, 9, 19254-19269.2021.
- [14] Q Wang, D Lin, P Yang, Z Zhang, An energy-efficient compressive sensing-based clustering routing protocol for WSNs. *IEEE Sensors Journal*, 19(10), 3950-3960.2019.
- [15] G. Kaur, P. Chanak, & M. Bhattacharya, Energy-efficient intelligent routing scheme for IoT-enabled WSNs. *IEEE Internet of Things Journal*, 8(14), 11440-11449. 2021.

- [16] C Wang, Y Zhang, X Wang, Z Zhang, Hybrid multihop partition-based clustering routing protocol for WSNs. *IEEE Sensors Letters*, 2(1), 1-4.2018.
- [17] D. Puri, & B. Bhushan, Enhancement of security and energy efficiency in WSNs: Machine Learning to the rescue. In 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 120-125). IEEE. 2019
- [18] Z Sun, L Wei, C Xu, T Wang, Y Nie, X Xing, J Lu, An energy-efficient cross-layer-sensing clustering method based on intelligent fog computing in WSNs. *IEEE Access*, 7, 144165-144177. 2019.
- [19] S. Huang, N. Cai, P.P. Pacheco, S. Narrandes, Y. Wang, & W. Xu, Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics & proteomics*, 15(1), 41-51. 2018.
- [20] S. Ray, A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE. 2019.
- [21] S. Sun, Z. Cao, H. Zhu, & J. Zhao, A survey of optimization methods from a machine learning perspective. *IEEE transactions on cybernetics*, 50(8), 3668-3681.2019.