JOIV : Int. J. Inform. Visualization, 6(2-2): A New Frontier in Informatics - August 2022 463-468

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Optimal Data Transmission and Improve Efficiency through Machine Learning in Wireless Sensor Networks

Hyunjoo Park [a], Junheon Jeon [b,*]

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

***Abstract*—Each sensor node in WSN is typically equipped with a limited capacity small battery. Energy-efficient communication is therefore considered a key component of network life extension. In addition, as the utilization of the sensor network increases, duplicate data and abnormal data is also collected to reduce the accuracy of the data in various environments. AI is used to recognize data anomaly values and increase packets' accuracy by removing out-of-range data. This can improve performance through optimal data transmission, resulting in increased network life, energy efficiency, and reliability. This paper proposes a protocol called MLQ-MAC that reflects the above. MLQ-MAC uses AI techniques to consider different types of data packets. The data collected by the sensor removes the measurement anomaly and duplicate data and stores it in a different transmission queue by priority. Efficient data transfer is possible by using an AI Discriminator for accurate classification before being stored on a transmission queue. The AI-Discriminator classifies a variety of factors, including the collection environment, characteristics of network applications, and so on. It also uses two new technologies: self-adaptation and scheduling for efficient transmission. In the protocol, the receiver adjusts the duty cycle according to to transmit urgency to improve network QoS. Finally, the simulation results show that the MLQ-MAC protocol reduces energy consumption at the receiver by up to 3.4% and per bit by up to 2.3% and improves packet delivery accuracy by up to 3%.**

***Keywords*— Artificial intelligence; energy efficient; machine learning; WSNs; QoS support; MAC protocol.**

Manuscript received 19 Oct. 2021; revised 29 Jan. 2022; accepted 8 Mar. 2022. Date of publication 31 Aug. 2022.

I. INTRODUCTION

In this paper, we propose a mechanism to use artificial intelligence (AI) to transmit using priority queues [1], [2] to improve transmission efficiency after the accurate determination of data. This mechanism is divided into two parts. First, it combines artificial intelligence (AI) technology to determine the exact content of data by determining and transmitting data and applying a transmission algorithm to the characteristics of the data. The second is to improve energy efficiency by adjusting the duty cycle for energy-efficient transmission [3]. This is one of the best ways to extend the life of the entire network in a battery-operated sensor network [4], [5]. The collection of disaster signals during the extended range of the sensor network has several important limitations [6]. It should be checked to process an urgent disaster signal quickly and determine whether the collected data is a meaningful value for system operation and whether the measurement value is exact [7]-[9].

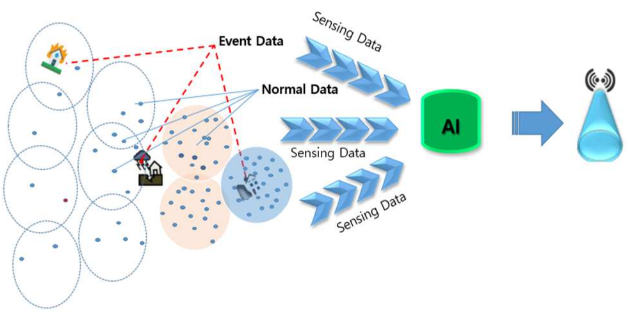

Fig. 1 Sensing Data in Disaster System

The system is connected to a network, which communicates with each other as their own language or protocols and determines how data is transmitted and received. Depending on the exact analysis and processing of the data, temperature and fire suppression can be detected, and disaster diffusion can be prevented in the bud. If there is an error in the transmission data value and a problem with the new signaling in the BS, there is a problem that does not guarantee the people's safety and drops the trust of the entire network. The range of the wireless sensor network is expanding and is also important in the disaster system, environmental monitoring, and wildlife tracking. Disaster situations are becoming more common, diversified, and large, and the need for a system to prepare for disasters is increasing as the damage is also increased if a disaster. The advancement of the disaster response system is deployed to prepare for the disaster situation that occurs regardless of the time and place and to establish systems that detect disaster situations or inform safety evacuation routes [4], [5]. However, as the disaster situation is diversified, the number of sensors that need to be managed as well as the number of sensors that need to be managed increases exponentially. In addition to the effective processing of data generated by the sensor, errors in data due to malfunction of the sensor also degrade the trust of the network as a whole, and the loss of resources [4] due to data processing impacts the life of the entire network. In addition, it is necessary to establish a system that can efficiently manage and utilize sensors because it can cause confusion among people if the system fails to detect a number of sensors properly or if the system is malfunctioning. In a disaster situation, we propose a way to reduce processing and transmission errors through accurate data determination using artificial intelligence. The proposed system is an artificial intelligence system that allows data to be fractionated by artificial intelligence to determine whether a message is a disaster signal, a general signal, or a general signal, which determines the priority and increases the efficiency of transmission.

II. Materials and Method

A. Data transmission and Duty Cycle in Sensor Node

The sensor node comes up after a period of time to save energy [8]. After the wakeup, collect data from other nodes. If data is not collected for a period of time, it switches to sleep mode to save energy.

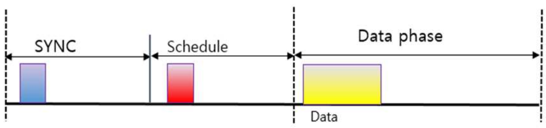

Fig. 2 Sensor Node Duty Cycle

To save energy, the S-MAC [10], [11] protocol iteratively repeats the active mode and sleep mode to address the unnecessary idle listening problem of the sensor network. In the periodic frame, the active mode and the sleep mode, which are inactive intervals, are operated at a fixed length. The sleep mode does not transmit data by powering down the sensor but saves energy by storing it in a buffer, thereby prolonging the overall network life. The active mode verifies whether or not the data is stored in the buffer and sends the data to the neighboring sensor node. However, there is a drawback that it cannot be applied to changes in traffic.

The T-MAC [12] protocol ensures reliability by the sensor nodes exchanging the RTS packet and the CTS packet to avoid collisions using data transfer and ACK (: Acknowledge). The T-MAC protocol can increase efficiency by reducing the energy consumed in idle listening if no data transmission and reception is detected in active mode. However, if the data occurrence interval is not constant, the transmission efficiency will be low due to the occurrence of a transfer after one frame. The SMAC has been sent to the active section, but TMAC does not have a constant interval when data occurs, so it is changed to the sleep interval after the Ta time and is not sent in the previous frame and is transmitted in the next active section.

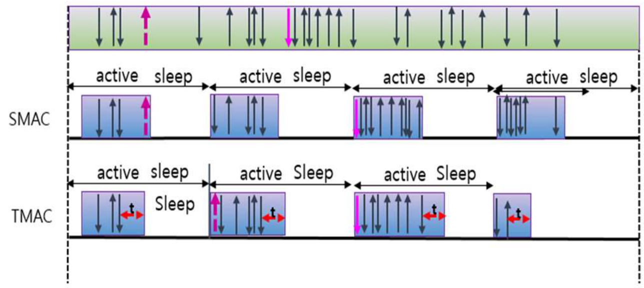

Fig. 3 Data transmission in SMAC and TMAC

The QAML-MAC [13] method can work well in an environment where low-priority data is not transmitted, even if it is not transmitted. However, event data that exceeds the threshold is transmitted continuously until the event is lost, resulting in a phenomenon that occupies network traffic.

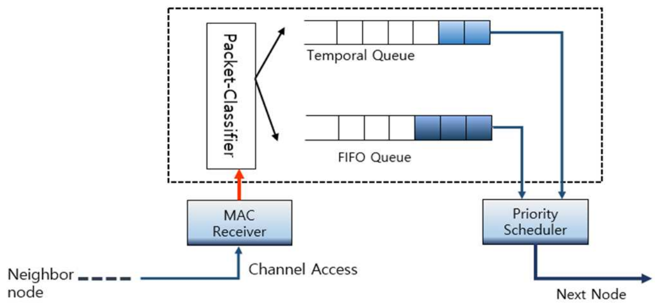

Fig. 4 Packet Classifier in QAML-MAC

B. Machine Learning Algorithms for WSNs

Machine learning algorithms for WSNs is shown in Fig. 7. It shows machine learning is divided into the following categories [14]-[16].

- Supervised Learning
- Unsupervised Learning
- Reinforced Learning.

A decision tree suitable for the algorithm to be applied in this paper during machine learning is one of the teaching methods. Supervision learning learns a function that maps input to output based on input-output training data pairs. In supervised learning, each example is a pair consisting of an input object (typically a vector) and the desired output value (also called the supervisory signal). The supervised learning algorithm analyzes the training data and generates an inferred function, which can be used to map the new example. The optimized scenario allows you to categorize classes of unknown instances correctly. This requires that the learning algorithm generalizes from training data to a situation that does not appear to be "rational". The decision tree algorithm is a subcategory of the classification of supervised learning branches in machine learning.

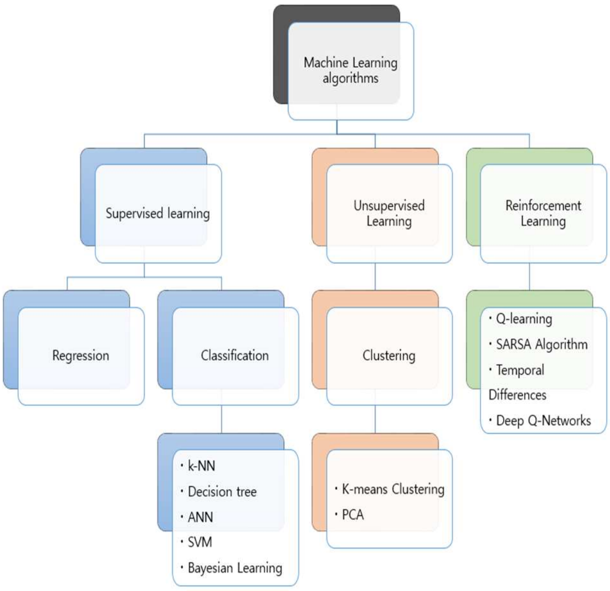

Fig. 5 Machine Learning Algorithm

C. Decision Tree

This paper uses a decision tree in a machine learning algorithm. Decision trees [16] are algorithms that correspond to the teaching of machine learning. Machine Learning Map Learning is a representative algorithm that gives you both a feature (input) and a label (output). It is an algorithm that classifies data. To output a class of input samples, output a class probability distribution with leaf nodes or determine a class to output. By calculating a simple question on a node, each time a question is processed, go down one step down and usually ends up in dozens of steps, so it is a very small amount of arithmetic. The decision tree has several characteristics [8], [15].

1) First, handle both metering data and non-metering data. When dealing with data that is not numeric, do not need to consider the numerical size of the data. For example, data that cannot be unified, such as (acids, weather, and temperature of 60 degrees), can be used as a training set pair.

2) Second, interpret the classification results. When looking at the questions in the absence, explain to the user why submitting these classification results is based on the root node's leaf node. This property is very useful and will automatically explain the classification results to the user in many applications to increase satisfaction.

3) Third, the decision tree is an unstable classifier. Even if the price of a training set or a stop condition changes a little bit, the tree's shape varies greatly. However, this is useful in other applications, such as ensemble methods. This is because combining an unstable classifier can lead to good performance.

The following is a DECISION TREE algorithm:

- Input: Decision Tree R, Test Sample x
- Output: Class k of Sample x

```
T = R
While (T ≠ Nil)
   x is the question of T and the result is r.
   If (r= Yes) Left child node of T = T
   Else T = T's right-hand node
   If (T is the leaf node)
      K = class of k = T
         T = Nil
```

A decision tree is made by comparing the relationship between the received data and the threshold.

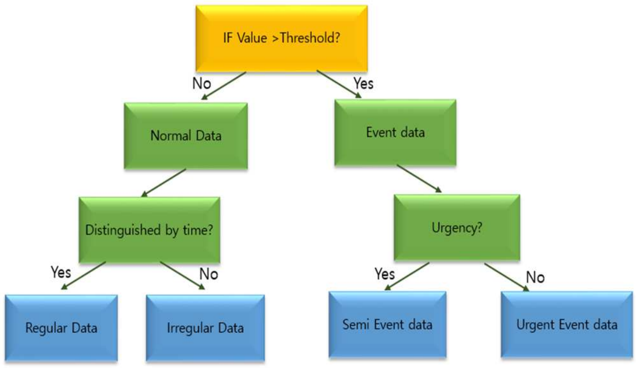

Fig. 6 AI Decision Tree

D. Data Classification and training

Using an artificial intelligence algorithm can transmit efficient data through data classification and training [15], [16]. When data is collected from another node after the wakeup, it classifies normal data and event data by a given algorithm, which means information that exists in the threshold range or is usually sent at a fixed time. Event data refers to values outside the threshold range, meaning urgent data.

However, data of errors can occur in the wireless sensor network environment using the sensor node. For example, it is not urgent data but can result in event data due to errors in the sensor node. In this case, time and energy consumption can occur to transmit event data.

The pre-learned optimal machine learning model predicts the classified and normal event data. The predicted data is divided into actual normal data and actual event data. The machine learning model has learned the collection data of sensor nodes in the existing wireless sensor network and sets the optimal model and optimal parameters by evaluating the learned data.

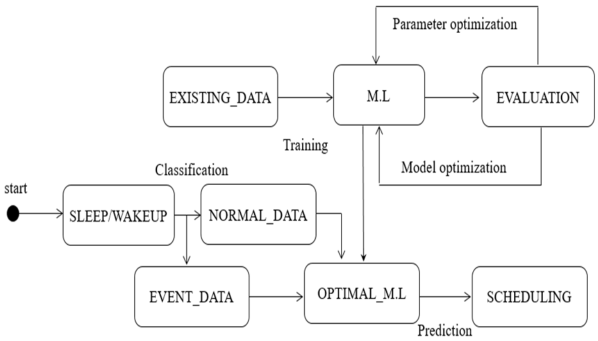

Fig. 7 AI data training Process

E. Proposed MAC Protocol

1) Eliminate redundant data with cosine similarity functions: Normal data and event data can be channel allocated by AI. The data received from the sensor is divided into Normal data and Event data based on thresholds. Normal data is divided into non-regular data that is transmitted at a certain amount of data, depending on the nature of the signal, and regular data that the data transmission time is fixed according to the time limit. Redundant data reduces the efficiency of the network at the time of transmission, thereby reducing network efficiency overall [17]. It is removed immediately if duplicate data is found using the cosine duplication function, and it also improves transmission efficiency by eliminating redundant data in the incoming data. The following formula shows the cosine similarity function. X, Y is an input/output pair; if the resulting value is 1, it is considered redundant and does not transmit data.

$$cs(X,Y) = \frac{\sum_{i=1}^{n}\sum_{k=1}^{m}(x_i \cdot y_{i+k})}{\sqrt{\sum_{i=1}^{n}(x^2)}\sqrt{\sum_{k=1}^{m}(y_k)^2}} \quad (1)$$

If the total amount of transmission is reduced, the network load will decrease, speeding up the transmission speed, increasing efficiency, and increasing the overall network life span.

2) Machine Learning Training Conditions: In this paper, 75 % of the existing data was used for the evaluation of the learning machine learning algorithm, and 25 % used the tester data. This is based on the ratio of training data for the classification of common sensor data and can vary the ratio depending on the sensor's environment.

3) Training Result and selection of parameters: The evaluation data selected the optimal parameter with an accuracy of about 73 % in the first study and a model with about 91 % accuracy after the optimal parameter selection. The data collected by the optimal machine learning model is classified as regular normal data, non-regular normal data, semi-event data, and very urgent event data.

4) Channel Allocation scheme: If event data and normal data are collected, the channel is allocated by machine learning. The classified data is sent on the allocated channel[18-20], and if there is a large amount of transmission data, the channel that has not been transmitted in time will be assigned data to the channel of TDMA according to the environment, such as time, location, and time, in the event data and normal data generation by the AI machine learning model. This is highly efficient when there is many transmission data.

5) Data Transmission Scheme: The transmission of the classified data is used by the CSMA scheme [21] of the existing wireless sensor network. This method is a channel allocation technique that ensures the QoS of the transmission data by the priority of the channel's sleep section and the wakeup interval.

- Normal data /non-Regular data

For regular normal data and non-regular normal data, it occurs at a given wake-up time and attempts to transmit. The receiving node receives the above data by waking up at a given time with a fixed sleep/wake-up time.

- Semi Event data

In the case of the semi-event data, the data transmission is attempted by a half-wave up than the defined wakeup time. It can be sent more quickly to the receiving node because it takes up half the wakeup time of the transmitting node. However, in this case, if the receiving node has been slopped just before the transmitting node's wake up, the transmitting node must have a sike of the receiving node. The receiving node receives the semi-event data by waking up at a given time with a fixed sleep/wake up time. The difference from the normal data transfer above is that the existing CSMA scheme does not schedule the nodes' sleep/wake up time. So, for normal data, if the transmitting node wakes up later than the receiving node, the transmitting node must wait and transmit a sike of duty cycle. An attempt is made to transmit data by adjusting the interval of the wakeup interval for quick receipt and transmission of the given event data.

In general, for data that requires emergency, the size is very short. Because the size of the data is short, the transmission time and the reception time on the receiving node can also be shortened, which can lead to quick action. For example, the SOS signal, which is a distress signal at sea, is very short. The Moss sign is also short and concise so that it can be easily transmitted. The purpose of the event data transfer is simply to use only the alert data or the emergency situation, so it attempts to transfer a different way from the large amount of data that is required for transmission efficiency (Throughput). In addition, it is important to reduce the delay in transmission because the most important criteria for transmission performance are emergency.

- Emergency Event data

For every emergency event data, the transmitting node immediately wakes up and attempts to transmit. So can use it to send and receive data faster. Emergency event data is important for both data discovery and transmission of data so that the wakeup cycle does not cause a delay in transmitting and receiving. Data classified using artificial intelligence reflects its characteristics and affects nodes' sending and receiving algorithms [22]. The data that requires the most urgency is also very important to the reliability of data recognition and transmission reception, so it is important to send safe and fast data to the sink node.

The channel allocation algorithm, which considers normal and event data's transmission efficiency, is shown in Fig. 8.

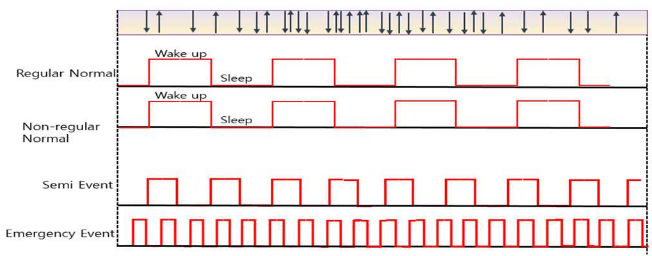

Fig. 8 Channel Allocation in MLQ-MAC

III. Results and Discussion

A. Packet Classifier Algorithm

Data Packet classifier in MLQ-MAC is performed in the following order.

1. The data collected from the sensor node is classified as normal data and event data in the classifier.
2. Classification of data collected from a classifier uses a cosine-like function to remove redundant data.
3. The normal data and event data are classified as non-normal, June event, and emergency event data that need to be transmitted according to the normal data/data amount of time required by the AI algorithm to be assigned to each channel.
4. The operation of the channel is operated by adjusting the active interval and the sleep interval according to the nature of the data. The normal data repeats the wakeup and sleep at a fixed interval, and the event data increases efficiency by running the event and the emergency event's wake up/sleep interval.

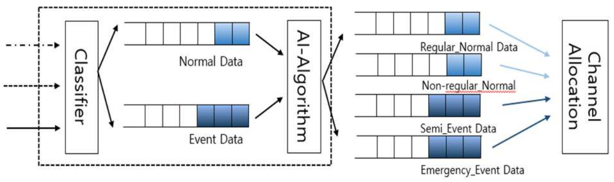

Fig. 9 MLQ-MAC

This is the data for the data that measured the average delay. In the transmission of data with priority, the MLQ-MAC exhibits some good performance.

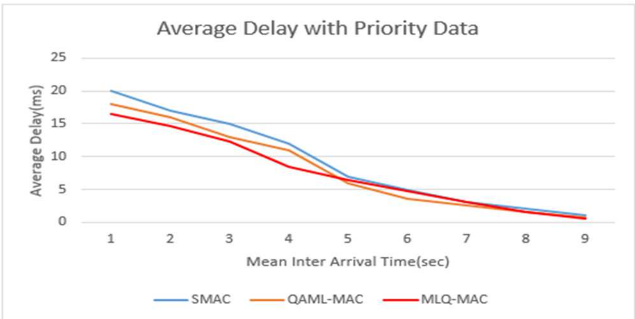

Fig. 10 Average Delay

The energy consumption per node is shown in fig. It can be seen that the sender node in MLQ-MAC consumes slightly higher energy of up to 3.7% compared to SMAC. The number of nodes for measurement was limited to 25.

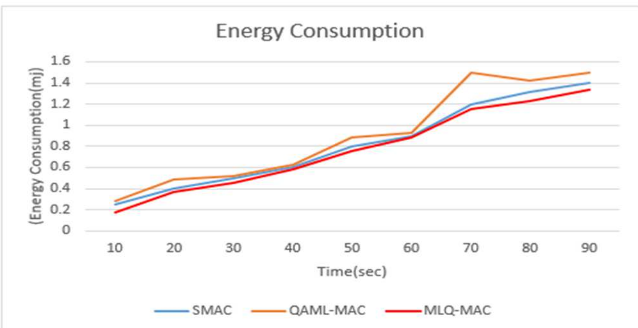

Fig. 11 Energy Consumption

IV. Conclusion

In this paper, we proposed an algorithm to enhance data reliability by increasing data transmission efficiency and using artificial intelligence to further refine the characteristics of data stored in priority queues. Data collected for this purpose efficiently manages channel resources by removing duplicate data from the classifier. In addition, data using artificial intelligence can be used to improve the efficiency of the entire network by operating channels to meet the characteristics of the data. The area where AI can be applied to sensor networks is getting bigger and bigger. As sensor network applications require the provision of various services, QoS support [22], [23] has become an essential element. Data's nature and exact classification is an important factor in improving the efficiency of the entire network and the service. In this paper, we proposed a MAC protocol to increase the accurate judgment of data and the efficiency of transmission using artificial intelligence. The incoming data is removed before it is stored in the transmission queue and divided by time. The transfer efficiency was good because the priority was divided and stored in the queue. A transmission queue with different priorities applies different wake-up cycles to the nature of each data, thus reducing transmission delays.

Future research will focus on the delay and energy efficiency of data in the sensor network, as well as other applications in the wireless network (e.g., body network, etc.) to provide customized transmission according to data characteristics.

Acknowledgment

This research was funded by a 2022 research Grant from Sangmyung University.

References

[1] Petrova, M.; Mähönen, P.; Osuna, A., "Multi-class classification of analog and digital signals in cognitive radios using support vector machines," In Proceedings of the 2010 7th International Symposium on Wireless Communication Systems (ISWCS), York, UK, 19–22 September 2010; pp. 986–990.

[2] Hyun Joo Park; Seong Cheol Kim; "An Efficient Packet Transmission Protocol Using Reinforcing Learning in Wireless Sensor Networks," Journal of System and Management Sciences Vol. 11 (2021) No. 2, pp. 65-76

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] Mateus P. Mota, Alvaro Valcarce, Jean-Marie Gorce, Jakob Hoydis, "The Emergence of Wireless MAC Protocols with Multi-Agent Reinforcement Learning," Electronics 2021, 10, 318. [Online]. Available: https://doi.org/10.3390/electronics10030318

[8] M. Mardiyono, R. Suryanita, and A. Adnan, "Intelligent Monitoring System on Prediction of Building Damage Index using Neural-Network," Telkomnika Telecommunication, Computing, Electronics and Control, vol. 10, no. 1, pp. 155-164, Mar. 2012. http://www.iaesjournal.com/online/index.php/IJEECS/ article/view/665

[9] K. Lorincz, D. J. Malan, T. R.F. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, and M. Welsh, "Sensor networks for emergency response: challenges and opportunities," IEEE Pervasive Computing, vol. 3, no. 4, pp. 16-23, Oct.-Dec 2004. https://doi.org/10.1109/MPRV.2004.18

[10] 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.

[11] 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.

[12] T. Dam, K. Langendoen, "An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks," ACM Sensys'03, Nov 2003.

[13] 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.

[14] Szymon Szott, Katarzyna Kosek-Szott, Piotr Gawłowicz, Jorge Torres Gómez, Boris Bellalta, Anatolij Zubow, Falko Dressler, "WiFi Meets ML: A Survey on Improving IEEE 802.11 Performance with Machine Learning," arXiv:2109.04786 [cs.NI], [Online]. Available: https://doi.org/10.48550/arXiv.2109.04786

[15] Kulin, M.; Kazaz, T.; De Poorter, E.; Moerman, I. "A Survey on Machine Learning-Based Performance Improvement of Wireless Networks: PHY, MAC and Network Layer," Electronics 2021, 10, 318. https://doi.org/10.3390/ electronics10030318

[16] Himanshu Sharma, Ahteshamul Haque, and Frede Blaabjerg, "Machine Learning in Wireless Sensor Networks for Smart Cities: A Survey," Electronics 2021, 10, 1012. https://doi.org/10.3390/ electronics10091012

[17] L. Tian, E. Lopez-Aguilera, E. Garcia-Villegas, M. T. Mehari, E. De Poorter, S. Latré, and J. Famaey, "Optimization-Oriented RAW Modeling of IEEE 802.11ah Heterogeneous Networks," IEEE Internet of Things Journal, vol. 6, no. 6, pp. 10 597–10 609, 2019.

[18] H Pan, SC Liew, "Information update: TDMA or FDMA?" IEEE Wireless Communications Letters, 9(6), 856-860, 2020.

[19] 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.

[20] JS Lee, YS Yoo, HS Choi, T Kim, "Energy-efficient TDMA scheduling for UVS tactical MANET," IEEE Communications Letters, 23(11), 2126-2129. 2019.

[21] A Maatouk, M Assaad, "On the age of information in a CSMA environment," IEEE/ACM Transactions on Networking, 28(2), 818-831.2020.

[22] Barbancho, J.; León, C.; Molina, J.; Barbancho, "A. Giving neurons to sensors. QoS management in wireless sensors networks," In Proceedings of the Emerging Technologies and Factory Automation, Diplomat Hotel Prague, Czech Republic, 20–22 September 2006; pp. 594–597

[23] M. Nivaashini and P. Thangaraj, "Computational intelligence techniques for automatic detection of Wi-Fi attacks in wireless IoT networks," Wireless Networks, vol. 27, no. 4, pp. 2761–2784, May 2021.