Unslotted CSMA/CA mechanism with reinforcement learning of Wi-SUN MAC layer

Sihyeon Lee, Sang-Hwa Chung
Department of Information Convergence Engineering
National University Busan, South Korea
shyeon0528@gmail.com, shchung@pusan.ac.kr

***Abstract*— Wi-SUN, a protocol necessary for realizing a smart city, uses IEEE 802.15.4 Unslotted CSMA/CA for MAC Layer to build a large-scale network. However, as the scale increases, the contention becomes more severe, and the performance of Unslotted CSMA/CA deteriorates rapidly. Also, there is a problem of wasting resources because there is no scheduling. Therefore, in this paper, the performance of the MAC layer is improved by 20% by applying reinforcement learning to each node and learning to select the optimal backoff delay value.

Keywords—IEEE 802.15.4; Wi-SUN; unslotted CSMA/CA; Q-learning

I. Introduction

Wi-SUN (wireless smart utility network) is an international wireless communication standard that does not require communication costs compared to other LPWA (Low Power Wide Area), and can automatically monitor radio wave conditions through multi-hop communication to switch routes. It has high reliability. Therefore, it is attracting attention as a communication technology that can build a large-scale mesh network necessary for the realization of a smart city [1].

II. Background

Wi-SUN adopts IEEE 802.15.4 MAC unslotted CSMA/CA to build a large-scale mesh network

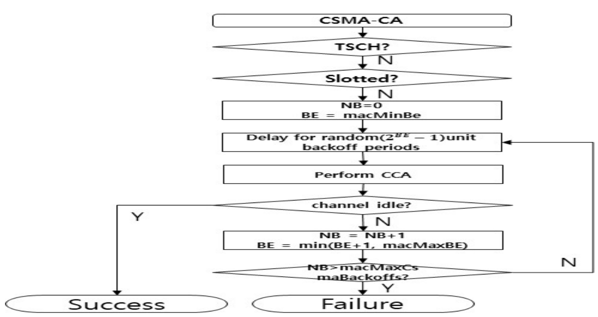

Figure 1. IEEE 802.15.4 Unslotted CSMA/CA Algorithm

Figure 1 shows the IEEE 802.15.4 unslotted CSMA/CA algorithm. A node that wants to transmit a packet can transmit after gaining access to the channel through the Unslotted CSMA/CA algorithm before accessing the channel. Each transmission node randomly selects a backoff value in the range [0, 2^BE-1] from the transmission time and delays transmission by the selected value. Thereafter, it is determined whether the channel to be accessed is in an idle state by performing CCA (Channel Clear Access). If the channel is idle, the node gains access to the channel and transmits the packet. If it is not idle, CCA is retryed by increasing the BE value by 1 and randomly setting the Backoff value in a wider range, and it can be repeated until the maximum retry chance is reached[2].

Q-learning is an off-policy algorithm based on the TD (Temporal-Difference) model. Unlike the value function V(s) that stores the reward only for the agent's state S, Q(s,a) stores the reward by reflecting the agent's state and action. The Q(s,a) value is updated by the following equations (1) and (2), and it has a policy to select the action to maximize the reward through the Q-Table of (s,a)[3].

$$\Delta Q(s,a) = \{r(s,a) + \beta \times \max_a Q(s',a')\} - Q(s,a) \tag{1}$$

$$Q(s,a) = (1-\alpha) \times Q(s,a) + \alpha \times \Delta Q(s,a) \tag{2}$$

Epsilon-Greedy is a simple method to balance exploration and exploitation by choosing between exploration and exploitation randomly. Since a better policy cannot be found only by taking an action that maximizes the Q value, the Epsilon-Greedy algorithm is used to find a policy with a better reward by making a new attempt with a certain probability. A random action is selected with a probability of ε, and a maximized action is selected with a probability of 1-ε.

III. Problem Analysys

Although IoT devices are increasing explosively with the Internet of Things, the increasing number of nodes deepens contention and adversely affects network performance[4]. CSMA/CA, which is a radio channel access method, is a method of avoiding transmitted packet collisions between different nodes. However, the packet collision problem still occurs. Although Wi-SUN is a communication standard targeting smart cities and smart grids with dozens or hundreds of nodes connected in a wide area, performance degradation due to collisions in dense nodes is inevitable. TSCH, another MAC protocol of IEEE 802.15.4, guarantees low-latency, high-reliability, contention-free packet transmission and reception by scheduling packet transmission. On the other hand, wireless nodes of Wi-SUN are mostly sensor nodes, and although they generate traffic at regular intervals, it is difficult to introduce a scheduling method according to time.

Two problems arise from the problem of not being able to schedule. First, it is a waste of channel resources. Specifically, since each node does not know the backoff delay of neighboring nodes overlapping with its own transmission range, it is not possible to use the network channel densely. That is, there are unused times when looking at the usage efficiency of the entire network channel.

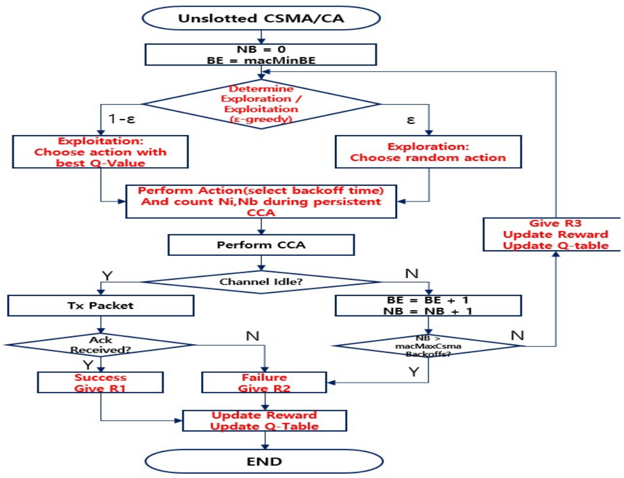

Figure 2. Q-learning Based Unslotted CSMA/CA Algorithm

Second, it is possible to minimize network channel waste and packet latency after learning. Each node selects an appropriate backoff value after learning to use the channel more efficiently, and the latency of packets is optimized and reduced.

IV. Proposed Algorithm

$$\mathbf{macIdleSum} = \frac{\mathbf{macIdleSum}}{\mathbf{2}} + \frac{\boldsymbol{N_i}}{\mathbf{2}} \tag{4}$$

$$\mathbf{macBusySum} = \frac{\mathbf{macBusySum}}{\mathbf{2}} + \frac{\mathbf{N_b}}{\mathbf{2}} \tag{5}$$

$$\mathbf{State} = \frac{\mathbf{macIdleSum}}{\mathbf{macIdleSum + macBusySum}} \tag{6}$$

$$\mathbf{state} = \mathbf{round}\left(\frac{\mathbf{macIdleSum}}{\mathbf{macIdleSum + macBusySum}}\right) * \mathbf{10} \tag{7}$$

Q-learning determines the action by referring to the Q-Table of state and action, so the number of states and actions affects performance. If you set it as in (4) to see the channel usage, the number of States becomes very large, so the result value of (6) is rounded up and used. Therefore, State has a total of 11 States from 0 to 10, and as a result, the size of Q-Table is Action(64) * State(11).

The rewards that the Agent receives are divided into 4 categories. First, when Agent selects Action (backoff) and the channel is idle at the time of backoff, packet is transmitted and ACK is received. Second, when Agent selects Action (backoff) and the channel is idle at the time of backoff, and packet is transmitted, but ACK is not received. Third, when the Agent selects Action (backoff), and the channel is busy at the time of backoff, it is necessary to select an Action by moving to the next State. Fourth, even though the agent continuously performed the third case, the channel access failed by exhausting all backoff opportunities (NB). The 1st case is defined as R1, 2nd and 4th cases are defined as R2, and the 3rd case is defined as R3.

$$\mathrm{P_b} = -\frac{Ni}{N_b + N_i} * N_{backoff} * \frac{1}{D} \tag{8}$$

$$\mathrm{R}_1 = 1 - P_b \tag{9}$$

$$R_2 = -1 - P_b \tag{10}$$

$$R_3 = 0 - P_b \tag{11}$$

The (1,-1,0) values before Pb are compensation for transmission success, transmission failure, and channel busy conditions, respectively. Pb is a formula to give a penalty when the agent selects an unnecessarily high backoff delay, and the variable D is a hyperparameter for how much penalty for latency will be given. That is, the smaller the D value, the larger the overall value of the equation, and the larger the D value, the smaller the value of the equation. to inflict a high penalty. For example, if the value of D is large, the penalty for latency is small, and the node gets a higher reward for successfully sending a packet than the latency, which increases the probability of success even though latency increases. By combining the two formulas, the agent can receive a higher reward as the transmission succeeds and the less wasteful backoff delay is selected as the action.

V. Simulation

Table1. Simulation Parameters

Parameter	Value
Simulator	Network Simulator 3 (NS3)
MAC protocol	Unslotted CSMA/CA
Number of Nodes	2/5/10/15/20/25/30/35/40/45/50/60
macMaxBE(2^BE)	5(32), 6(64), 7(128)
Topology	Star Topology
Traffic Periodicity	1s
Simulator running time	1000s
Hyperparameter D	64
TX retransmission	0

The sink node is fixed to one, and the child nodes of 1hop gradually increase as shown in Table 1. The macMaxBE of the proposed method is 6, and the existing method was tested by comparing BE=5,6,7. Traffic creates a packet every second from each child node to the sink node and transmits it 1000 times in 1000 seconds. Therefore, since the start time of every packet is the same for all nodes, a contention situation occurs. Hyperparameter D was set to 64 by matching the BE value.

In general, when the number of competing nodes is the same, the higher the PDR (Packet Delivery Ratio), the higher the latency. Since these two are in a trade-off relationship, it is difficult to properly compare their performance. Therefore, the performance improvement is shown by comparing the latency after matching the PDR of the proposed method similarly to the existing method. The PDR of the proposed method was matched by adjusting hyperparameter D.

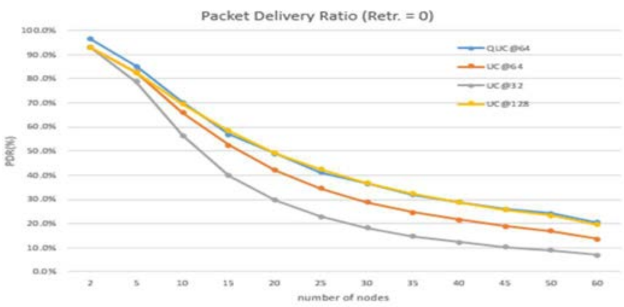

Figure 3. Packet delivery ratio graph

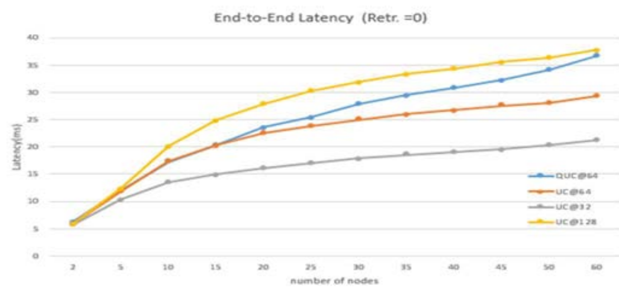

Figure 4. End to End Latency graph

Figure 3 and Figure 4 are PDR and latency graphs according to the number of competing nodes, respectively. QUC (Q-learning based Unslotted CSMA/CA) means the proposed method, and UC stands for Unslotted CSMA/CA, which is an existing method. The number after @ means 2^BE, and the larger the number, the wider the Backoff delay can be selected. As BE increases, PDR increases, but latency also increases, indicating that the two are in a trade-off relationship.

QUC (Q-learning based Unslotted CSMA/CA) means the proposed method, and UC stands for Unslotted CSMA/CA, which is an existing method. The number after @ means 2^BE, and the larger the number, the wider the backoff delay can be selected. As BE increases, PDR increases, but latency also increases, indicating that the two are in a trade-off relationship.

Important to look at are QUC@64 and UC@128. The PDRs of QUC@64 and UC@128 were made similar by adjusting the hyperparameter D of QUC. After that, looking at the latency in Fig 4., QUC@64 is lower. UC@128 randomly selects the backoff delay in the range of [0,127], whereas QUC@64 selects the best backoff delay in the narrower range [0,64], which improves performance by up to 20%.

Comparing the proposed method (QUC@64) with the same resource (backoff range) with the existing method (UC@64), the PDR of the proposed method showed an average performance improvement of 23.3%.

The proposed method (QUC@64) has a higher TX-ACK rate than the existing methods (UC@64, UC@128) in all situations. This is because it gets a higher number of ACKs with fewer TXs, so the proposed method is better in terms of energy efficiency consumed in TX. The proposed method

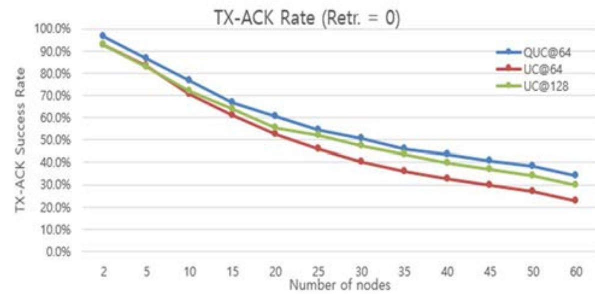

Figure 5. TX-ACK Rate

improved performance by 23% on average compared to UC@64 and 7.7% on average compared to UC@128.

VI. Conclusion

In this paper, we point out the problems of Unslotted CSMA/CA, a Wi-SUN MAC protocol, and propose a mechanism to improve MAC layer throughput by using reinforcement learning to each node for when channel access competition intensifies.

As a result of the experiment, the performance of PDR was improved by 20% within the range of less Backoff Time Selection, and it was better than existing method in terms of transmission energy consumption by better receiving ACKs during transmission using the proposed method.

The proposed mechanism can be applied not only to Wi-SUN but also to other protocols using Unslotted CSMA/CA (e.g. Zigbee).

In future experiments, we plan to apply Deep-Reinforcement learning to unslotted CSMA/CA and to reduce calculating resource for porting to hardware.

Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program(IITP-2022-2016-0-00318) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2020-0-01450, Artificial Intelligence Convergence Research Center [Pusan National University])

References

[1] R. Wayong, R. Okumura, K. Mizutani and H. Harada (2020, May). "A scheduling scheme for channel hopping in Wi-SUN fan systems toward data throughput enhancement." In 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring) (pp. 1-5). IEEE.

[2] IEEE Std 802.15.4™-2015, IEEE Standard for Low-Rate Wireless Networks

[3] Watkins, Christopher JCH, and Peter Dayan. "Q-learning." Machine learning 8.3 (1992): 279-292.

[4] Anastasi, Giuseppe, Marco Conti, and Mario Di Francesco. "The MAC unreliability problem in IEEE 802.15. 4 wireless sensor networks." Proceedings of the 12th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems. 2009.