IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 25, 2021, accepted June 15, 2021, date of publication June 29, 2021, date of current version July 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3093421

Autonomous Decentralized Traffic Control Using Q-Learning in LPWAN

AOTO KABURAKI[1], (Graduate Student Member, IEEE), KOICHI ADACHI[1], (Senior Member, IEEE), OSAMU TAKYU[2], (Member, IEEE), MAI OHTA[3], (Member, IEEE), AND TAKEO FUJII[1], (Member, IEEE)

[1]Advanced Wireless and Communication Research Center, The University of Electro-Communications, Tokyo 182-8585, Japan
[2]Department of Electrical and Computer Engineering, Shinshu University, Nagano 380-8553, Japan
[3]Department of Electronics Engineering and Computer Science, Fukuoka University, Fukuoka 814-0180, Japan

Corresponding author: Aoto Kaburaki (kaburaki@awcc.uec.ac.jp)

This work was supported by the MIC/SCOPE under Grant JP205004001.

This work did not involve human subjects or animals in its research.

ABSTRACT Owing to the recent research and development on the Internet-of-Things (IoT) and machine-to-machine (M2M) communication, wireless sensor networks have attracted considerable attention. Among these networks, low power wide area networks (LPWANs), which realize low power, low data rate, and wide communication area, are most commonly used for long-range communication. These networks adopt asynchronous random-access protocols, such as the pure ALOHA protocol in the medium access control (MAC) layer. Thus, there is a high possibility that multiple nodes transmit packets simultaneously on the same frequency channel, resulting in packet collisions. Carrier-sense multiple access/collision avoidance (CSMA/CA) and centralized resource allocation are effective for avoiding packet collisions. However, these schemes increase the energy consumption of battery-powered LPWAN nodes. In addition, LPWAN has a large coverage area; hence, there is a high possibility that the carrier sense will not work successfully. Thus, this paper proposes a simple but effective machine-learning-based scheme that tackles the packet collision problem by offsetting the transmission timings and avoiding unnecessary packet transmission in an autonomous decentralized manner. Each LPWAN node adjusts the transmission probability and timing using the Q-learning technique. The proposed scheme provides effective packet collision avoidance for LPWAN nodes without the need for an additional control signal. The computer simulation results show that the proposed scheme can improve the average packet delivery ratio (PDR) by 60% compared to the pure ALOHA protocol.

INDEX TERMS Internet of Things (IoT), LoRaWAN, low power wide area networks (LPWAN), machine learning, resource allocation.

I. INTRODUCTION

Owing to the advancements in the field of the Internet of things (IoT) and machine-to-machine (M2M) communications [1], wireless sensor networks (WSNs) that gather sensing information have become important infrastructures for IoT applications. WSNs aim to collect event information such as facts regarding fire and floods in a particular area [2] as well as environmental information through periodic sensing. Low power wide area networks (LPWANs) have become the focus of considerable research attention because they can achieve low power, low data rate, and a wide communication area, which are crucial characteristics in realizing the WSNs [3]. Thus, LPWAN is expected to be applied in smart cities, environmental monitoring, and smart meters. It adopts interference and noise-tolerant modulation schemes, such as binary phase-shift keying (BPSK) and chirp spread spectrum (CSS) modulations [4] as a physical layer modulation scheme to enable long-distance transmission. Because LPWAN nodes need to be inexpensive, they generally possess low quality and low-cost circuit configurations [3]. Therefore, LPWAN adopts asynchronous random-access protocols, such as the pure ALOHA protocol in the medium access control (MAC) layer.

The associate editor coordinating the review of this manuscript and approving it for publication was Chien-Fu Cheng.

In asynchronous random-access protocols, packet collisions occur when multiple nodes simultaneously transmit packets using the same frequency channel. Once a packet collision occurs, the receiver cannot demodulate the packets correctly. The traffic in LPWAN is dominated by uplink communication for data collection [5], and packet collisions become a serious problem when the number of LPWAN nodes is large. One cause of packet collision in LPWAN is the simultaneous packet transmission triggered by event detection [5]. An example of an event is a change in the soil moisture content in land under smart farming that sets off a trigger [6]. If multiple nodes detect an event, they simultaneously transmit data packets. Thus, owing to packet collisions along the way, the gateway (GW) may not detect the event. Furthermore, there are limitations on the duty cycle (DC) in LPWAN. The DC determines the transmission time ratio at which each LPWAN node can occupy a specific frequency channel. The LPWAN node accessing the unlicensed bands should set DC in accordance with the regulation of each county or region for frequency sharing among other systems [7]. If a packet collision occurs, the LPWAN cannot retransmit the packet immediately but has to wait for a certain period decided by the DC. The latency caused by the DC for packet retransmission is unacceptable for applications that need to recognize events immediately. Therefore, developing a packet collision avoidance technology is urgently needed.

One such technology in LPWAN is the carrier-sense multiple access/collision avoidance (CSMA/CA), also known as listen-before-talk (LBT). Each node checks the frequency channel before transmitting a packet [8]. However, the CSMA/CA approach may not be an appropriate solution for LPWAN because of the increased node energy consumption that takes place due to carrier sensing [9]. The success of the carrier sense (CS) depends on the CS threshold. Moreover, the hidden node problem occurs because there is a high probability that signals cannot be detected by the CS owing to the large coverage area in LPWAN [10].

Another approach to avoid packet collision is to allocate appropriate resources to the nodes. Various resource management schemes have been proposed to address this issue [11]–[13]. In [11], the optimization problem has been formulated to maximize the average system packet success probability in LPWAN, and LoRaWAN. A quasi-optimal spreading factor (SF) allocation algorithm is proposed based on the aforementioned formulated problem. A centralized channel allocation algorithm based on the matching theory was proposed by [12]. In [13], a joint allocation problem of SF and power was formulated, and an algorithm based on it was proposed. These approaches are optimized in a static environment that depends on a formulated mathematical model. Recently, machine-learning-based resource management schemes for IoT networks have been proposed [14]–[19] by applying reinforcement learning, such as Q-learning [20] and multi-armed bandit learning [21]. Reinforcement learning can realize dynamic resource allocation in response to the environment because the learning process calculates the reward based on feedback from the environment [22]. However, these models do not capture probabilistic factors, such as event-triggered traffic. In [5], the traffic of LPWAN was modeled and the network performance was evaluated by considering an event-triggered traffic. However, packet collision avoidance in event-triggered traffic has not yet been discussed. In [5], a resource allocation scheme for the IoT considered unusual traffic due to anomaly detection. This approach improves the throughput by allocating frequency channels and backoff window sizes to the nodes. However, dynamic resource allocation using this scheme may require the implementation of complex protocols and synchronization mechanisms [23].

Therefore, the LPWAN needs a dynamic and low-overhead resource allocation strategy. It also needs to accommodate event-triggered traffic by suppressing packet collisions.

This paper proposes an autonomous decentralized traffic control system using Q-learning that ensures packet collision avoidance and improves the communication quality. It should be noted that similar to the ALOHA and CSMA/CA, the proposed scheme does not require any synchronization between LPWAN nodes; in particular, it consists of two steps. First, the LPWAN node detecting the event determines the transmission timing offset autonomously using Q-learning. When an event occurs, several LPWAN nodes may detect an event almost simultaneously. Strategically shifting event packet transmission timing can effectively avoid packet collisions. Second, the LPWAN nodes probabilistically transmit event packets. When multiple nodes observe the same event, the information transmitted by them is highly correlated. Thus, there is little need for all LPWAN nodes to transmit packets to a GW. A decrease in the number of event packets leads to a decrease in the probability of packet collisions. The proposed scheme has the following three advantages:

(i) It does not need complex control signals and synchronization between the LPWAN nodes to allocate wireless resources. In other words, asynchronous management between LPWAN nodes is possible.
(ii) It can realize dynamic resource allocation independent of the LPWAN system model as a result of adopting reinforcement learning. The LPWAN nodes converge to a state suitable for the given system.
(iii) It is a powerful and simple algorithm that can be run on LPWAN nodes with a low computational capability.

Computer simulation results show that the proposed scheme can improve the average packet delivery rate (PDR) by 60% compared to the conventional ALOHA protocol in the LoRaWAN environment [24].

The remainder of this paper is organized as follows. In Section II, we describe the proposed system model. In Section III, an autonomous decentralized traffic control is proposed. In Section IV, simulation results are provided to demonstrate the effectiveness of the proposed scheme. Section VI concludes the paper.

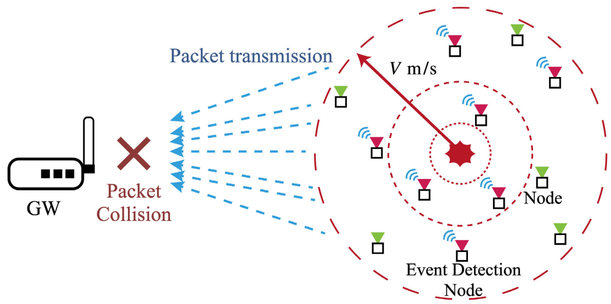


FIGURE 1. Event model.

II. SYSTEM MODEL

A. NETWORK MODEL

We consider a WSN based on LPWAN, where I LPWAN nodes ($\mathcal{I} = \{1, \cdots, i, \cdots, I\}$) are randomly and uniformly distributed within the communication area of $L \times L$ km^2. The LPWAN nodes are connected to a single GW located at the center of the communication area. The system has K orthogonal frequency channels ($\mathcal{K} = \{1, \cdots, k, \cdots, K\}$), and each LPWAN node selects a frequency channel and transmits a packet to the GW.

B. EVENT GENERATION AND DETECTION

LPWAN nodes detect events that occur within a communication area. An event occurs at a specific location, which is determined randomly. An LPWAN node detects the event; the LPWAN transforms the event into numeric data called event true data, $x \in [x_{\min}, x_{\max}]$, where $x_{\min}$ and $x_{\max}$ are the minimum and maximum numerical values, respectively. The event occurrence information with event true data x propagates from the event epicenter with a speed of V m/s outwardly in a circle, as shown in Fig. 1. LPWAN node i detects an event with a probability of δ_i, which is given by [5]

$$\delta_i = e^{-\alpha d_{\mathrm{e},i}}, \tag{1}$$

where α is the event propagation coefficient and $d_{\mathrm{e},i}$ m is the distance between the LPWAN node i and the event epicenter. The LPWAN node i that detects an event observes the sensing data x_i^{sens}, which is the event true value with an error based on the sensor accuracy. The sensing data, x_i^{sens}, are given by

$$x_i^{\mathrm{sens}} = x + e_i, \tag{2}$$

where e_i is a Gaussian random variable that follows a standard normal distribution $\mathcal{N}(0, 1)$.

C. TRAFFIC MODEL

1) PACKET GENERATION

This study assumes two types of traffic: *regular traffic* and *event-triggered traffic*. Regular traffic represents the data packets transmitted periodically from each LPWAN node. Such data packets are generated at every predetermined packet generation interval G_{p} sec. The LPWAN node i transmits the first periodic packet at $T_{\mathrm{offset},i}$ determined by a random number generated according to $\mathcal{U}\left(0, G_{\mathrm{p}}\right)$, where $\mathcal{U}(a, b)$ is a uniform random variable generated in the range (a, b). Event-triggered traffic represents the data packets generated by event detection.

2) PACKET TRANSMISSION

In this paper, we define the transmission phase as the period from the start of packet transmission to its end. Without loss of generality, all packets are assumed to have the same packet length and DC constraint for packet transmission. Thus, each LPWAN node must satisfy the DC constraint. We define the DC constraint as

$$T_{\mathrm{DC},i} = \left(\frac{1 - D_{\mathrm{c}}}{D_{\mathrm{c}}}\right) T_{\mathrm{L},i}, \tag{3}$$

where $D_{\mathrm{c}} \in (0, 1]$ is the DC, and $T_{\mathrm{L},i}$ sec is the packet transmission duration. $T_{\mathrm{DC},i}$ sec denotes the minimum time required to wait for the next packet to be transmitted to satisfy the DC constraint. In the first case, the LPWAN node i generates a new packet during the packet transmission phase. In the second case, the LPWAN node i generates a new packet within $T_{\mathrm{DC},i}$.

Because the collection of event information is essential, the LPWAN node transmits the event packet as a confirmed message to guarantee successful communication. When the GW successfully receives the event packet of the confirmed message, it replies with an acknowledgment (ACK) signal to the sender LPWAN node.

III. PROPOSED SCHEME

This section describes the proposed autonomous decentralized traffic control scheme using Q-learning. The scheme comprises two steps: controlling the event packet transmission probability and controlling the event packet transmission timing.

A. EVENT PACKET TRANSMISSION PROBABILITY

When an event occurs, multiple LPWAN nodes detect it almost simultaneously. Thus, once an event occurs, an event-triggered traffic dominates the network traffic. It is challenging to determine the number of LPWAN nodes that simultaneously detect an event precisely. As all event-triggered packets contain the event true data with an observation error at each LPWAN node, there is a high correlation among the obtained data. Therefore, the GW can accurately estimate the true event data with a small number of event packets. Thus, limiting the number of nodes that transmit event packets can reduce the packet collision probability and improve the communication quality. To limit the number of nodes transmitting event packets, we define an event packet transmission probability $p_{\mathrm{s},i} \in (0, 1]$ for the LPWAN node i. The event packet transmission probability $p_{\mathrm{s},i}$ determines whether to transmit an event packet, as shown in Fig.2.

The event packet transmission probability $p_{\mathrm{s},i}$ at the LPWAN node i is automatically controlled based on the ACK signal as

$$p_{\mathrm{s},i} = \frac{1 + M_i^{\mathrm{ack}}}{1 + M_i^{\mathrm{tran}}}, \tag{4}$$

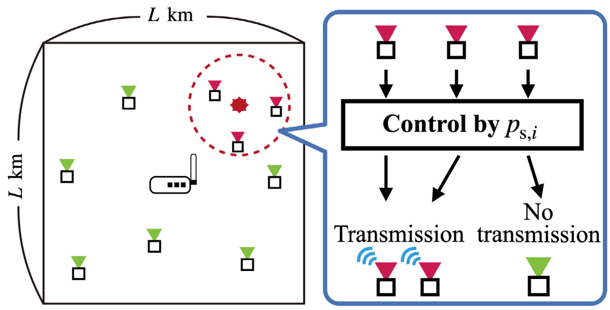


FIGURE 2. Reduction in the number of transmitting devices by $p_{s,i}$.

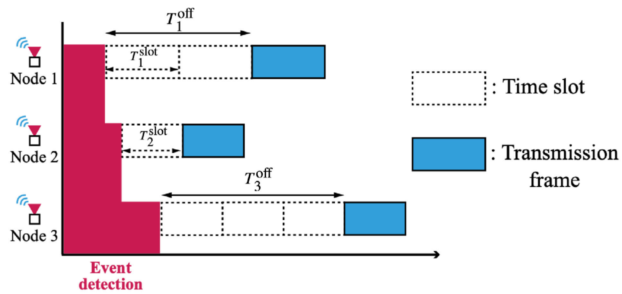


FIGURE 3. Event packet transmission with a time slot.

where M_i^{ack} is the number of ACK signals and M_i^{tran} is the total number of event packets transmitted by the LPWAN node i. By evaluating the event packet success rate at each node using (4), the transmission probability of nodes with a low event packet success rate is reduced. Nodes with low event packet success rates are most likely to undergo packet collisions, which may negatively affect the other LPWAN nodes.

B. ADAPTIVE RESOURCE ALLOCATION USING Q-LEARNING

The communication resources generally used in LPWAN are the time and frequency channels. As stated earlier, when an event occurs, multiple LPWAN nodes detect it almost simultaneously and transmit event packets. As a result, packets are transmitted densely in the time domain, which causes packet collisions. Thus, packet collisions can be avoided by strategically shifting the transmission timing. In the proposed approach, each LPWAN node detects an event and then autonomously selects the transmission timing offset determined by reinforcement learning.

1) Q-LEARNING

Q-learning is a popular reinforcement learning technique; it learns actions that maximize the reward for the environment through repeated trial and error. Because Q-learning does not require computationally intensive calculations, such as matrix operations, it is easy to run it on LPWAN nodes. In Q-learning, a learning machine is called an agent. In our proposed scheme, each LPWAN node was installed as a Q-learning agent. The Q-learning agent observes a state that indicates environmental information and decides an action based on a Q-table. The Q-table stores Q-values, which represent the evaluation values of actions in a specific state. The agent updates the Q-table based on the rewards received from the environment through its actions, which is called exploration. If the exploration is sufficiently successful, it is guaranteed to converge to the optimal solution that maximizes the reward [25]. However, the larger the Q-table size, the larger the exploration volume becomes; thus, it is difficult to conduct an exploration in a real system. Thus, we propose a resource allocation scheme using Q-learning with as small a Q-table as possible. The agent can observe its own frequency channel and the transmission timing offset. However, it is possible to reduce the amount of exploration by considering only the transmission timing offset in the Q-table. Therefore, the frequency channel used by the LPWAN node is first randomly assigned, and subsequently, only the assigned frequency channel is used.

2) TRANSMISSION TIMING OFFSET

Let $T_{L,i}$ be the packet transmission duration of LPWAN node i. The LPWAN node i is randomly allocated a set of transmission timing offset indices $\mathcal{D}_i = \{0, D_{i,1}, \cdots, D_{i,j}, \cdots, D_{i,J} | D_{i,j} \sim \mathcal{U}'(1, D_{\max})\}$, where J is the total number of candidate transmission timing offset index, $\mathcal{U}'(a, b)$ is an integer uniform random number generated in the range $[a, b]$, and $D_{\max}$ is the maximum transmission offset index. Upon detecting an event, LPWAN node i waits for the transmission timing offset T_i^{off} sec and then transmits an event packet, as shown in Fig. 3. The transmission timing offset T_i^{off} is given as follows:

$$T_i^{off} = T_i^{slot} \times D_{i,j^\star}, \tag{5}$$

where T_i^{slot} sec is a time slot of the same length as the packet transmission duration $T_{L,i}$ and $D_{i,j^\star} \in \mathcal{D}_i$ is the selected transmission timing offset index with $j^\star$ being the transmission timing offset index determined by Q-learning.

3) LEARNING MODEL

In this paper, epochs represent periods of real-time length that are pre-determined for Q-learning. Thus, an epoch is defined as follows:

(i) LPWAN node i determines its own transmission timing offset index $D_{i,j^\star}$.
(ii) An event occurs and the node detects it.
(iii) After waiting for T_i^{off}, the LPWAN node i transmits an event packet with a probability $p_{s,i}$.
(iv) Reward is calculated and Q-value updated.

Let us define the set of states and set of actions $\mathcal{A}$ in Q-learning as follows:

- Set of state $\mathcal{D}_i$: The set of selectable transmission timing offset indices of the LPWAN node i.
- Set of action $\mathcal{A}$: The set of changes in the transmission timing offset T_i^{off}. The elements are denoted by $\mathcal{A} = \{1, 0, -1\}$, where $1, 0, -1$ indicates that $D_{i,j^\star}$ is set larger, kept unchanged, or set smaller, respectively.

The agent of LPWAN node i is defined as

- State $s_{i,z} \in \mathcal{D}_i$: The transmission timing offset index of node i observed by the agent in epoch z
- Action $a_{i,z} \in \mathcal{A}$: The transmission timing offset index change in epoch z.
- Reward $r_{i,z}$: Reward value from the environment by taking action $a_{i,z}$.
- Q-value $Q\left(s_{i,z}, a_{i,z}\right)$: Value of action $a_{i,z}$ at state $s_{i,z}$.

For an efficient exploration of the Q-table, we adopt the ε-greedy algorithm that uses probability $\varepsilon \in [0, 1]$ to explore and exploit the Q-table. At epoch z, $\varepsilon_{i,z}$ is given as

$$\varepsilon_{i,z} = 1 - \frac{M_i^{\mathrm{tran}}}{Z}, \tag{6}$$

where Z is the total number of epochs. The exploration is performed with a probability $\varepsilon_{i,z}$, and the exploitation is performed with a probability $1 - \varepsilon_{i,z}$. Then, the Q-value is updated as

$$Q\left(s_{i,z}, a_{i,z}\right) = Q\left(s_{i,z}, a_{i,z}\right) + \eta E_{i,z}^{\mathrm{TD}}, \tag{7}$$

where $E_{i,z}^{\mathrm{TD}}$ is the temporal difference error and η is the Q-learning rate. A temporal difference error $E_{i,z}^{\mathrm{TD}}$ is given as

$$E_{i,z}^{\mathrm{TD}} = r_{i,z+1} + \beta \left(\max_{a^\star \in \mathcal{A}(s_{i,z+1})} Q(s_{i,z+1}, a^\star) - Q\left(s_{i,z}, a_{i,z}\right) \right), \tag{8}$$

where β is the discount rate. In Q-learning, the Q-value is updated by considering the maximum reward value that can be expected in a transitioning state at an epoch $z + 1$.

4) REWARD FUNCTION

In Q-learning, the reward function is vital because it learns the action that maximizes the sum of the rewards. We aim to determine the timing of all the packet transmissions to avoid any packet collisions. Therefore, we adopted the ACK signal as a measure of packet collision avoidance. If the LPWAN node receives the ACK signal, it positively evaluates the transmission timing offset used for event packet transmission. However, if the LPWAN node does not receive the ACK signal, the Q-learning agent will negatively evaluate the transmission timing offset. Thus, the reward function is given as

$$r_{i,z} = \begin{cases} 1 & \text{if ACK is received} \\ -1 & \text{otherwise} \end{cases}. \tag{9}$$

C. TRAFFIC CONTROL ALGORITHM

The LPWAN node i transmits a packet using the frequency channel k_i, which is randomly selected from $\mathcal{K}$ and is then fixed. Algorithm 1 shows the traffic control algorithm for node i. The LPWAN node i that detects an event decides to transmit the event packet by taking into consideration the event packet transmission probability $p_{\mathrm{s},i}$. When an event packet is transmitted, the LPWAN node i determines the transmission timing offset T_i^{off} based on Q-learning. When the predetermined total number of epochs Z has elapsed,

Algorithm 1 Traffic Control Algorithm for Node i

Input:
  $\mathcal{K} = \{1, \cdots, k, \cdots, K\}, T > 0, J > 0,$
  $\mathcal{D}_i = \{0, D_{i,1}, \cdots, D_{i,j}, \cdots, D_{i,J} | D_{i,j} \sim \mathcal{U}'(1, D_{\max})\}$
Initialization:
  Epoch counter $z = 1$
  Allocate frequency channel $k_i \in \mathcal{K}$
  Transmission probability $p_{\mathrm{s},i} = 1$
  Set of states $\mathcal{D}_i$
  Set of actions $\mathcal{A} = \{1, 0, -1\}$
  Q-table $Q\left(s_{i,z}, a_{i,z}\right) \sim \mathcal{U}(0, 1)$
while $z < Z$ **do**
  Determine $D_{i,j^\star}$ by Q-learning agent
  if Generate event packet **then**
    if $p_{\mathrm{s},i} > \mathcal{U}(0, 1)$ **then**
      Transmit event packet
      Update $M_i^{\mathrm{tran}} \leftarrow M_i^{\mathrm{tran}} + 1$
      Calculate reward $r_{i,z}$ by (9)
      Update Q-table $Q\left(s_{i,z}, a_{i,z}\right)$ by (7)
    else
      Discard event packet
    end if
    Update $p_{\mathrm{s},i}$ by (4)
  end if
  Update $z \leftarrow z + 1$
end while
Output:
  Fix $p_{\mathrm{s},i}$ and $D_{i,j^\star}$

Q-learning is terminated and LPWAN node i fixes the event packet transmission probability $p_{\mathrm{s},i}$ and transmission timing offset T_i^{off}.

IV. SIMULATION AND RESULTS

A. LoRaWAN MODEL

For the computer simulation of our proposed scheme, we adopted LoRaWAN. LoRaWAN is a popular LPWAN technology developed and commercialized by Semtech Corporation. It adopts a unique modulation scheme in the physical layer based on CSS modulation and frequency shift keying (FSK) called LoRa modulation [26]. LoRa modulation has a high level of interference tolerance. This is because it spreads a narrowband signal into a wider bandwidth, thus reducing the noise level of the output signal [27]. Thus, LoRa modulation can realize long-range communication. An important parameter in LoRa modulation is the SF, which determines the number of transmitted bits in a CSS-modulated symbol [28]. The SF takes values from 7 to 12. Let SF be $S_i^{\mathrm{SF}} \in \mathcal{S} = \{7, 8, 9, 10, 11, 12\}$ at LoRaWAN node i, and one CSS-modulated symbol time length $T_{\mathrm{s},i}(S_i^{\mathrm{SF}})$ is given by

$$T_{\mathrm{s},i}(S_i^{\mathrm{SF}}) = \frac{2^{S_i^{\mathrm{SF}}}}{W}, \tag{10}$$

where W Hz is the frequency bandwidth. The LoRaWAN packet physical structure consists of a preamble, a synchronization word, a physical header, a header cyclic redundancy check, a physical payload, and a payload cyclic redundancy check (CRC) [29]. Let the number of symbols other than the physical payload and CRC be the number of overhead symbols $O_{\rm sym}$. Let the number of packet data sizes be $B_{{\rm data},i}$ bit that includes the physical payload and CRC. Thus, the required number of CSS symbols $N_{{\rm s},i}(S_i^{\rm SF})$ to transmit one data packet is given by [30]

$$N_{{\rm s},i}(S_i^{\rm SF}) = O_{\rm sym} + \left\lceil \frac{B_{{\rm data},i}/R}{S_i^{\rm SF}} \right\rceil, \tag{11}$$

where $\lceil x \rceil$ is the ceiling function of x, and R is the coding rate. Therefore, packet transmission duration $T_{{\rm L},i}$ is given as

$$T_{{\rm L},i} = T_{{\rm s},i}(S_i^{\rm SF}) \times N_{{\rm s},i}(S_i^{\rm SF}). \tag{12}$$

B. CHANNEL MODEL

In this paper, path loss and log-normally distributed shadowing loss were considered for a channel model. The received signal power of the LoRaWAN node i at the GW is given as

$$P_{{\rm r},i} = P_{\rm t} - P_{\rm Loss}(d_i) - \psi, \tag{13}$$

where $P_{\rm t}$ dBm is the transmission power common to all LoRaWAN nodes, and ψ dB is a log-normally distributed shadowing. From [31], the path loss component $P_{\rm Loss}(d_i)$ dB is given as

$$P_{\rm Loss}(d_i) = 10\mu \log_{10} d_i + \nu + 10\xi \log_{10} f_{\rm c} + \zeta, \tag{14}$$

where d_i m is the physical distance between LoRaWAN node i and GW, propagation parameters μ, ν, ξ are the path loss coefficient, offset, and frequency loss component, respectively, $f_{\rm c}$ GHz is the carrier frequency, and ζ is a Gaussian random variable of the distribution of $\mathcal{N}(0, \sigma)$. The signal-to-noise power ratio (SNR) and signal-to-interference power ratio (SIR) of LoRaWAN node i observed at the GW are given by

$$\begin{cases} \gamma_{{\rm SNR},i} = P_{{\rm r},i} - \left(N_0 + 10\log_{10} W + NF\right) \\ \gamma_{{\rm SIR},i} = P_{{\rm r},i} - \sum_{i' \in \mathcal{I}_i} P_{{\rm r},i'}, \end{cases} \tag{15}$$

where N_0 dBm/Hz is the noise power spectrum density, NF is the noise figure, and $\mathcal{I}_i$ is the set of interfering LoRaWAN nodes that transmit packets using the same frequency channel as the LoRaWAN node i. In this study, we determine the success of packet demodulation through the SNR threshold $\Gamma_{{\rm SNR},S_i^{\rm SF}}$, and the SIR threshold $\Gamma_{{\rm SIR},S_i^{\rm SF}}$ at GW [14]. We use the SNR threshold and the SIR threshold to consider the capture effect due to an imperfect orthogonality between different SFs in LoRa modulation [32].

When only one packet is received by the GW, it is assumed that the packet is successfully demodulated if SNR $\gamma_{{\rm SNR},i}$ is above the threshold $\Gamma_{{\rm SNR},S_i^{\rm SF}}$. When multiple packets are received on the same frequency channel, we consider two cases [34].

- When the SF is different among the received signals, SIR thresholds $\Gamma_{{\rm SIR},S_i^{\rm SF}}$, which are shown in Table 1, are used [32], [33].
- When all received signals use the same SF $S^{\rm SF}$, SIR threshold $\Gamma_{{\rm SIR},S_i^{\rm SF}}$ is set to 6 dB.

TABLE 1. SNR and SIR thresholds [32], [33].

SF	SNR threshold $\Gamma_{{\rm SNR},S_i^{\rm SF}}$ dB	SIR threshold $\Gamma_{{\rm SIR},S_i^{\rm SF}}$ dB
7	-7.5	-11
8	-10	-13
9	-12.5	-16
10	-15	-19

TABLE 2. Simulation parameters.

Simulation area, $L \times L$	2×2 km^2
Simulation time	10 min
Number of LoRaWAN nodes, I	$\{500, 1000\}$
Transmit power, $P_{\rm t}$	13 dBm
Carrier frequency, $f_{\rm c}$	0.923 GHz
Bandwith, W	125 kHz
Number of frequency channels, K	$\{1, 2, 4, 8, 16\}$
SF, $\mathcal{S}$	$\{7, 8, 9, 10\}$
Coding rate, R	4/7
Duty cycle, $D_{\rm c}$	0.01
Noise power spectrum density, N_0	-174 dBm/Hz
Noise figure, NF	10 dB
Path loss coefficient, μ	4.0
Propagation offset, ν	9.5
Frequency loss component, ξ	4.5
Pathloss standard deviation, σ	7.6 dB
Shadowing standard deviation	3.48 dB
Overhead symbol, $O_{\rm sym}$	20.25
Packet data size, $B_{{\rm data},i}$	160 bit
Packet generation interval, $G_{\rm p}$	10 min
Max transmit offset slots, $D_{\rm max}$	64
Frequency of event occurrence	1 /epoch
Event propagation coefficient, α	0.01
Event propagation speed, V	1000 m/s
Minimum event true data value, $x_{\rm min}$	-50
Maximum event true data value, $x_{\rm max}$	50

TABLE 3. Q-learning parameters.

Time length of one epoch	10 min
Total number of epochs, Z	$\{50, 100, 500, 1000, 2000, 5000, 10000\}$
Q-learning rate, η	0.3
Discount rate, β	0.95
Total number of candidates transmission timing offsets indices, J	3

C. PARAMETER SET

The LoRaWAN system parameters and Q-learning parameters are listed in Tables 2 and 3. The LoRaWAN system parameters follow the Japanese parameter configuration AS923 [29]. In this study, the periodic packet is transmitted immediately after generation as an unconfirmed message that does not require an ACK signal from the GW. By contrast, an event packet is transmitted as a confirmed message. When the GW receives the event packet successfully, it transmits an ACK signal to the LoRaWAN node that transmits a confirmation message. We assume that the ACK signal is received by the LoRaWAN node ideally. This study considers a scenario where all LPWAN nodes observe the same target. Thus, it is reasonable to assume that the packet formats of all LPWAN nodes are the same, which results in the length of the packet being the same. Even if the packet length is different for different LPWAN nodes, the proposed scheme works as it does not assume any synchronization between the LPWAN nodes.

D. EVALUATION CRITERIA

In this study, we evaluate the communication quality based on the event PDR. The average event PDR is defined as

$$\mathrm{PDR} \triangleq \frac{N_{\mathrm{succ}}}{N_{\mathrm{tran}}}, \tag{16}$$

where N_{succ} is the number of event packets successfully received by the GW, and N_{tran} is the total number of event packets transmitted by the LoRaWAN nodes. The GW calculates an estimate of the event data $\hat{x}$ by averaging the successfully received data x_i^{sens} as

$$\hat{x} \triangleq \frac{1}{N_{\mathrm{succ}}} \sum_{i \in \mathcal{I}_{\mathrm{s}}} x_i^{\mathrm{sens}}, \tag{17}$$

where $\mathcal{I}_{\mathrm{s}}$ is the set of LoRaWAN nodes whose event packets have been successfully received by the GW. The squared error x_{SE} between the event estimate data $\hat{x}$ and the event true data x is defined as

$$x_{\mathrm{SE}} \triangleq (\hat{x} - x)^2. \tag{18}$$

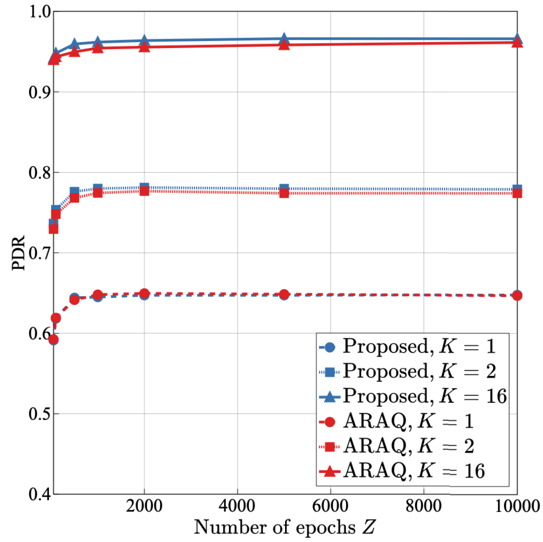
The squared error x_{SE} averaged over the total number of simulation runs is defined as the mean squared error (MSE), which is defined as

$$\mathrm{MSE} \triangleq \mathbb{E}[x_{\mathrm{SE}}], \tag{19}$$

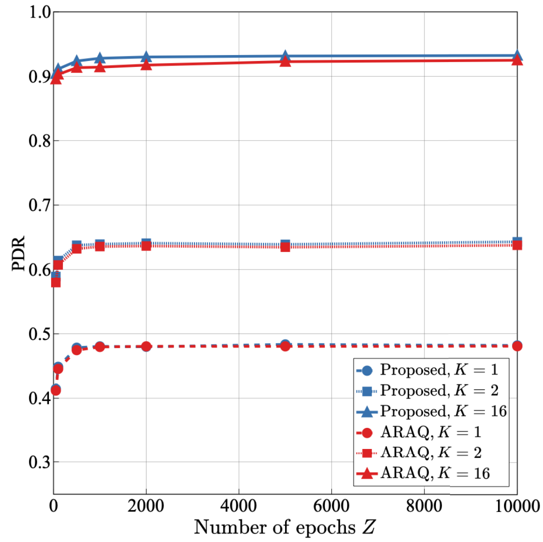
where $\mathbb{E}[a]$ is a function that averages a over the total number of simulation runs. We define the event detection probability at the GW. When the GW successfully receives one or more packets generated by the event, the event detection is considered successful. The event detection probability is calculated as the ratio of the number of event detections to the number of event occurrences.

E. BENCHMARK SCHEME

We describe three schemes to demonstrate the effectiveness of the proposed scheme in this paper. 1. The pure ALOHA protocol, which is the most common MAC layer access protocol used in LoRaWAN. The ALOHA scheme transmits an event packet immediately after the event packet is generated. 2. The RANDOM scheme, which determines the transmission timing offset index $D_{i,j^\star}$ randomly at each transmission of an event packet. 3. All resource allocation by Q-learning (ARAQ) scheme, which selects both the frequency channel and the transmission timing offset index using Q-learning. The frequency channel and transmission timing can be arbitrarily selected by the LPWAN node. Thus, the ARAQ scheme selects the frequency channel $f_{i,z} \in \mathcal{K}$ in addition to the time slot offset through Q-learning at epoch z. Thus, in ARAQ scheme, the set of Q-learning action is defined as $\mathcal{A}'$, which contains combinations of frequency channel and transmission timing offset change action.



(a) $I = 500$



(b) $I = 1000$

FIGURE 4. **Impact of number of epochs Z on average PDR performance.**

F. NUMERICAL RESULTS

1) IMPACT OF EPOCHS

Fig. 4 shows the impact of learning epochs Z on the average PDR performance. Fig. 4 shows that the performance of the proposed method significantly improves at 500 epochs, which equals 5000 min in the environment considered in this study. Note that because the proposed scheme learns based on the experience of transmitting event packets, the required learning period depends on the frequency of event occurrence. The results show that the average PDR performance of the proposed scheme converges at approximately 2000 epochs. When multiple frequency channels are available, the proposed scheme performs better than the ARAQ scheme. The number of elements in the Q-table is $(J+1) \times n(\mathcal{A})$ for the proposed scheme, where $n(\mathcal{X})$ is the number of elements in the set $\mathcal{X}$. On the other hand, it is $(J+1) \times n(\mathcal{A}) \times n(\mathcal{K})$ in the ARAQ scheme because of the action set $\mathcal{A}'$. Thus, the ARAQ scheme is likely to be insufficient for exploration compared to the proposed scheme owing to the large Q-table size. If the LPWAN nodes are given sufficient learning time, we presume that the ARAQ scheme may outperform the proposed scheme because the ARAQ scheme considers the time and frequency channels. However, Fig. 4 shows that increasing the $Z = 2000$ to $Z = 10000$ only slightly improves the performance of the ARAQ scheme; thus, it is difficult for the ARAQ scheme to outperform the proposed scheme in real time. Therefore, our proposed scheme can achieve a good average PDR performance even when the frequency channels are randomly allocated. Hereafter, the performance of all the schemes with Q-learning was obtained after the number of epochs $Z = 2000$.

2) AVERAGE PDR PERFORMANCE OF EACH SCHEME

Fig. 5 shows the impact of the number of channels K on the average PDR performance. In Fig. 5(a), the proposed scheme provides a better average PDR performance compared with the ALOHA scheme irrespective of the number of frequency channels K. When the number of available frequency channels is small, the improvement in the average PDR performance by the proposed scheme is high. The average PDR performance of the proposed scheme improves by 60% compared to that of the ALOHA scheme when the number of available frequency channels is $K = 2$. The proposed scheme depicts a better average PDR performance compared to the case where the allocated frequency channel is not fixed and is changed for each packet transmission. This is because, after allocating a frequency channel, the proposed scheme can learn an appropriate transmission timing offset index $D_{i,j^\star}$ on the allocated channel by fixing the frequency channel to be used. The proposed scheme improves the average PDR performance irrespective of the number of frequency channels K to the scheme without calculating the event packet transmission probability $p_{s,i}$. The smaller the number of available frequency channels, the higher is the contribution of the event transmission probability $p_{s,i}$ to the average PDR performance. This is because the proposed scheme reduces the packet collision probability by reducing the number of transmitting LoRaWAN nodes with an event transmission probability $p_{s,i}$ based on the ACK signal. In addition, the proposed scheme can significantly improve the average PDR performance compared to the RANDOM scheme. Although

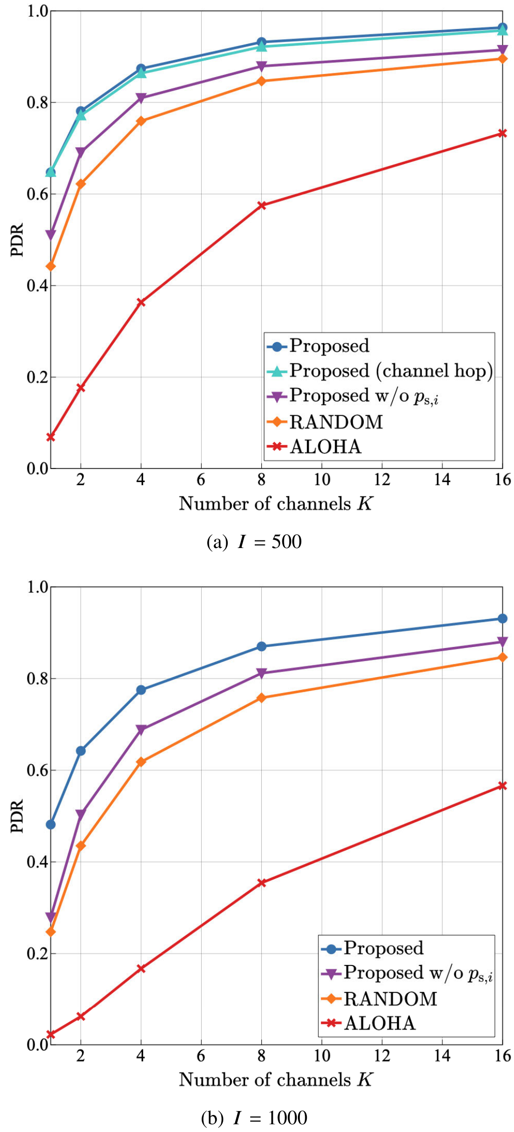
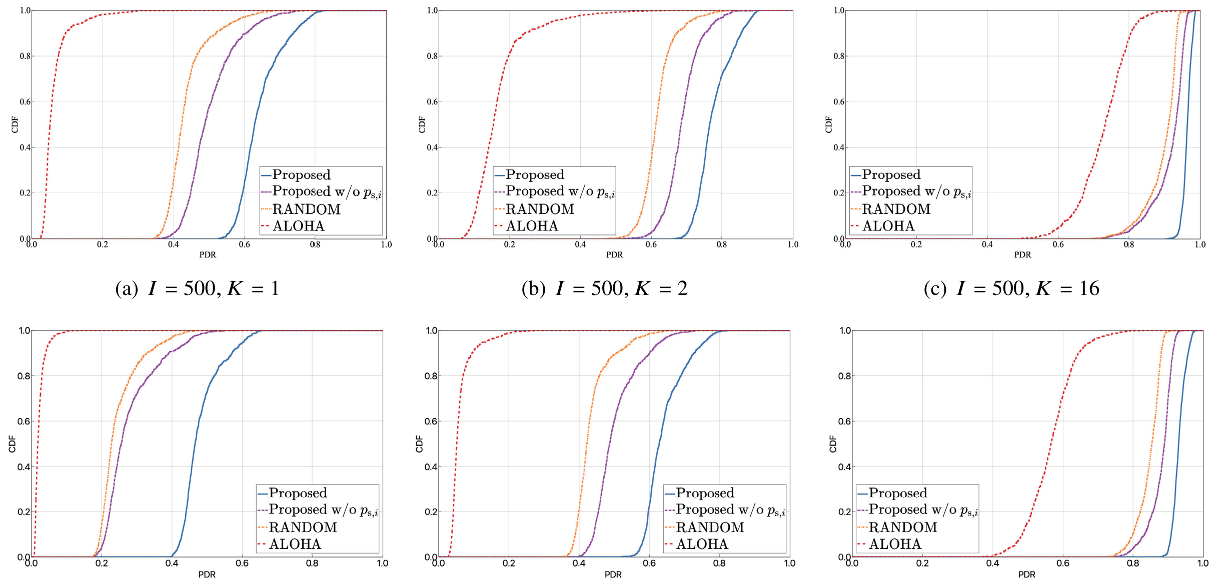


FIGURE 5. Impact of number of channels K on average PDR performance.

the proposed scheme requires a learning period, it can reduce the probability of packet collision because it can allocate an appropriate transmission timing offset index. Fig. 5 (b) shows that the proposed scheme is more effective than the other schemes for all channel numbers in the case of a large number of LoRaWAN nodes with a high network load. When the number of LoRaWAN nodes is $I = 1000$, the proposed scheme can improve by 61% compared to the ALOHA scheme when $K = 4$.

The cumulative distribution function (CDF) of the average PDR performance is shown in Fig. 6. Fig. 6 shows that the proposed scheme has a better CDF of average PDR than the other schemes and is independent of the number of LoRaWAN nodes I and number of frequency channels K. Moreover, the proposed scheme exhibits no variation in the average PDR performance when the number of frequency channels is large.



(a) $I = 500, K = 1$ (b) $I = 500, K = 2$ (c) $I = 500, K = 16$

(d) $I = 1000, K = 1$ (e) $I = 1000, K = 2$ (f) $I = 1000, K = 16$

FIGURE 6. **CDF of average PDR.**

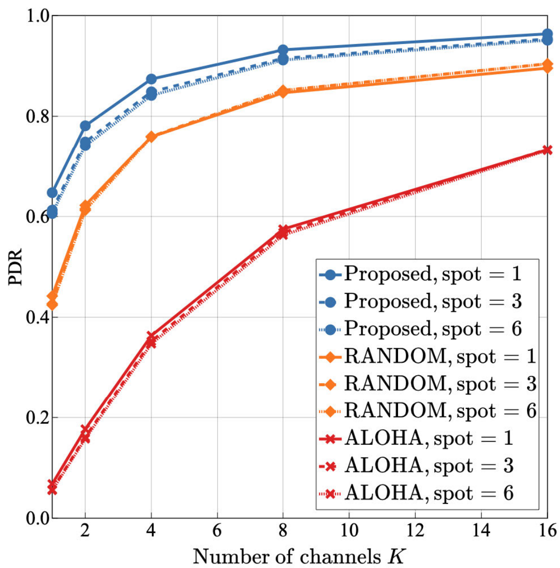


FIGURE 7. **Effect of the number of locations where events occur.**

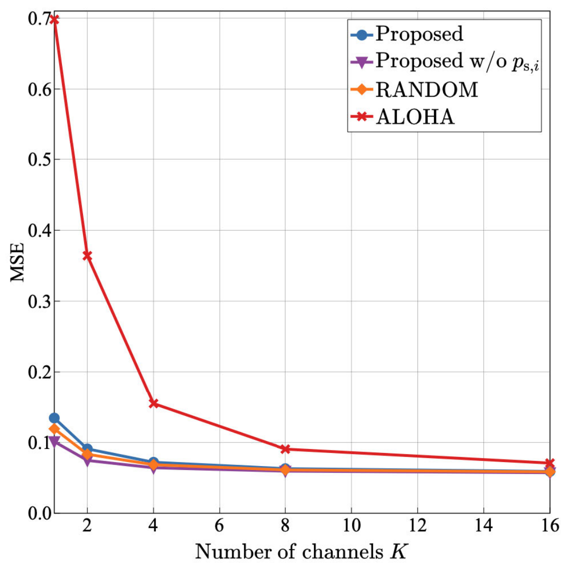


FIGURE 8. **MSE performance.**

Fig. 7 shows the impact of the number of candidate event occurrence spots on the average PDR. The average PDR performance of the proposed scheme degrades as the number of candidate event occurrence points increases. This is because the number of events detected by the LoRaWAN node within the learning period decreases, and the amount of Q-table exploration decreases; hence, it is not possible to allocate an appropriate transmission timing offset index.

3) EVENT DETECTION CAPABILITY

Fig. 8 shows the MSE performance of an event true value estimation. The proposed scheme provides good MSE performance independent of the number of channels K compared to the ALOHA scheme. In particular, when the number of frequency channels K is small, the proposed scheme has a better MSE performance compared with the ALOHA scheme and can reduce the MSE by 81% when $K = 1$. This is because the proposed scheme dramatically reduces packet collision and increases the number of event packets that can be successfully demodulated at the GW by using the event packet transmission probability $p_{s,i}$ and adaptive resource allocation using Q-learning. Conversely, the scheme without an event packet transmission probability $p_{s,i}$ exhibits the best MSE performance when $K = 1, 2$ and 4. This is because the probability of the number of event packets successfully demodulated by the GW is high as the number of transmission

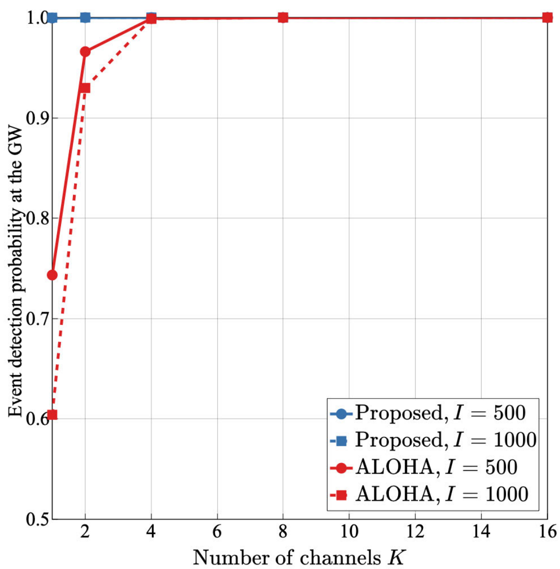


FIGURE 9. Event detection probability at the GW.

LoRaWAN nodes is not reduced by the event packet transmission probability $p_{s,i}$. In contrast, when $K = 8$ and 16, the impact of probability-based transmission control on the MSE performance is small because the network capacity is large and a high event packet transmission probability is likely to be allocated.

Fig. 9 shows the event detection probability at the GW. When the number of channels is small, the GW may not be able to detect the event occurrence due to severe packet collision with the ALOHA protocol. Conversely, the proposed scheme can detect an event at the GW with a probability 1, which is independent of the number of frequency channels.

V. CONCLUSION

In this paper, an autonomous decentralized traffic control method for wireless sensor nodes using Q-learning is proposed. The proposed scheme is a powerful autonomous decentralized resource allocation algorithm that aims to avoid packet collisions without the need for any additional control signals. In the proposed scheme, the packet transmission is controlled probabilistically based on the success rate of packet transmission, and the transmission offset is adaptively allocated by reinforcement learning.

We focused on the event-triggered traffic, which has a huge network load, and the proposed scheme was evaluated using computer simulations. The numerical results show that the proposed scheme can improve the average PDR performance by approximately 60% compared with the ALOHA scheme. Furthermore, the proposed scheme can detect events with a probability 1 at the GW with a reduced MSE. These results indicate that the proposed scheme can significantly contribute to the improvement of communication quality in WSNs.

REFERENCES

[1] O. Georgiou and U. Raza, "Low power wide area network analysis: Can LoRa scale?" *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 162–165, Apr. 2017.

[2] F.-Y. Leu, H.-L. Chen, and J.-C. Liu, "Improving multi-path congestion control for event-driven wireless sensor networks by using TDMA," in *Proc. 9th Int. Conf. Broadband Wireless Comput., Commun. Appl.*, Nov. 2014, pp. 300–305.

[3] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, 2nd Quart., 2017.

[4] A. Springer, W. Gugler, M. Huemer, L. Reindl, C. C. W. Ruppel, and R. Weigel, "Spread spectrum communications using chirp signals," in *Proc. Inf. Syst. Enhanced Public Saf. Secur. (IEEE/AFCEA EUROCOMM)*, 2000, pp. 166–170.

[5] V. Gupta, S. K. Devar, N. H. Kumar, and K. P. Bagadi, "Modelling of IoT traffic and its impact on LoRaWAN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[6] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu, and M. I. Ali, "Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications," in *Proc. IEEE 3rd World Forum Internet Things (WF-IoT)*, Dec. 2016, pp. 442–447.

[7] M. Chen, Y. Miao, X. Jian, X. Wang, and I. Humar, "Cognitive-LPWAN: Towards intelligent wireless services in hybrid low power wide area networks," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 2, pp. 409–417, Jun. 2019.

[8] J. Ortin, M. Cesana, and A. Redondi, "Augmenting LoRaWAN performance with listen before talk," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3113–3128, Jun. 2019.

[9] Z. Xie, R. Xu, and L. Lei, "A study of clear channel assessment performance for low power wide area networks," in *Proc. 10th Int. Conf. Wireless Commun., Netw. Mobile Comput. (WiCOM)*, 2014, pp. 311–315.

[10] L. Beltramelli, A. Mahmood, P. Österberg, and M. Gidlund, "LoRa beyond ALOHA: An investigation of alternative random access protocols," *IEEE Trans. Ind. Informat.*, vol. 17, no. 5, pp. 3544–3554, May 2021.

[11] J.-T. Lim and Y. Han, "Spreading factor allocation for massive connectivity in LoRa systems," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 800–803, Apr. 2018.

[12] Z. Qin and J. A. McCann, "Resource efficiency in low-power wide-area networks for IoT applications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–7.

[13] L. Amichi, M. Kaneko, E. H. Fukuda, N. El Rachkidy, and A. Guitton, "Joint allocation strategies of power and spreading factors with imperfect orthogonality in LoRa networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3750–3765, Jun. 2020.

[14] N. Aihara, K. Adachi, O. Takyu, M. Ohta, and T. Fujii, "Q-learning aided resource allocation and environment recognition in LoRaWAN with CSMA/CA," *IEEE Access*, vol. 7, pp. 152126–152137, 2019.

[15] A. Azari and C. Cavdar, "Self-organized low-power IoT networks: A distributed learning approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.

[16] R. Bonnefoi, L. Besson, C. Moy, E. Kaufmann, and J. Palicot, "Multi-armed bandit learning in IoT networks: Learning helps even in non-stationary settings," in *Proc. Int. Conf. Cognit. Radio Oriented Wireless Netw.*, 2018, pp. 173–185.

[17] D.-T. Ta, K. Khawam, S. Lahoud, C. Adjih, and S. Martin, "LoRa-MAB: Toward an intelligent resource allocation approach for LoRaWAN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[18] N. Jiang, Y. Deng, A. Nallanathan, and J. A. Chambers, "Reinforcement learning for real-time optimization in NB-IoT networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1424–1440, Jun. 2019.

[19] I. Ilahi, M. Usama, M. O. Farooq, M. U. Janjua, and J. Qadir, "LoRaDRL: Deep reinforcement learning based adaptive PHY layer transmission parameters selection for LoRaWAN," in *Proc. IEEE 45th Conf. Local Comput. Netw. (LCN)*, Nov. 2020, pp. 457–460.

[20] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[21] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," 2012, *arXiv:1204.5721*. [Online]. Available: http://arxiv.org/abs/1204.5721

[22] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3072–3108, Jun. 2019.

[23] L.-H. Shen, C.-H. Wu, W.-C. Su, and K.-T. Feng, "Analysis and implementation for traffic-aware channel assignment and contention scheme in LoRa based IoT networks," *IEEE Internet Things J.*, early access, Jan. 13, 2021, doi: 10.1109/JIOT.2021.3051347.

[24] N. Sornin, M. Luis, T. Eirich, T. Kramp, and O. Hersent, "LoRaWAN specification," LoRa Alliance, Fremont, CA, USA, Tech. Rep. v1.0, 2015.

[25] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, 1992.

[26] G. Zhu, C.-H. Liao, M. Suzuki, Y. Narusue, and H. Morikawa, "Evaluation of LoRa receiver performance under co-technology interference," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2018, pp. 1–7.

[27] B. Reynders, W. Meert, and S. Pollin, "Range and coexistence analysis of long range unlicensed communication," in *Proc. 23rd Int. Conf. Telecommun. (ICT)*, May 2016, pp. 1–6.

[28] D. Bankov, E. Khorov, and A. Lyakhov, "On the limits of LoRaWAN channel access," in *Proc. Int. Conf. Eng. Telecommun. (EnT)*, Nov. 2016, pp. 10–14.

[29] LoRa Alliance. (Dec. 2018). *LoRaWAN Regional Parameters V1.1rB*. [Online]. Available: https://lora-alliance.org/sites/default/files/2018-04/lorawantm_regional_parameters_v1.1rb_-_final.pdf

[30] L. Vangelista, "Frequency shift chirp modulation: The LoRa modulation," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1818–1821, Dec. 2017.

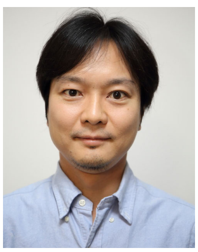
[31] P. Series, *Propagation Data and Prediction Methods for the Planning of Short-Range Outdoor Radiocommunication Systems and Radio Local Area Networks in the Frequency Range 300 MHz to 100 GHz*, document TR ITU-R, 2017. [online]. Available:https://www.itu.int/rec/R-REC-P.1411-9-201706-S/en

[32] D. Croce, M. Gucciardo, S. Mangione, G. Santaromita, and I. Tinnirello, "Impact of LoRa imperfect orthogonality: Analysis of link-level performance," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 796–799, Apr. 2018.

[33] *Semtech SX1272-Long Range, Low Power RF Transceiver 860-1000 MHz With LoRa Technology*. Accessed: Oct. 9, 2020. [Online]. Available: https://www.semtech.com/products/wireless-rf/lora-transceivers/sx1272/sx1272Datasheet

[34] A. Waret, M. Kaneko, A. Guitton, and N. El Rachkidy, "LoRa throughput analysis with imperfect spreading factor orthogonality," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 408–411, Apr. 2019.

AOTO KABURAKI (Graduate Student Member, IEEE) received the B.E. degree in information and communication engineering from The University of Electro-Communications, in 2020, where he is currently pursuing the M.E. degree. His research interest includes machine learning and its application to wireless communication.



KOICHI ADACHI (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from Keio University, Japan, in 2005, 2007, and 2009, respectively.

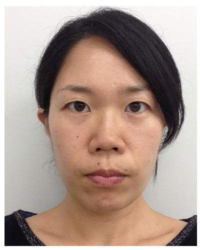
From 2007 to 2010, he was a Research Fellow of the Japan Society for the Promotion of Science (JSPS). He was a Visiting Researcher with the City University of Hong Kong, in April 2009, and a Visiting Research Fellow with the University of Kent, from June 2009 to August 2009. From May 2010 to May 2016, he was with the Institute for Infocomm Research, A*STAR, Singapore. He is currently an Associate Professor with The University of Electro-Communications, Japan. His research interests include cooperative communications and energy efficient communication technologies.

Dr. Adachi is a member of the IEICE. He is the coauthor of WPMC2020 Best Student Paper Award. He was awarded the Excellent Editor Award from IEEE ComSoc MMTC, in 2013. He served as the General Co-Chair for the 10th and 11th IEEE Vehicular Technology Society Asia Pacific Wireless Communications Symposium (APWCS), the Track Co-Chair for Transmission Technologies and Communication Theory of the 78th and 80th IEEE Vehicular Technology Conference, in 2013 and 2014, the Symposium Co-Chair for Communication Theory Symposium of IEEE GLOBECOM, in 2018, the Tutorial Co-Chair of IEEE ICC 2019, and the Symposium Co-Chair of Wireless Communications Symposium of IEEE GLOBECOM, in 2020. He has been an Associate Editor of *IET Communications*, from 2015 to 2017, IEEE WIRELESS COMMUNICATIONS LETTERS, since 2016, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, from 2016 to 2018, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, since 2016, and IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY, since 2019. He was recognized as an Exemplary Reviewer from IEEE WIRELESS COMMUNICATIONS LETTERS, in 2012, 2013, 2014, and 2015.

OSAMU TAKYU (Member, IEEE) received the B.E. degree in electrical engineering from the Tokyo University of Science, Chiba, Japan, in 2002, and the M.E. and Ph.D. degrees in open and environmental systems from Keio University, Yokohama, Japan, in 2003 and 2006, respectively.

From 2003 to 2007, he was a Research Associate with the Department of Information and Computer Science, Keio University. From 2004 to 2005, he was a Visiting Scholar with the School of Electrical and Information Engineering, The University of Sydney. From 2007 to 2011, he was an Assistant Professor with the Department of Electrical Engineering, Tokyo University of Science. From 2011 to 2013, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Shinshu University, where he has been an Associate Professor, since 2013. His current research interests include wireless communication systems and distributed wireless communication technology.

Dr. Takyu was a recipient of the Young Researcher's Award of IEICE 2010, the Active Research Award in radio communication systems (RCS) from IEICE Technical Committee on RCS, in 2010, and the Best Paper Award in smart radio (SR) from IEICE Technical Committee on SR, in 2018.



MAI OHTA (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from The University of Electro-Communications, Tokyo, Japan, in 2008, 2010, and 2013, respectively.

Since 2013, she has been an Assistant Professor with the Department of Electronics Engineering and Computer Science, Fukuoka University. Her research interests include cognitive radio, spectrum sensing, LPWAN, and sensor networks.

Dr. Ohta was a recipient of the Young Researcher's Award from IEICE, in 2013.

TAKEO FUJII (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1997, 1999, and 2002, respectively.

From 2000 to 2002, he was a Research Associate with the Department of Information and Computer Science, Keio University. From 2002 to 2006, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. From 2006 to 2014, he was an Associate Professor with the Advanced Wireless Communication Research Center, The University of Electro-Communications, where he is currently a Professor and the Director of the Advanced Wireless and Communication Research Center. His current research interests include cognitive radio and *ad hoc* wireless networks.

Dr. Fujii is a fellow of the IEICE. He was a recipient of the Best Paper Award in the IEEE VTC, in Fall 1999, the Active Research Award in radio communication systems (RCS) from the IEICE Technical Committee of RCS, in 2001, the Ericsson Young Scientist Award, in 2001, the Young Researcher's Award from the IEICE, in 2004, the Young Researcher Study Encouragement Award from IEICE Technical Committee of AN, in 2009, the Best Paper Award in the IEEE CCNC, in 2013, and the IEICE Communication Society Best Paper Award, in 2016.