Design and Implementation of SF Selection Based on Distance and SNR Using Autonomous Distributed Reinforcement Learning in LoRa Networks

Ikumi Urabe[1](✉), Aohan Li[2], Minoru Fujisawa[1], Song-Ju Kim[1,3],
and Mikio Hasegawa[1]

[1] Tokyo University of Science, Shinjuku City, Tokyo 125-8585, Japan
4321510@ed.tus.ac.jp
[2] The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan
[3] SOBIN Institute LLC, Kawanishi, Hyogo 666-0145, Japan

Abstract. LoRaWAN, one of Low-Power-Wide-Area (LPWA), has been deployed in many IoT applications due to its ability to communicate over long distances and low power consumption. However, the scalability and communication performance of LoRaWAN is highly dependent on Spreading Factor (SF) and Channel (CH) allocation. In particular, it is important to configure SF appropriately according to the distance of the LoRa device from the GateWay (GW) and the environment. In this paper, we implement and evaluate lightweight distributed reinforcement learning-based SF selection methods. This method allows each IoT device to make appropriate parameter selections without requiring any prior information, but only utilizing ACKnowledge obtained from its own transmissions. We then conducted real experiments in a small area indoors to verify that each LoRa device can autonomously and decentralized perform appropriate SF selection in response to the distance from the GW and SNR that varies depending on the surrounding environment. The results show that the implemented methods can select appropriate SF and achieve a better Frame Success Rate (FSR) than other lightweight approaches.

Keywords: IoT · LoRaWAN · Lightweight Distributed Reinforcement Learning · SF Selection

1 Introduction

IoT communication requires high energy efficiency, long-distance communication, and low cost, and a communication method called Low-Power-Wide-Area (LPWA), which satisfies these requirements, is attracting attention. LoRa, Sigfox, LTE-Cat, and NB-IoT are well-known LPWA networks. Among them, LoRaWAN has been adopted in many IoT applications because it is a low-cost and open-standard communication method.

Y. Kambayashi et al. (Eds.): AICON 2022, LNICST 477, pp. 34–42, 2023.
https://doi.org/10.1007/978-3-031-29126-5_3

With the rapid growth of LoRa devices, improving the scalability of LoRaWAN has become a major challenge.

LoRa networks use chirped spread spectrum (CSS) modulation. CSS is a spread spectrum technique that uses a chirp signal whose frequency increases linearly with time, making LoRaWAN highly resistant to interference. And its Spreading Factor (SF) determines the bit rate and reception sensitivity, which are in a trade-off relationship. In other words, The smaller the SF, the higher the transmission speed and the shorter the transmission time, but a high Signal to Noise Ratio (SNR) is required. Therefore, packets cannot be received by the GW unless the SF is appropriate for each LoRa device's distance from the GW and the surrounding conditions. From these, SF allocation has a significant impact on communication performance in LoRaWAN and must be optimized for each IoT application. Besides the access SF, the Channel (CH) is another important factor affecting the communication performance of LoRaWAN. And since packet collisions occur only when two or more devices choose the same CH and SF, the scalability of LoRaWAN is highly dependent on those parameter assignments.

Most of the research on parameter selection methods in LoRaWAN is based on a scheme called the centralized approach. In this approach, the GW assigns transmission parameters to each device. However, in the centralized approach, the GW needs to know all the information in advance, such as the number of devices, the location of each device, and the probability of an event occurring. Furthermore, each IoT device is required to be in a waking state to receive control signals from the GW, which increases battery consumption, and the occupancy of the channel by control signals increases transmission resource consumption. It also fails to account for interference from IoT devices in other applications that the GW is not aware of or from electronic devices emitting leaky radio waves, which has serious implications in the current environment of rapidly growing IoT devices. On the other hand, the distributed approach allows each device to make autonomous choices without requiring prior information, and the use of lightweight reinforcement learning algorithms allows for low-cost parameter optimization in terms of power and memory consumption.

In this paper, we implement and evaluate the Multi-Armed-Bandit (MAB) based decentralized SF selection methods. Each device is considered an intelligent agent, which is rewarded with Acknowledgement (ACK) information and chooses transmission parameters to maximize the cumulative reward. In [5], a channel selection method using the MAB algorithm was proposed, but it takes a huge number of iteration for the parameter selection of each device to converge, which is impractical in real environments where communication congestion is constantly and dynamically changing. In [6, 7], an algorithm called Tug-of-War (ToW) -dynamics was proposed, which is a small amount of calculation and a highly efficient search method for dynamically changing environments. Also in [8–10], ToW-dynamics is applied to the channel selection problem in wireless networks, and the proposed approach is implemented in IoT devices to verify its effectiveness. In this paper, we focus on SF selection in LoRa, implement the proposed method based ToW-dynamics [8] in a real LoRa device, and conduct experiments in a small indoor area. The goal of ToW-dynamics is to improve scalability by allowing each LoRa device to autonomously and decentrally select the SF with the least collisions and based on distance and surrounding environment. To demonstrate the performance of the proposed approach, we evaluate the Frame Success Rates (FSR) and compare them with other lightweight techniques using our LoRa devices equipped with the SF selection algorithms, which are located in an experimental field with different distances.

2 System Model and Problem Formulation

Fig. 1. System Model.

In this paper, we consider the star topology of LoRaWAN and assume a network with one GW and L LoRa devices. D is the LoRa device set, where D_l denotes the lth LoRa device. The LoRa devices are located in several locations, each at a different distance from the GW (Fig. 1). Assuming that the number of SFs is S, each LoRa device selects one SF among them to transmit packets. As shown in Sect. 1, LoRa employs chirp spread spectrum modulation so that signals with different SF (7–12) can be identified and received even if they are transmitted simultaneously on the same channel. And different SFs have different transmission speeds and different thresholds for the SNR that can be received. Table 1 shows the relationship between the bit rate and reception threshold. Theoretically, each spreading code is orthogonal, so collisions occur when two or more LoRa devices choose the same SF and CH. In practice, however, perfect orthogonality is not guaranteed, and interference between transmission using different SFs on the same CH must be considered [11, 12].

Each device selects not only the SF, but also the bandwidth from 62.5, 125, 250, and 500 [kHz], the transmit power from −4–13 [dBm], and the channel from 1–15 [CH]. In this paper, for ease of performance evaluation, we assume that all devices set the bandwidth to 125 kHz, the transmit power to maximum, and use the same channel. We also assume that all devices transmit M-byte packets of the same length with the same transmission interval TI.

A lightweight distributed reinforcement learning approach is implemented in each LoRa device and learns using ACK. At each decision, the device is considered as an intelligent agent that needs to strategically select SF based on the reinforcement learning approach. After decision and transmission, the LoRa device waits for ACK from the GW. If the transmission is successful, the LoRa device receives the ACK; otherwise, it does not. It is also assumed that there is no collision between this ACK and the uplink transmission. If ACK is not received, the packet is considered lost due to a collision with another packet transmitted on the same SF, or due to inter-SF interference.

In this paper, FSR was used to evaluate each approach. FSR in the LoRaWAN system at the t-th decision is defined as the ratio of the number of successful transmissions to

Table 1. LoRa Modulation Parameters at BW = 125 kHz.

SF	Bit Rate [kbps]	Receiver Sensitility [dBm]	SNR Thresh [dB]	Inter-SF collision Thresh [dB]
7	5.47	−123	−6	−7.5
8	3.13	−126	−9	−9
9	1.76	−129	−12	−13.5
10	0.98	−132	−15	−15
11	0.54	−133	−17.5	−18
12	0.29	−136	−20	−22.5

the total number of transmission attempts at the t-th decision and is expressed as:

$$FSR(t) = \sum_{m=1}^{M} \frac{r_t(t)}{n_l(t)} \tag{1}$$

where n_l is the number of transmission attempts at device l and r_l is the number of successful transmissions, i.e., the number of ACKs received. The objective function is expressed by the following equation:

$$(P) \qquad \max \sum_{t=1}^{T} FSR(t) \tag{2}$$

3 SF Parameter Selection for ToW-Dynamics

3.1 Multi-armed Bandit Problem

In the Multi-Armed Bandit (MAB) problem, a player selects a slot machine to play from among several slot machines. The player aims to maximize the amount of coins he or she earns by playing repeatedly. The player does not know the coin payout probability of each slot machine but finds the slot machine that gives the most coins by repeated play. In order to find a good slot machine that pays more coins, the player must search by playing various slot machines, i.e., by playing slot machines other than the one that currently has the best probability of paying out more coins. On the other hand, if we search more than necessary, we will not be able to increase the number of coins, so if we can estimate a good slot machine, we must increase the number of coins by playing that slot machine.

As shown in Sect. 2 our goal is to maximize the cumulative FSR by having each device autonomously select the appropriate SF using only the local information available to it (received ACKs), and this learning can be addressed in the MAB framework: an IoT device (player) has S SFs (slot machines), and the objective is to maximize the cumulative FSR (cumulative rewards).

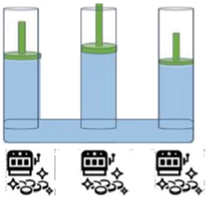


Fig. 2. ToW-dynamics.

3.2 ToW-Dynamics Based SF Selection

In this paper, we implement an approach for solving two MAB problems with SF selection using ToW-dynamics [6–8]. ToW-dynamics has been analytically validated for its high efficiency, despite its simplicity and low computational complexity, in making a series of decisions to maximize the probabilistic reward obtained, even in dynamic environments where reward probabilities frequently change [6–8]. The essential element of ToW-dynamics is a volume-conserving physical object, assuming, for example, that each slot machine is assigned to multiple cylinders with branches containing incompressible fluid, as shown in Fig. 2. The volume is updated by pushing and pulling the corresponding cylinder depending on whether or not the slot machine is rewarded for a trial at time t. Since the cylinders are connected, as in Fig. 2, the volume increase in one part immediately compensates for volume decreases in others. In other words, ToW-dynamics can update reward estimates for slot machines that are not actually played, which is the reason for its high performance. In ToW-dynamics, the arm k^* with the height cylinder interface value X is selected, and X_k is expressed by the following formula:

$$X_k(t) = Q_k(t-1) - \frac{1}{U - 1_{k' \neq k}} \sum Qk'(t-1) + osc_k(t) \tag{3}$$

$$osc_k(t) = Acos\left(\frac{2\pi(t+k-1)}{U}\right) \tag{4}$$

The osc_k is a fluctuation term, and the inclusion of this fluctuation allows the algorithm to perform properly for exploration and exploitation. U is the number of arms, which in SF selection is the number of available SFs $S.Q_k$ is the reward estimate of the kth arm and is derived by the following equation:

$$Q_k(t) = \begin{cases} \alpha Q_k(t-1) + \Delta Q_k(t), & \text{if } k = k^* \\ \alpha Q_k(t-1) & \text{otherwise.} \end{cases} \tag{5}$$

where α ($0 < \alpha \leq 1$) is the decay parameter for reward estimation. ΔQ_k is given by the following Eq. (6).

$$\Delta Q_k(t) = \begin{cases} +1 & \text{if receiving ACK} \\ -\omega_j & \text{if not receiving ACK} \end{cases} \tag{6}$$

In other words, if the transmission is successful and ACK is received, the Q value of the selected arm's (parameter's) gains " +1" as a reward, increasing the height of the fluid interface. Conversely, when the transmission fails and ACK is not received, the corresponding arm (parameter) is updated with the punishment − ω, decreasing the interface value of the arm and increasing the interface value of the other arm. Here − ω is expressed:

$$\omega(t) = \frac{p_{1st}(t) + p_{2nd}(t)}{2 - p_{1st}(t) - p\quad(t)} \tag{7}$$

where p_{1st} and p_{2nd} are the highest and second highest reward probabilities in the arm at time t, respectively. The reward probability p is given by the following equation:

$$p(t) = \frac{R_k(t)}{N_k(t)} \tag{8}$$

where N_k is the number of times arm k was selected by time t, and R_k is the number of successful transmissions. N_k and R_k are derived by:

$$N_k(t) = \begin{cases} 1 + \beta N_k(t-1) & \text{if } k = k \\ \beta N_k(t-1) & \text{otherwise} \end{cases} \tag{9}$$

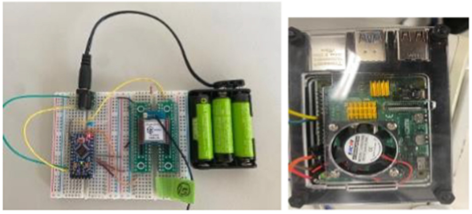
$$R_k(t) = \begin{cases} 1 + \beta R_k(t-1) \text{ if } k = k^* & \text{and receiving ACK} \\ \beta R_k(t-1) & \text{otherwise} \end{cases} \tag{10}$$

where β (0 < β ≤ 1) is the forgetting rate for learning and success experience.

In this method, each device does not need to keep information about packets it has sent in the past and only needs to keep the values of Q_k, N_k and R_k, hence memory consumption is only $S \times \{\textit{size of bits}\,(Q_k, + N_k + R_k)\}$ is all that is needed. Also, since each device only needs to decide which SF to use in each iteration, the computational complexity is $O(1)$.

4 Implementation and Performance Evaluation

The proposed method is implemented in a wireless module that supports LoRaWAN communication in the 920 MHz band. A Raspberry Pi was used for the GW and a battery-powered Arduino mini pro was used for the LoRa device. And each device is equipped with an ES 920LR module that supports LoRaWAN communication (Fig. 3). The implemented LoRaWAN system is used to evaluate the performance of the ToW-dynamics-based transmit parameter selection approach compared to i) UCB1-based, ii) ε-greedy based, iii) random-based transmit parameter selection. The experiments were conducted indoors and LoRa devices were placed in several rooms (Fig. 4). The LoRa device in position 1 was installed in the same room as the GW, guaranteeing a LOS path, while the devices in positions 2 and 3 were installed in another room at a distance from the GW, resulting in an NLOS path, and the Received Signal Strength Indication (RSSI) from each location is shown in Table 2. Therefore, devices at those distances from the GW should select a higher SF due to the lower received strength at the GW. To evaluate the impact of parameter selection in the network, we considered two scenarios, one with only SF selection and the other with SF-CH selection, using the proposed approach.



(a) (b)

Fig. 3. (a) LoRa Device. (b) GateWay

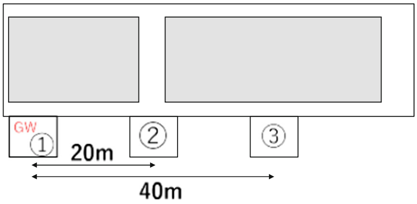


Fig. 4. Experimental Field

Table 2. RSSI from each location.

Location	RSSI [dBm]
1	−57
2	−106
3	−116

In this scenario, the performance of SF selection is evaluated when all LoRa devices use a single channel. In this experiment, the transmission interval was set as 30 s. The forgetting rate α and β in ToW-dynamics were set to 0.9, and the parameter ε, which determines the search rate in ε-greedy, was set to 0.1. Other parameters used in the experiments are summarized in Table 3.

Table 3. Parameters for performance analysis in SF selection.

Parameter	Value
BandWidth	125 kHz
Number of devices *L*	3,9,15,30
SF	7, 8, 9
Transmission Interval *I*	20
Payload Length *M*	50 bytes
Transmission power *P*	13 dBm

Figure 5 summarizes the FSR for each selected approach when the number of LoRa devices is varied from 3 to 30 (3, 9, 15, 30). The number of devices at each location is equal, i.e., one device at each location for 3 devices and 10 devices at each location for 30 devices. From Fig. 5, it can be seen that the FSR decreases as the number of LoRa devices increases for all approaches. This is due to the increase in packet collisions as the number of LoRa devices increases. In addition, the method using the MAB algorithm has a higher FSR, indicating that the autonomous distributed reinforcement learning approach is effective. Among them, the proposed method obtains higher FSR than other methods using reinforcement learning, which makes it suitable for LoRaWAN systems. Figure 6 shows the ratio of SF selection per deployment location when there are 30 LoRa devices using ToW-dynamics approach. Packets from LoRa devices located at positions 2 and 3, which are far away from the GW, have weak reception strength at the GW, so a high SF should be selected. In other words, the proposed method does not require prior information on the distance to the GW, and the SF selection is appropriate for each LoRa device.

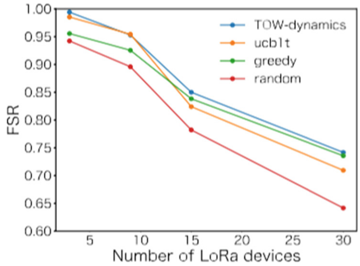


Fig. 5. FSR for SF selection

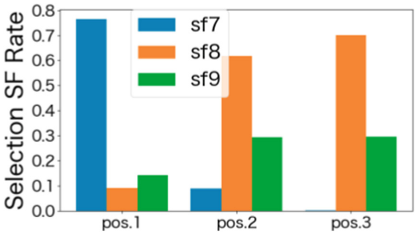


Fig. 6. SF selection at each location (TOW-dynamics)

5 Conclusion

In this paper, we propose a lightweight distributed reinforcement learning-based parameter selection method to address the LoRa network congestion problem. The approach can be implemented for LoRa devices with limited memory and computing power, and was evaluated in real-world experiments. Experimental results show that the proposed method can achieve higher FSR than other lightweight approaches. For future perspectives, there remains a need to select other parameters such as transmit power and bandwidth.

References

1. Shen, L.-H., et al.: Analysis and implementation for traffic-aware channel assignment and contention scheme in LoRa-based IoT networks. IEEE Internet Things J. **8**(14), 11368–11383 (2021)
2. Saluja, D., et al.: Scalability analysis of LoRa network for SNR-based SF allocation scheme. IEEE Trans. Ind. Inf. **17**(10), 6709–6719 (2021)
3. Lim, J.-T., et al.: Spreading factor allocation for massive connectivity in LoRa systems. IEEE Commun. Lett. **22**(4), 800–803 (2018)
4. Hamdi, R., et al.: Dynamic spreading factor assignment in LoRa wireless networks. In: ICC, pp. 1–5 (2020)
5. Ta, D.-T., et al.: LoRa-MAB: toward an intelligent resource allocation approach for LoRaWAN. In: GLOBECOM, pp. 1–6 (2019)
6. Kim, S.-J., et al.: Tug-of-war model for the two-bandit problem: nonlocally-correlated parallel exploration via resource conservation. BioSystems **101**, 29–36 (2010)

7. Kim, S.-J., et al.: Efficient decision-making by volume-conserving physical object. New J. Phys. **17**, 083023 (2015)
8. Ma, J., et al.: A reinforcement-learning-based distributed resource selection algorithm for massive IoT. Appl. Sci. **9**(18), 3730–3745 (2019)
9. Yamamoto, D., et al.: Performance evaluation of reinforcement learning based distributed channel selection algorithm in massive IoT networks. IEEE Access. **10**, 67870–67882 (2022)
10. Hasegawa, S., et al.: Multi-armed-bandit based channel selection algorithm for massive heterogeneous internet of things networks. Appl. Sci. **12**(15), 7424–7443 (2022)
11. Mahmood, A., et al.: Scalability analysis of a LoRa network under imperfect orthogonality. IEEE Trans. Ind. Inf. **15**(3), 1425–1436 (2019)
12. Li, A., et al.: A lightweight decentralized reinforcement learning based channel selection approach for high-density LoRaWAN. In: DySPAN, pp. 9–14 (2021)