Performance Evaluation of Highly Efficient Information Collection Methods by Trend Analysis of Sensor Information Using Pre-learning

Ryuji Miyamoto*, Osamu Takyu*, Hiroshi Fujiwara*, Koichi Adachi†, Mai Ohta‡, Takeo Fujii†
* Shinshu University, Nagano Japan, E-mail: takyu@shinshu-u.ac.jp
† The University of Electro-Communications, Tokyo Japan
‡ Fukuoka University, Fukuoka Japan

***Abstract*—LPWA (Low Power Wide Area), a low-power, long-distance communication method, has been attracting attention. Level Index Modulation (PLIM) has a problem of missing information due to packet collision when multiple sensors select the same index. In this study, we established an optimal design based on a mathematical model that minimizes the probability of packet collision by using a trend analysis of prior sensor information in the index design, which is the correspondence between packet transmission time, selected channel, and transmitted information. In this paper, we proceeded with the characterization by computer simulation and showed that the proposed design method can suppress the packet collision probability better than the conventional method.**

I. Introduction

In recent years, the IoT is becoming more and more popular, such as turning on the air conditioner at home or boiling the bath if you have a smartphone even when you are away from home [1]. For such remote control and remote monitoring in IoT, long-distance communication using a large number of sensors by wireless sensor network is indispensable.

LPWA (Low Power Wide Area) is a wireless communication method that consumes low power and covers a wide area, and is expected to be used in outdoor wireless sensor network development, and is attracting attention [2]. LoRa [3] and Wi-SUN [4] are widespread as LPWA standards in Japan for LPWA of specified low power wireless communication in the 920MHz band. LPWA makes it possible to widen the communication range by narrowing the transmission band and reducing the noise power. However, the problem is the decrease in throughput due to the transmission time limitation in the 920MHz band.

Packet Level Index Modulation (PLIM) [5] has been proposed as a method to increase the throughput in compliance with the LPWA standard. PLIM uses the combination of the transmission time and transmission frequency channel of the transmission packet as an index of transmission information for transmission. As a result, it becomes possible to transmit additional information even under the transmission time constraint of the 920MHz band, and the throughput can be expanded. However, as the number of sensors to be aggregated increases, the packets of each sensor select the same index, and there is a risk that information due to packet collision will be lost. So far, there is a report that collision can be suppressed by randomly assigning the mapping, which is the correspondence between the index and the transmitted information, at each sensor [6]. Therefore, it may be possible to further suppress packet collisions by improving the mapping design, but the mapping design policy has not yet been clarified.

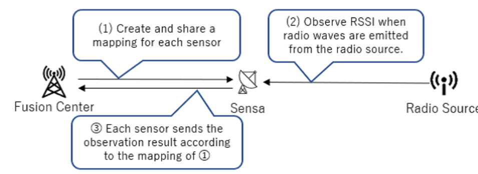


Fig. 1. System Model

In this paper, we derived the optimum mapping that minimizes the collision probability for the design of the mapping that is the correspondence between the transmission information and the index of the transmission packet in LPWA communication in PLIM. In the proposed method, a variable indicating the allocation of each index was set, and the collision probability between the two sensors was derived by a mathematical formula. With the collision probability as the objective function, the constraints that occur during packet transmission are similarly modeled by mathematical formulas. As a result, it becomes possible to apply the solver by the numerical analysis method as a combinatorial optimization problem. By applying the solver, we derived an optimized transmission mapping method that minimizes packet collisions between the two sensors. Assuming a radio wave sensor, we constructed a simulation assuming that the received electric field strength (RSSI), which is the strength of the radio wave detected by the sensor that spreads over the surface, is transmitted using PLIM when notifying the aggregation station. From the simulation results, the packet collision probability is reduced compared to the conventional design method [6].

II. System Model

A. LPWA

The system model of LPWA communication in this paper is shown in Figure1. This time, we assume a radio wave sensor that monitors the radio wave environment. However, the method of this study can be applied to other sensors such as temperature and humidity. The sensor detects the radio waves generated from the radio source and measures the received electric field strength (RSSI). The section closest to the observed RSSI is selected for the section of RSSI with a fixed width interval, and the number of that section is used as the sensor information to be transmitted. Since the sensor information is transmitted by PLIM, the transmission channel and transmission time are switched according to the sensor information.

978-1-6654-8550-0/22/$31.00 ©2022 IEEE

TABLE I
MAPPING EXAMPLE

		Index Number			
		①	②	③	④
Sensa	A	00	01	10	11
	B	00	01	10	11

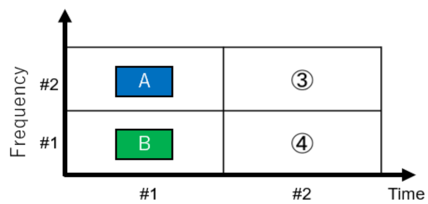


Fig. 2. PLIM model diagram

B. PLIM

In PLIM, the information to be sent for each sensor and the index number mapping table are shared between the sender and the receiver in advance. A mapping example is shown in Table I, and a PLIM model diagram is shown in Figure 2. For example, from Table I and Figure 2, the packet of sensor A is located at the index number ①, so the information "00" was sent. This can be seen on the receiving side. Similarly, since the packet of sensor B is located at the index number ②, it can be seen that the information "01" was sent. In this paper, the quantization number of the received electric field strength (RSSI) of the radio wave is converted into digital data and transmitted as the information to be sent.

III. OPTIMAL INDEX DESIGN

Conventional PLIM has the problem that as the number of sensors increases, the transmitted packets select the same index, and information is lost due to collision between packets. Therefore, in this paper, we use mathematical optimization to design to minimize collisions between packets using data learned in advance.

A. Definition of Variables

First, we define the variables used in mathematical optimization. Define $x_{i,j,k}$ as an expression (1).

$$x_{i,j,k} \in \{0,1\} \qquad \forall i \in I, \forall j \in J, \forall k \in K \tag{1}$$

i, j, k are variables representing the sensor number, RSSI number, and index number, and I, J, K represent the total number of aggregate sensors, the total number of RSSI numbers, and the total number of indexes. The RSSI number is the number assigned to each class by quantizing the RSSI value. When the variable $x_{i,j,k} = 1$, the j th RSSI number of the i th sensor is used. It means that it is assigned to the k th index, and it is not assigned when $x_{i,j,k} = 0$.

B. Objective Function

The objective function of mathematical optimization is defined as an equation (2).

$$\min \sum_{k=1}^{K} \sum_{i_1=1}^{I} \sum_{i_2=i_1+1}^{I} \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} P_{i_1,j_1} P_{i_2,j_2} x_{i_1,j_1,k} x_{i_2,j_2,k} \tag{2}$$

The above equation is an objective function that minimizes the collision probability of two packets. The two sensors are represented as i_1, i_2, and the RSSI number of each sensor is represented as j_1, j_2. $P_{i,j}$ represents the probability that the i th sensor observes the j th RSSI number, and is a random variable obtained from the result of pre-learning. Then, the sum of the collision probabilities of all two sensors was derived as the two-packet collision probability.

Here, the difference between the two-packet collision probability and the actual packet collision will be explained. As an example, Figure 3 shows a diagram showing the relationship between each sensor and the packet generation probability at a certain index when the number of sensors is 3. In the figure, let X be the whole set and the subset of X be A, B, C. A, B, C indicate the probability that the 1st, 2nd and 3rd sensors will select the index. And $\bar{A}$ indicates the case where the first sensor does not transmit at that index. From the figure, the probability P_3 of packet collision is given as the following equation.

$$\begin{aligned} P_3 = {} & P(A \cap B \cap C) + P(\bar{A} \cap B \cap C) \\ & + P(A \cap \bar{B} \cap C) + P(A \cap B \cap \bar{C}) \end{aligned} \tag{3}$$

Here, the first term on the right side shows the probability that the three sensors will transmit at the same time and the packets will collide. The other terms on the right side are the probabilities that packets from the two sensors will be transmitted at the same time and the packets will collide. Next, the two-packet collision probability is given by the following equation.

$$P_2 = P(A \cap B) + P(B \cap C) + P(A \cap C) \tag{4}$$

Here, the following relationship holds between A and B.

$$A \cap B = A \cap B \cap \bar{C} + A \cap B \cap C \tag{5}$$

Therefore,

$$\begin{aligned} P_2 &= 3P(A \cap B \cap C) + P(\bar{A} \cap B \cap C) \\ &\quad + P(A \cap \bar{B} \cap C) + P(A \cap B \cap \bar{C}) \\ &= P_3 + 2P(A \cap B \cap C) \end{aligned} \tag{6}$$

Therefore, P_2 calculates the collision probability higher than P_3 by the 3 packet simultaneous collision probability $P(A \cap B \cap C)$. In this way, the two-packet collision probability

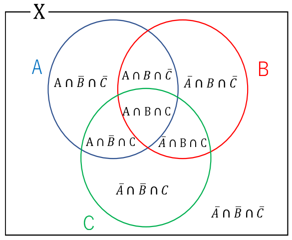


Fig. 3. Example of Allocating 3 Packets to The Same Index

is calculated with a higher collision probability than it should be. Here, this surplus corresponds to the probability that three or more sensors transmit packets at the same time in three or more sensors, but it is smaller than the two-packet collision. In addition, since the probability of two-packet collision includes the probability that three or more sensors send packets at the same time, the probability that three or more sensors send packets at the same time tends to decrease if the minimization is advanced by mathematical optimization. It is in. Furthermore, the collision probability of two packets can be expressed by a quadratic model, which is a quadratic integer programming problem of mathematical optimization, and has the advantage that the optimum solution can be obtained with a small amount of calculation compared to a high-order integer programming problem. Therefore, in this study, the two-packet collision probability was used as the objective function at the time of optimal design.

C. Constraint Expression

Formulate the constraint equations required to create the mapping for each sensor. Constraint expressions are shown in the expressions (7), (8), (9).

$$\sum_{k=1}^{K} x_{i,j,k} = 1 \quad \forall i \in I, \forall j \in J \tag{7}$$

The above equation shows that in all i and all j, the i th sensor always assigns one j th RSSI number to any index.

$$\sum_{j=1}^{J} x_{i,j,k} \leq 1 \quad \forall i \in I, \forall k \in K \tag{8}$$

The above equation shows that for all i and all k, the i th sensor assigns at most one RSSI number to the k th index. This means that the inequality $J \leq K$ holds.

$$1 \leq \sum_{j=1}^{J}\sum_{i=1}^{I} x_{i,j,k} \leq I \quad \forall k \in K \tag{9}$$

The above equation shows that the total number of sensors (number of duplicates) assigned to the k th index is 1 or more and I or less in all k.

Summarizing the defined variables, objective functions, and constraint expressions, index allocation can be formulated as a quadratic integer programming problem.

TABLE II
EXPERIMENTAL SPECIFICATIONS

Center Frequency	923.6MHz
Bandwidth	125kHz
Transmission Power	10mW
Transmission Interval	20s
Observation Point	83

$$\min \sum_{k=1}^{K}\sum_{i_1=1}^{I}\sum_{i_2=i_1+1}^{I}\sum_{j_1=1}^{J}\sum_{j_2=1}^{J} P_{i_1,j_1} P_{i_2,j_2} x_{i_1,j_1,k} x_{i_2,j_2,k}$$

$$\begin{aligned} \text{subject to} \quad & x_{i,j,k} \in \{0,1\} \quad \forall i \in I, \forall j \in J, \forall k \in K \\ & \sum_{k=1}^{K} x_{i,j,k} = 1 \quad \forall i \in I, \forall j \in J \\ & \sum_{j=1}^{J} x_{i,j,k} \leq 1 \quad \forall i \in I, \forall k \in K \\ & 1 \leq \sum_{j=1}^{J}\sum_{i=1}^{I} x_{i,j,k} \leq I \quad \forall k \in K \end{aligned}$$

The optimal solution is derived by Gurobi, an optimization solver that can calculate the above mathematical model in a quadratic integer programming problem.

IV. Numerical Result

A. Experiment Outline

In order to design the mapping by mathematical optimization, the frequency distribution of each RSSI for each sensor is required in advance. Therefore, pre-learning is performed to acquire the RSSI frequency distribution. In this paper, pre-learning was performed using LoRa in the 920MHz band, with the campus of the Faculty of Engineering, Shinshu University as the centralized area. Table II shows the experimental specifications, and Figure 4 shows the experimental environment. As shown in Figure 4, 30 sensors are placed on the campus of the Faculty of Engineering, Shinshu University, the radio wave source is moved in the area at regular time intervals, and the received electric field when the radio wave is emitted in the 920 MHz band LoRa. Record the intensity (RSSI) for each sensor.

B. Experimental Result

The cdf evaluation of the RSSI frequency distribution was performed from the results aggregated for each sensor. Table III shows the specifications of the cdf evaluation, and Figure 5 shows the results of the cdf evaluation. The horizontal axis represents RSSI and the vertical axis represents cumulative probability. RSSI is divided into 8 classes at intervals of about 13[dB]. In addition, when RSSI is less than -130[dBm], it is regarded as noise or less. It can be seen that many sensors have a high probability of observing RSSI below noise.

Fig. 4. Measurement Environment

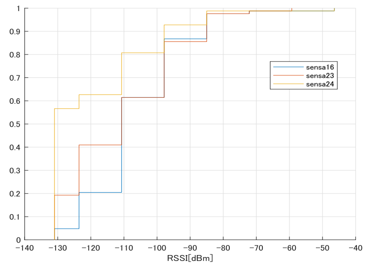
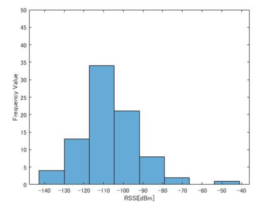


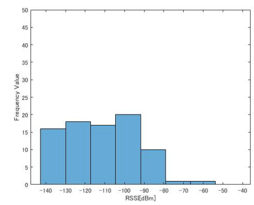
Fig. 6. CDF of 3 sensors

TABLE III
SPECIFICATIONS OF CDF EVALUATION

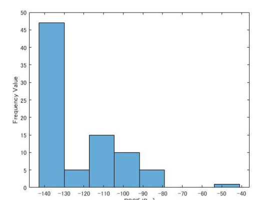
Minimum RSSI	-131dBm
Maximum RSSI	-41dBm
Number of RSSI numbers	8
Quantization Interval	12.86dB



(a) Histogram of 16th sensor



(b) Histogram of 23rd sensor



(c) Histogram of 24th sensor

Fig. 7. Histogram of 3 sensors

A sensor with a characteristic cdf result is picked up from Figure 5 and shown in Figure 6. The 24th sensor with the highest RSSI value below noise is indicated by the yellow line, the lowest 16th sensor is indicated by the blue line, and the average 23rd sensor is indicated by the orange line. The histogram of each sensor is shown in Figure 7. (a) is the histogram of sensor No. 16, (b) is the histogram of sensor No. 23, (c) is the histogram of sensor No. 24, the horizontal axis is RSSI, and the vertical axis is the frequency value. Since the number of RSSI numbers is set to 8, the maximum number of bins is 8. From Figure 6 and Figure 7, the 16th sensor tends to be close to the normal distribution, the 23rd sensor tends to be close to the uniform distribution, and the 24th sensor tends to be less than noise. It can be seen that the distribution has an extremely high ratio. From the tendency of the histogram, the probability of occurrence of a specific RSSI number tends to occur frequently in the 16th and 24th sensors, and it is possible to suppress the collision probability by improving the allocation design. be.

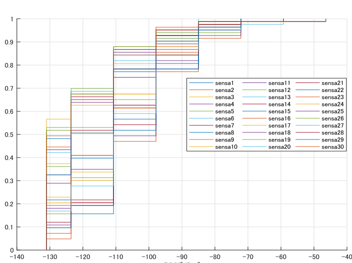


Fig. 5. CDF of 30 sensors

C. Simulation Overview

The evaluation contents of the simulation performed in this paper assume an aggregated environment as shown in Figure.4, and compare the packet collision probabilities when aggregated by the following three methods when aggregated by PLIM.

-Fixed mapping (conventional method): When the mapping of each sensor is the same (the index number corresponds to the RSSI number)

-Mapping random (conventional method): When the mapping of each sensor is randomly determined

-Optimization mapping (proposal method): When the mapping of each sensor is designed by optimization so that the probability of two-packet collision is minimized.

In this simulation, the selection probability of RSSI information obtained by the outdoor measurement experiment in the previous section is used.

D. Index Allocation Design

In order to confirm the effect of optimization, the index allocation design is confirmed by paying attention to the mode number of each sensor and the RSSI number adjacent to the mode number. For example, in the case of sensor 16 in Figure

TABLE IV
SIMULATION SPECIFICATIONS OF HEAT MAP IN OUTDOOR ENVIRONMENT

Number of Sensors (Sensor Number)	8 (3, 5, 8, 12, 17, 21, 24, 26)
Mode RSSI Number (Sensor Number)	3,3,3,1,3,3,1,1 (3, 5, 8, 12, 17, 21, 24, 26)
Number of Target RSSI Numbers	3 (Mode RSSI Number $\pm$ 1)
Number of Indexes	10

TABLE V
SIMULATION SPECIFICATIONS OF PACKET COLLISION PROBABILITY

Number of Sensors	4,5,6,7,8
Number of RSSI numbers	8
Quantization Interval	12.86dB
Number of Indexes	10
Number of Trials	100

7 (a), the mode RSSI number is 3, so RSSI numbers 2, 3 and 4 are targeted. The simulation specifications are shown in Table IV. The sensors used for mapping were fixed in order to confirm the difference in allocation due to the three different mapping design methods. By narrowing down the number of RSSI numbers to the mode number of each sensor and the adjacent RSSI number, it is confirmed that the relatively easy-to-observe RSSI number is not assigned to the same index by mathematical optimization.

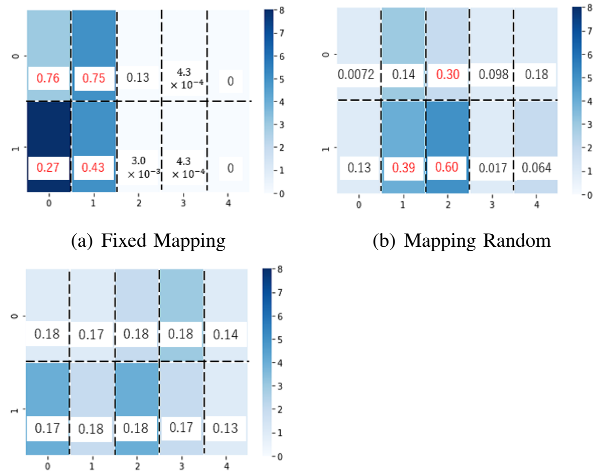
E. Packet Collision Probability

The simulation specifications are shown in Table V. Initially, it was assumed that the mapping of all sensors would be designed by mathematical optimization using 240 indexes, but it was found that the amount of calculation is enormous and it takes a considerable amount of processing time to obtain the optimum solution. Therefore, in this paper, 4-8 sensors were randomly selected from all the sensors, and the number of indexes was reduced to 10. Since the sensors are randomly selected, the number of trials was set to 100 for averaging.

V. SIMULATION RESULT

A. Results of Index Allocation Design

The simulation results are shown in Figure 8. (a), (b), and (c) show the cases where the mapping of all sensors is fixed to one, the mapping of each sensor is randomly determined, and the mapping is designed by optimization. It is assumed that the horizontal axis of the graph represents the time slot, the vertical axis represents the frequency channel, and the number of indexes is 10. The number in each index of Figure 8 is the collision rate of the packet, and the case of 0.2 or more is shown in red. The color depth of the index indicates the number of allocations. The darker the color, the larger the number of allocations, and the lighter the color, the smaller the number of allocations.Since this time we are limited to assigning the most frequent RSSI number of each sensor and the adjacent RSSI number, it is considered that collisions are more likely to occur with darker indexes, and collisions are



(a) Fixed Mapping

(b) Mapping Random

(c) Optimal Mapping

Fig. 8. Index Allocation Design Results

less likely to occur with lighter colors. In addition, since no white index is assigned, information is not given at the time of transmission, resulting in a useless index. Therefore, our ideal is to allocate a small number evenly to each index and to design a large number of allocations with a low actual packet collision probability.

Since the index number in (a) is designed to be the same as the RSSI number, the smaller the RSSI number (weak reception), the more it is assigned to the index on the left side. From Table V, it can be seen that the mode is concentrated on low values, so the allocation is also concentrated on a specific index, and the collision probability is high. In (b), it can be seen that the allocation has spread to the whole. However, the collision probability of dark indexes is relatively high. This means that the most frequently received RSSI number or a value close to it collides with each other and is missing. Therefore, the overall collision probability is lower than (a), but the important information that we originally wanted to receive remains unreceived. In case of (c), it can be seen that the number of allocations is relatively equal and the collision probability of each index is also equal. Therefore, it can be said that the effect of optimization appears because important information is less likely to be lost while suppressing the overall collision.

B. Results of Packet Collision Probability

Since the effect of optimization was confirmed from the result of the allocation design, the collision probabilities were compared by simulation. The simulation results are shown in Figure 9. The horizontal axis represents the number of sensors, the vertical axis represents the packet collision probability, the red line represents the case of fixed mapping, the blue line represents the mapping random, and the green line represents the case of optimal mapping. For any number of sensors, we succeeded in suppressing packet collisions more than the conventional method by designing the mapping by optimization.

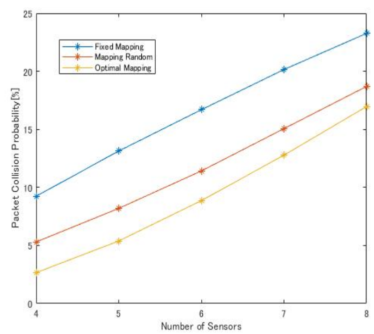


Fig. 9. Result of Packet Collision Probability

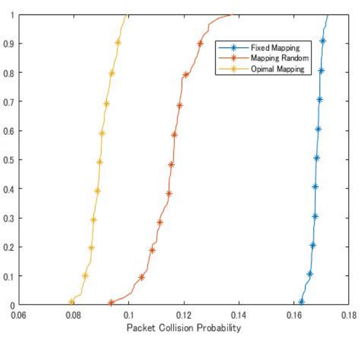


Fig. 10. Packet Collision Probability CDF with 6 Sensors

In addition, Figure 10 shows the result of the packet collision probability when the number of sensors is limited to 6 and 100 combinations of sensors are performed. The horizontal axis represents the packet collision probability, the vertical axis represents the cumulative probability, the red line represents the fixed mapping, the blue line represents the mapping random, and the green line represents the optimal mapping. Since the blue line determines the mapping at random, it can be seen that the slope is gentle and the collision probability varies compared to the others. In addition, while the collision probability can be suppressed to almost 10% or less with optimal mapping, the probability that the collision probability can be suppressed to 10% or less even with random mapping does not reach 5% in the past.

VI. Conclusion

When the number of sensors increases when transmitting by packet level index modulation (PLIM), the possibility of selecting the same index increases, and there is a problem that information is lost due to packet collision. In this paper, we proceeded with the tendency analysis of sensor information in advance and obtained the probability of occurrence of each sensor information. Then, the probability of packet collision was theoretically modeled, and the correspondence between the transmission information in PLIM and the transmission frequency and transmission timing of the packet was optimized under the minimum collision probability condition. It was clarified by the characteristic evaluation by computer simulation that the proposed design method can suppress the packet collision probability compared with the conventional method.

Acknowledgement

A part of this research project is sponsored by Ministry of Internal Affairs and Communications in Japan under the project name of Strategic Information and Communications R&D Promotion Programme (SCOPE JP 205004001) .

References

[1] L. Foschini, T. Taleb, A. Corradi and D. Bottazzi, ”M2M-based metropolitan platform for IMS-enabled road traffic management in IoT,” in IEEE Communications Magazine, vol. 49, no. 11, pp. 50-57, November 2011.
[2] Takahashi, Kakiuchi, ”LPWA (Low Power Wide Area) Standards and Technological Trends” Journal of Academic Society, Vol.100 No.9 pp.982-986 September 2017
[3] C. Shao, O. Muta, W. Wang and W. Lee, ”Toward Ubiquitous Connectivity via LoRaWAN: An Overview of Signal Collision Resolving Solutions,” in IEEE Internet of Things Magazine, vol. 4, no. 4, pp. 114-119, December 2021.
[4] T. Junjalearnvong, R. Okumura, K. Mizutani and H. Harada, ”Performance Evaluation of Multi-hop Network Configuration for Wi-SUN FAN Systems,” 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2019, pp. 1-6.
[5] K. Adachi, et. al., “Packet-Level Index Modulation for LoRaWAN” IEEE Access 2020
[6] Takeda, Taku, Adachi, Ota, Fujii, ”Implementation of Packet Level Index Modulation in 429MHz LoRa / FSK” Proc. 2022, B-17-10, March 16, 2022