Channel Capacity for a Model of Packet Level Index Modulation in LPWA Networks

Riku Yamabe, Mikihiko Nishiara, Osamu Takyu
Shinshu University, Japan
E-mail: takyu@shinshu-u.ac.jp Tel/Fax: +81-26-5255/+81-26-5220

***Abstract*—Packet level index modulation is a modulation scheme that can expand the transmission capacity without changing the existing LPWA (Low Power Wide Area) communication standard. In this paper, we calculate the theoretical capacity of the channel. In this modulation scheme, time is treated as a unit of slots and there are multiple channels that can transmit packets. To comply with the 10 $\%$ duty cycle of LPWA specification, we assume that only 1 packet is transmitted per 10 slot. In general, it is possible to transmit N packets in $10N$ slots. We calculate the capacity of the channel for this model.**

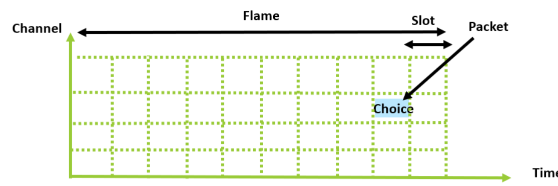

Fig. 1. Packet Level Index Modulation

I. Introduction

Low Power Wide Area (LPWA) is a low-power, wide-area wireless communication method, which is expected to be used in long range wireless sensor networks.[1]. In order to secure the wide area, the ultra high frequency (UHF) band such as 920MHz is used and the transmission band is narrowed to achieve low noise power. However, as a result of narrowing the transmission band, the throughput is limited and it is difficult to secure the bandwidth for expanding the throughput in the UHF band where there are many existing systems. In addition, as a result of securing a wide communication range, it is necessary to share the same frequency with many other systems. Therefore, in the 920MHz band, carrier sense, which is an access check before transmission, and transmission time limit are established.[2].As a result, the utilization time of the frequency band is further shortened and the throughput is reduced. In recent years, the various kinds of sensing information are simultaneously gathered by wireless sensor networks [3], and thus the additional increment of throughput is required for LPWA.

On the other hand, the deregulation of LPWA is being studied to promote the spread of LPWA [2]. Since LPWA is a narrowband transmission, the definition of many channels and the relaxation of the transmission time limit for distributed use in multiple channels are being studied. However, the throughput improvement effect is limited due to the fixed transmission time limit.

Recently, index modulation, which transmits additional information by assigning an index of the transmitted information to the frequency and time of transmission, has attracted much attention [4]. The index modulation is based on orthogonal frequency-division multiplexing (OFDM), and the method of adding the index by switching the subcarrier wave not used for transmission [5] and the method of switching the transmission time and frequency according to the information have been studied [6]. However, both of these methods require changes in the communication system, and it takes time to introduce them due to changes in the standards. Therefore, Packet Level Index Modulation (PLIM) has been proposed, which switches the frequency and time timing of transmission per packet and assigns an index of information without changing the existing wireless communication system [7]. LPWA in the 920MHz band has a transmission time limit, but the timing of transmission is selective and can be used as an index. On the other hand, since LPWA is a narrowband transmission, many channels are defined and the selected transmission channel can be used as an index. Therefore, PLIM is suitable for LPWA and is expected to increase the throughput.

In PLIM, when packet loss occurs due to packet collision, demodulation of index information also becomes difficult. Therefore, it is useful to clarify the channel capacity through the LPWA with PLIM in order to establish a compensation method for packet loss. In addition, when the index is extended to a structure consisting of multiple packets, the index can be increased without increasing the amount of packets transmitted for a certain time interval. As a result, there is a possibility that the amount of information transmitted can be increased.

In this paper, we derive the channel capacity when the index modulation is formed by multiple packets, assuming PLIM. In reality, we consider a situation where multiple sensors transmit information, but in this paper, we focus on a single sensor. The channel capacity for an arbitrary number of packets is derived by an equation, and the channel capacity is clarified. From the numerical results, we confirmed the effect of the expansion of the channel capacity for the expansion of the number of connected packets.

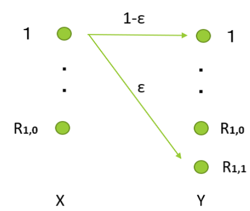

Fig. 2. Model of 1 packet 1 frame

II. About the channel model

In this paper, because we use packet level index modulation (Figure 1), the transmission of packets is handled on a slot-by-slot basis, and there are multiple channels on which packets can be transmitted. To comply with the LPWA specification, we assume that only one packet is transmitted per 10 slot. Only the 1 slot is used to transmit 1 packets, and the other slots are used for waiting time. Therefore, 10 slots are needed to transmit 1 packets and 20 slots are needed to transmit 2 packets. In general, $10N$ slots are required to transmit N packets. The slots and channels used to transmit packets can be identified by the receiver, so information can be transmitted using these slots and channels. In this paper, based on such a model, we calculate the channel capacity when N packets are transmitted in $10N$ slots.

We assume that the channel used for transmission is a vanishing channel. We assume that the channel used for transmission is a vanishing channel, i.e., transmitted packets vanish with a constant vanishing probability ε. When a packet is lost, the receiver cannot detect the contents of the packet as well as the slot and channel used for transmission. In addition, each packet does not have the information of how many packets it is among N packets. In other words, a packet does not have an identifier.

Denote the packet length as L. We call 10 slots a frame. The number of slots in a frame is denoted by T ($T = 10$). The number of channels is denoted by K. In the following sections, we derive the channel capacity by considering the input and output of the channel as random variables X and Y, respectively. The logarithms in this paper all have a base of 2.

A. Channel model for 1 packet transmission in 1 frame

To prepare for this approach, in this section we check the capacity of the channel in the case of 1 frame and 1 packet. First, we check the number of code words. Since the packet length is L, there are 2^L types of packets. In addition, we select 1 transmission slot from T slots, and then select 1 transmission channel from K channels. Therefore, there are $KT \cdot 2^L$ code words. In other words, this is the number of values that X can take. As shown in the appendix, the capacity of the channel is achieved when these code words are used with equal probability.

Next, to determine the mutual information between X and Y, we consider the reception state of the receiver. Under the condition that the transmitter sends one packet, the events observed at the receiver are

1) Send 1 packet and no disappearance
2) 1packet sent but lost

is classified as . The probability of each occurrence is

1) $1 - \varepsilon$
2) ε

The number of events observed at the receiver is If we count the number of events observed at the receiver in terms of channel, slot, and packet content, we get

1) $R_{1,0} = K^1 \times \binom{T}{1} \times (2^L)^1$Pattern
2) $R_{1,1} = K^0 \times \binom{T}{0} \times (2^L)^0$Pattern

(Fig. 2). The $R_{1,0}$ is the same as the number of code words. The conditional entropy of the received word Y under the transmission of one of the $R_{1,0}$ code words x is

$$\begin{aligned} H(Y|X=x) &= -(1-\varepsilon)\log(1-\varepsilon) - \varepsilon\log\varepsilon & (1) \\ &= h(\varepsilon) & (2) \end{aligned}$$

so it becomes

$$\begin{aligned} H(Y|X) &= \sum_x P_X(x)H(Y|X=x) & (3) \\ &= h(\varepsilon) & (4) \end{aligned}$$

where $h(\varepsilon)$ is the 2-valued entropy. However, $h(\varepsilon)$ represents the 2-valued entropy

$$h(\varepsilon) = -(1-\varepsilon)\log(1-\varepsilon) - \varepsilon\log\varepsilon \quad (5)$$

The entropy of the received state is

$$H(Y) = R_{1,0} \times \left(-\frac{1-\varepsilon}{R_{1,0}}\log\frac{1-\varepsilon}{R_{1,0}}\right) - \varepsilon\log\varepsilon \quad (6)$$

From the results in Appendix 1, the mutual information is maximized when the probability of occurrence of the transmission state is equal. Therefore, the frame-by-frame capacity $C_{1,\varepsilon}$ for this channel model is expressed by the mutual information when the transmission state has equal probability. Therefore

$$\begin{aligned} C_{1,\varepsilon} &= I(X;Y) & (7) \\ &= H(Y) - H(Y|X) & (8) \\ &= R_{1,0} \times \left(-\frac{1-\varepsilon}{R_{1,0}}\log\frac{1-\varepsilon}{R_{1,0}}\right) - \varepsilon\log\varepsilon - h(\varepsilon) & (9) \\ &= -(1-\varepsilon)(\log(1-\varepsilon) - \log R_{1,0}) - \varepsilon\log\varepsilon - h(\varepsilon) & (10) \\ &= (1-\varepsilon)\log(K \times T \times 2^L) & (11) \end{aligned}$$

This is the same as the well-known result for vanishing channels. This is the same as the well-known result for the vanishing channel. Therefore, the number of symbols that can be input to this channel is $KT \times 2^L$. In other words, this is the number of values that X can take. Now, since the channel model is a simple vanishing channel model, the channel capacity $C_{1,\varepsilon}$ is given by the well-known result

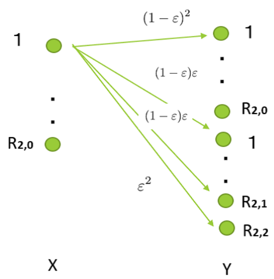

Fig. 3. Model of 2 packet 2 frame

$$C_{1,\varepsilon} = (1-\varepsilon)(L+\log KT) \tag{12}$$

The basic effect of index modulation can be ensured because the capacity of the channel is $(1-\varepsilon)L$[bit/frame] when only the packet contents are communicated without index modulation.

B. Channel model for sending 2 packets in 2 frames

In this section, we derive the capacity of the channel for 2 frames and 2 packets in preparation for generalization.

The number of code words is $K^2 \times \binom{2T}{2} \times (2^L)^2$. In this case, the capacity of the channel is achieved when the code words are used with equal probability.

When 2 packets are sent, the events observed at the receiver are

1) Send 2 packet and no disappearance
2) Sent 2 packets but lost 1 packets
3) 2 packets sent but 2 packets lost

is classified as one of the following. The probability of each of these occurrences is

1) $(1-\varepsilon)^2$
2) $(1-\varepsilon)\varepsilon$
3) ε^2

The number of events observed at the receiver is If we count the number of types of events observed at the receiver, taking into account the channel, slot, and packet content, we get

1) $R_{2,0} = K^2 \times \binom{2T}{2} \times (2^L)^2$Pattern
2) $R_{2,1} = K^1 \times \binom{2T}{1} \times (2^L)^1$Pattern
3) $R_{2,2} = K^0 \times \binom{2T}{0} \times (2^L)^0$Pattern

(Figure 3). Note that $R_{2,0}$ is the same as the number of code words. The conditional entropy of the received word Y under the transmission of one of the $R_{2,0}$ code words x is

$$\begin{aligned} H(Y|X=x) &= -(\varepsilon^2)\log(\varepsilon^2) - (1-\varepsilon)^2\log(1-\varepsilon)^2 \\ &+ 2(-(1-\varepsilon)\varepsilon)\log((1-\varepsilon)\varepsilon) && (13) \\ &= -2\varepsilon\log\varepsilon - 2(1-\varepsilon)\log(1-\varepsilon) && (14) \\ &= 2h(\varepsilon) && (15) \end{aligned}$$

so it becomes

$$\begin{aligned} H(Y|X) &= \sum_x P_X(x)H(Y|X=x) && (16) \\ &= 2h(\varepsilon) && (17) \end{aligned}$$

There are $R_{2,0}$ patterns in all. If we denote each pattern as $R_{0,i}$, $i = 1, \cdots, R_{2,0}$, then

$$\begin{aligned} P_Y(r_{0,i}) &= \sum_x P_x(x)W(r_{0,i}|x) && (18) \\ &= P_x(r_{2,i}) \times W(r_{0,i}|r_{0,i}) && (19) \\ &= \frac{(1-\varepsilon)^2}{R_{2,0}} && (20) \end{aligned}$$

If we denote each pattern of 1-packet vanishing events as $r_{1,i}$, $i = 1, \cdots, R_{2,1}$ as well

$$\begin{aligned} P_Y(r_{1,i}) &= \sum_x P_x(x)W(r_{1,i}|x) && (21) \\ &= \frac{1}{R_{2,0}}(K(2T-1)(2^L)) \times (1-\varepsilon)\varepsilon && (22) \\ &= \frac{(1-\varepsilon)\varepsilon}{K \times T \times 2^L} && (23) \end{aligned}$$

There is only one 2-packet vanishing event, which we denote by $r_{2,1}$.

$$P_Y(r_{2,i}) = \varepsilon^2 \tag{24}$$

Therefore, the The entropy of the received state is

$$\begin{aligned} H(Y) &= -\varepsilon^2\log\varepsilon^2 \\ &+ R_{2,1} \times \left(-\frac{(1-\varepsilon)\varepsilon}{K \times T \times 2^L}\right)\log\frac{(1-\varepsilon)\varepsilon}{K \times T \times 2^L} \\ &+ R_{2,0} \times \left(-\frac{(1-\varepsilon)^2}{R_{2,0}}\right)\log\frac{(1-\varepsilon)^2}{R_{2,0}} && (25) \\ &= 2h(\varepsilon) + 2(1-\varepsilon)\varepsilon\log(K \times T \times 2^L) \\ &+ (1-\varepsilon)^2\log(R_{2,0}) && (26) \end{aligned}$$

In this case too, the mutual information content is maximized when the code words are used with equal probability. From this, the channel capacity $C_{2,\varepsilon}$ per 2 frame is

$$\begin{aligned} C_{2,\varepsilon} &= I(X;Y) && (27) \\ &= H(Y) - H(Y|X) && (28) \\ &= 2(1-\varepsilon)\varepsilon\log(K \times T \times 2^L) \\ &+ (1-\varepsilon)^2\log\left(K^2 \times \binom{2T}{2} \times (2^L)^2\right) && (29) \end{aligned}$$

C. Channel model for sending N packets in N frames

The following theorem1 holds for the capacity of the channel for N frames and N packets. Checking the number of code words, it is $K^N \times \binom{NT}{N} \times (2^L)^N$.

Theorem 1:

$$\begin{aligned} C_{N,\varepsilon} &= \log\left(K^N\binom{NT}{N}(2^L)^N\right) \\ &- \sum_{j=0}^{N}\binom{N}{N-j}\varepsilon^j(1-\varepsilon)^{N-j}\log\left(K^j\binom{NT-(N-j)}{j}(2^L)^j\right) \end{aligned} \tag{30}$$

Proof: When N packets are sent, the events observed at the receiver are $i = 1, 2, 3, ...N$, then
(i) N packets are sent and i are lost.
The probability of occurrence of each is
(i) $(1-\varepsilon)^{N-i}\varepsilon^i$
If the type of event observed at the receiver is $l = 0, 1, 2, 3, ...$, taking into account the channel, slot, and packet content N, we have

(l) $R_{N,l} = K^{N-l} \times \binom{NT}{N-l} \times (2^L)^{N-l}$Pattern
As before, $R_{N,0}$ is the same as the number of code words. Find the conditional entropy of the received word Y under which one of the $R_{N,0}$ code words x has been sent.We have

$$\begin{aligned} H(Y|X=x) &= -\binom{N}{0}\varepsilon^N(1-\varepsilon)^0 \log \varepsilon^N(1-\varepsilon)^0 \\ &- \binom{N}{1}\varepsilon^{N-1}(1-\varepsilon)^1 \log \varepsilon^{N-1}(1-\varepsilon)^1 \\ &\quad .... - \binom{N}{N}\varepsilon^0(1-\varepsilon)^N \log \varepsilon^0(1-\varepsilon)^N \qquad (31) \\ &= -\sum_{k=0}^{N}\binom{N}{k}\varepsilon^{N-k}(1-\varepsilon)^k \log \varepsilon^{N-k}(1-\varepsilon)^k \qquad (32) \end{aligned}$$

Therefore

$$\begin{aligned} H(Y|X) &= \sum_x P_X(x) H(Y|X=x) \qquad (33) \\ &= -\sum_{k=0}^{N}\binom{N}{k}\varepsilon^{N-k}(1-\varepsilon)^k \log \varepsilon^{N-k}(1-\varepsilon)^k \qquad (34) \end{aligned}$$

If we denote each pattern as $R_{n,i}$, $i = 1, \ldots, R_{N,n}$, the probability of occurrence of each reception state is

$$\begin{aligned} P_Y(r_{n,i}) &= -\sum_x P_x W(r_{n,i}|x) \qquad (35) \\ &= \frac{1}{R_{N,0}} \times \left(K^n \binom{NT-(N-n)}{n} (2^L)^n\right) \\ &\times \quad \varepsilon^n(1-\varepsilon)^{N-n} \qquad (36) \end{aligned}$$

Therefore, the entropy of the received state is

$$\begin{aligned} & H(Y) \\ &= -\sum_{j=0}^{N}\sum_i P_Y(r_{j,i}) \log P_Y(r_{j,i}) \qquad (37) \\ &= -\sum_{j=0}^{N}\sum_i \frac{1}{R_{N,0}} \left(K^{N-j}\binom{NT-j}{N-j}(2^L)^{N-j}\right)\varepsilon^{N-j}(1-\varepsilon)^j \\ &\quad \times \log\left(\frac{1}{R_{N,0}}\left(K^{N-j}\binom{NT-j}{N-j}(2^L)^{N-j}\right)\varepsilon^{N-j}(1-\varepsilon)^j\right) \qquad (38) \\ &= -\sum_{j=0}^{N}\binom{N}{j}\varepsilon^{N-j}(1-\varepsilon)^j \\ &\quad \times \left(\log\frac{1}{R_{N,0}} + \log\left(K^{N-j}\binom{NT-j}{N-j}(2^L)^{N-j}\varepsilon^{N-j}(1-\varepsilon)^j\right)\right) \qquad (39) \\ &= -\log\frac{1}{R_{N,0}} - \sum_{j=0}^{N}\binom{N}{j}\varepsilon^{N-j}(1-\varepsilon)^j \\ &\quad \times \log\left(K^{N-j}\binom{NT-j}{N-j}(2^L)^{N-j} \times \varepsilon^{N-j}(1-\varepsilon)^j\right) \qquad (40) \end{aligned}$$

From the above, the channel capacity $C_{N,\varepsilon}$ in units of N frames is

$$\begin{aligned} C_{N,\varepsilon} &= I(X;Y) \qquad (41) \\ &= H(Y) - H(Y|X) \qquad (42) \\ &= \log\left(K^N\binom{NT}{N}(2^L)^N\right) \\ &- \sum_{j=0}^{N}\binom{N}{N-j}\varepsilon^j(1-\varepsilon)^{N-j} \log\left(K^j\binom{NT-(N-j)}{j}(2^L)^j\right) \qquad (43) \end{aligned}$$

■

III. Limits of channel capacity

Intuitively, a unit of many frames is expected to increase the channel capacity due to the greater freedom in choosing slots to transmit packets.

Therefore, we show the extreme values of the channel capacity.

A. Limit for $\varepsilon = 0$.

Assuming the loss probability is zero, the following theorem 2 holds in the limit of the per-frame channel capacity when the number of frames N is increased.

Theorem 2:

$$\begin{aligned} & \lim_{N\to\infty}\frac{1}{N}C_{N,0} \\ &= \log K(2^L) + T\log T - (T-1)\log(T-1) \qquad (44) \end{aligned}$$

Proof:

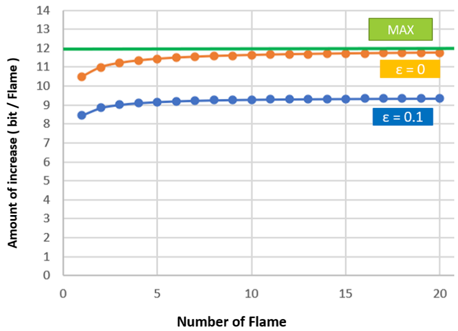

Fig. 4. Amount of increase in channel capacity (10 bits)

In general, the channel capacity per frame is

$$\frac{1}{N} \times C_{N,0} = \frac{1}{N} \times \log \binom{NT}{N} K^N (2^L)^N \tag{45}$$

$$= \log K(2^L) + \frac{1}{N} \log \frac{NT!}{N!(NT-N)!} \tag{46}$$

$$= \log K(2^L) + \frac{1}{N} \log NT! - \frac{1}{N} \log N! - \frac{1}{N} \log (NT-N)! \tag{47}$$

where the factorial can be approximated by Stirling's formula[1]

$$\frac{1}{N} \times C_{N,0} \simeq \log K(2^L) + \frac{1}{N} \log \sqrt{2\pi NT} \left(\frac{NT}{e}\right)^{NT} - \frac{1}{N} \log \sqrt{2\pi N} \left(\frac{N}{e}\right)^{N} - \frac{1}{N} \log \sqrt{2\pi (NT-N)} \left(\frac{NT-N}{e}\right)^{NT-N} \tag{48}$$

$$= \log K(2^L) + T \log T - (T-1) \log (T-1) + \frac{1}{N} (\log \sqrt{2\pi NT} - \log \sqrt{2\pi N} - \log \sqrt{2\pi (NT-N)}) \tag{49}$$

In the above equation, the term for N is

$$\lim_{N \to \infty} \frac{1}{N} (\log \sqrt{2\pi NT} - \log \sqrt{2\pi N} - \log \sqrt{2\pi (NT-N)}) = 0 \tag{50}$$

Therefore, when the loss probability is 0 and the unit is [bit/frame], the limit of the channel capacity is,

$$\lim_{N \to \infty} \frac{1}{N} C_{N,0} = \log K(2^L) + T \log T - (T-1) \log (T-1) \tag{51}$$

■

[1] $N! \simeq \sqrt{2\pi N} \left(\frac{N}{e}\right)^N$

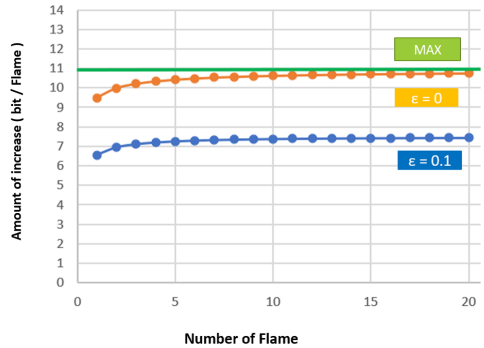

Fig. 5. Amount of increase in channel capacity (20 bits)

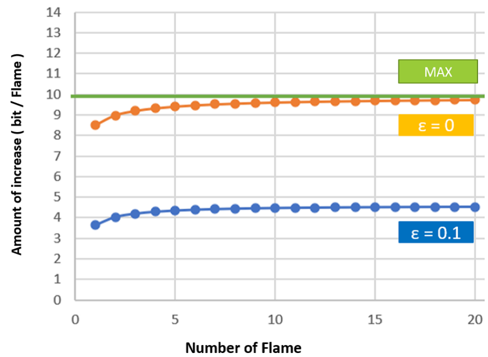

Fig. 6. Amount of increase in channel capacity (40 bits)

IV. DISCUSSION OF RESULTS WITH GRAPHS

When communicating using only the contents of a packet of length L [bit], the capacity of the channel is $(1-\varepsilon)L$[bit/frame], regardless of the number of frames, but by employing index modulation, it is possible to convey a greater amount of information than this.

In this chapter, we examine the increase of the channel capacity by index modulation. Figs. 4, 5, and 6 show the increase in the capacity per frame when the packet length is 10, 20, and 40[bit], respectively.

The number of channels is set to $K = 15$. From the above

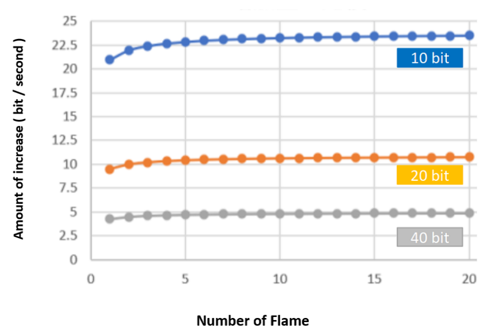

Fig. 7. Channel capacity comparison

figure, it can be seen that the capacity of the channel increases as the number of frames increases. We can see that the capacity of the channel increases as the number of frames increases, because the number of frames increases the degree of freedom in choosing slots. This graph is considered to be the result of the increase in the capacity of the channel. However, the amount of increase after 10 frames is relatively small, so it is desirable to use the number of frames up to that point.

Next, we consider the amount of increase in the capacity of the channel per time as the packet length changes. Figure 7 shows the amount of increase in the capacity of the channel when viewed in [bit/sec]. The same change is observed when the number of bits is increased, and it is found that the increase is larger for smaller packets. This can be attributed to the fact that the number of time slots per frame increased due to the shorter transmission time caused by the smaller packet length. The increase in the number of time slots increases the indexing effect, which also affects the capacity of the channel. Therefore, it is considered to be more efficient to transmit 10[bit] in 4 times than to transmit 40[bit] at a time.

V. Summary

In this study, the capacity of the channel is derived by considering the channel model of packet level index modulation. The limit of the channel capacity is also derived by setting the number of frames to ∞, since the channel capacity increases with the number of frames. In this study, we considered only the case where packets are always sent and the case where there is only one sensor, so the case where packets are not sent and the case where there are multiple sensors are not considered. In the future, we would like to model and derive the channel capacity based on this fact.

Acknowledgements

This work was supported by the Ministry of Internal Affairs and Communications (SCOPE JP205004001) and JSPS Grant-in-Aid for Scientific Research (JP18K04131).

References

[1] Takahashi、Kakiuchi、”The Trend for LPWA Specification” IEICE Journal、Vol.100 No.9 pp.982-986 Sept. 2017

[2] Ministry of Internal Affairs and Communications,Technical Requirements for the Upgrading of 920 MHz Band Low Power Radio Systems

[3] W. A. Hussein, B. M. Ali, M. F. A. Rasid and F. Hashim, “Design and performance analysis of high reliability-optimal routing protocol for mobile wireless multimedia sensor networks,” 2017 IEEE 13th Malaysia International Conference on Communications (MICC), Johor Bahru, 2017, pp. 136-140.

[4] N. Ishikawa, S. Sugiura and L. Hanzo, “50 Years of Permutation, Spatial and Index Modulation: From Classic RF to Visible Light Communications and Data Storage,” in IEEE Communications Surveys and Tutorials, vol. 20, no. 3, pp. 1905-1938, thirdquarter 2018.

[5] N. Ishikawa, S. Sugiura and L. Hanzo, “Subcarrier-Index Modulation Aided OFDM - Will It Work?,” in IEEE Access, vol. 4, pp. 2580-2593, 2016.

[6] M. I. Kadir, “Generalized Space–Time–Frequency Index Modulation,” in IEEE Communications Letters, vol. 23, no. 2, pp. 250-253, Feb. 2019.

[7] Koichi Adachi, Osamu Takyu, Mai Ohta, and Takeo Fujii, “Flexible LPWA Based on Environmental Dynamics” IEICE ICETC 2020, Dec. 2020

Appendix 1

The probability of occurrence of a transmission state when maximizing the mutual information is investigated using the Lagrange’s undecided multiplier method. In this paper, we consider the case of one packet transmission in one frame, but the same idea can be used for two packets in two frames, so that we can solve the problem for N packets in N frames as well.

Maximization by Lagrange’s undecided multiplier method

Let the objective function be

$$I(X;Y) \tag{52}$$

As, the condition

$$\sum_x P_x(x) = 1 \tag{53}$$

The above is the objective function and the Lagrange’s undecided multiplier method is used as the condition. Let R be the total pattern of transmission states.

$$\begin{aligned}
\phi &\simeq I(X;Y) + \lambda(\sum_x P_x(x) - 1) && (54)\\
&= \sum_{x,y} P_x(x) W(y|x) \log \frac{W(y|x)}{P_y(y)} + \lambda \left(\sum_x P_x(x) - 1 \right) && (55)\\
&= \sum_{x=1}^{R} P_x(x) \left((1-\varepsilon) \log \frac{1-\varepsilon}{P_x(x)(1-\varepsilon)} + \varepsilon \log \frac{\varepsilon}{\varepsilon} \right)\\
&\quad + \lambda \left(\sum_x P_x(x) - 1 \right) && (56)
\end{aligned}$$

$\frac{\partial \phi}{\partial P_x(x)} = 0$ is calculated by partial differentiation to obtain the maximum value.

Now, x は $(x = 1, 2, 3, ...R)$ we have,

$$\frac{\partial \phi}{\partial P_x(x)} = (1-\varepsilon) \log \frac{1}{P_x(x)} - P_x(x)(1-\varepsilon) \frac{1}{P_x(x)} + \lambda \tag{57}$$

$$\begin{aligned}
0 &= (1-\varepsilon) \log \frac{1}{P_x(x)} - (1-\varepsilon) + \lambda && (58)\\
\log \frac{1}{P_x(x)} &= 1 - \frac{\lambda}{1-\varepsilon} && (59)\\
P_x(x) &= e^{\frac{\lambda}{1-\varepsilon} - 1} && (60)
\end{aligned}$$

The maximum probability of occurrence of the transmission state is calculated by substituting the condition,

$$\begin{aligned}
\sum_x e^{\frac{\lambda}{1-\varepsilon} - 1} &= 1 && (61)\\
R \times e^{\frac{\lambda}{1-\varepsilon} - 1} &= 1 && (62)\\
e^{\frac{\lambda}{1-\varepsilon} - 1} &= \frac{1}{R} && (63)\\
P_x(x) &= \frac{1}{R} && (64)
\end{aligned}$$