

# Derivation and Development of Noise Channel Information Transmission Capacity

Yutong Li

North China Electric Power University, Baoding, Hebei Province, China

815707230@qq.com

**Abstract.** In this paper, we derive one of Shannon's theorems, the maximum capacity theorem of noisy channels  $C = B \log_2(1 + S / N)$ . The derivation of noisy continuous channel capacity is based on the derivation of discrete channels and most of the conclusions are a limit discretization of the corresponding theorems of discrete channels. The channel capacity is the maximum transmission rate of the information, which is the difference between the source entropy and the loss information entropy. By using both the noise and the signal independently, the channel capacity can be obtained as the entropy difference between the received signal and the white Gaussian noise. Based on this theorem, the channel capacity can be further improved in the project, such as spread spectrum communication.

**Key words:** Channel; theorem; derivation; spectrum communication; a limit discretization.

## INFORMATION UNCERTAINTY AND ENTROPY

In 1948, Shannon proposed the concept of "information entropy". For the first time, he clarified the relationship between probability and information redundancy in mathematical language to solve the problem of quantitative measurement of information: the size of a piece of information and its value Determinism has a direct relationship. The term entropy is borrowed from thermodynamics by Shannon. Thermal entropy in thermodynamics is a physical quantity that indicates the degree of disorder in a molecule's state. Shannon used the notion of information entropy to describe the uncertainty of the source. Shannon is a specific definition of information entropy:

Assuming there is a set of possible events, the probability of occurrence of these events is  $p_1, p_2, \dots, p_n$ , that these probabilities are known, but it is not known as to which event will occur, now defining a measure that measures how many species are involved in the selection of events "Choose", or how much uncertainty there is in the output.

Assuming  $H(p_1, p_2, \dots, p_n)$  first, it must have the following properties:

$H$  should be about  $p_i$  continuous;

If all  $p_i$  are equal, that is,  $p_i = 1/n$ , then  $H$  should be a monotonically increasing function of  $n$ . If the possibilities of events are equal, the more events, the more choices or the more uncertainties,

If a choice is broken down into two consecutive choices, the original  $H$  should be a weighted sum of each  $H$  value.

$H$  that can uniquely satisfy the above three assumptions has the following form:

$$H = -k \sum_{i=1}^n p_i \log p_i \quad (1)$$

Where  $k$  is a positive number, in order to measure the choice of units.

## ENTROPY OF SOURCE

Considering the discrete sources discussed above with a finite number of states, there is a set of probabilities  $p_i(j)$  for each possible state  $i$  that represents the various possible states  $j$  generated. Therefore, for each state, there exists an entropy  $H_i$ , the entropy of the source is defined as the weighted average of these  $H_i$ , and the weighted value is the probability of occurrence of the considered state:

$$H = \sum_i P_i H_i = - \sum_{i,j} P_i p_i(j) \log p_i(j) \quad (2)$$

## DISCRETE CHANNEL CAPACITY DEFINITION

If a noisy discrete channel receives content from a source, there are two statistical processes at work: source and noise. Therefore, there are multiple entropies, entropy  $H(x)$  of the source, entropy  $H(y)$  of the channel output,  $H(y) = H(x)$ , the combined entropy  $H(xy)$  of the input and output, and the conditional entropies  $H_x(y)$  and  $H_y(x)$  under no noise case, respectively, output and output when the input is known. Known input entropy, the following relationship:

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x) \quad (3)$$

In a discrete noisy channel, the received signal is not consistent with the transmitted signal due to the interference of the received noise, and there is a lost amount of information, that is, there is uncertainty in receiving the signal of the actual transmitted content. Knowing the received signal, using the conditional entropy of the information as an information metric for this loss information, according to this idea, the actual rate of subtraction of the conditional entropy from the rate of generation (ie, the entropy of the source) Transmission rate  $R$ .

$$R = H(x) - H_y(x) \quad (4)$$

In the formula,  $H_y(x)$  is called "the degree of doubt", which measures the average ambiguity of the received signal.

Due to the equations given above, the transfer rate  $R$  can be credited to two other forms:

$$R = H(x) - H_y(x) = H(y) - H_x(y) = H(x) + H(y) - H(x, y) \quad (5)$$

The first definition expression has been interpreted as the amount of information transferred minus the uncertainty of the content being sent. The second expression measures the accepted content minus the part due to noise. The third expression is two in number Minus the joint entropy, in a sense, is the number of bits shared by each of the two, so the meaning of these three expressions is intuitively consistent.

The capacity  $C$  of a noisy channel shall be the maximum achievable transmission rate, ie the rate at which the source and channel exactly match, so we define the channel capacity as:

$$C = \max[H(x) - H_y(x)] \quad (6)$$

The maximum of these is found relative to all sources that can be used as channel inputs. If the channel is noise-free,  $H_y(x) = 0$ , the maximum entropy of the channel, is its capacity.

## CONTINUOUS NOISY CHANNEL

In a continuous channel, the input signal or the transmitted signal is the continuous time signal  $f(t)$  in a particular set, the output signal or the received signal is the result of interference received by  $f(t)$ , considering only the transmitted and received signals Bandwidth does not exceed  $B$ , so in time  $T$ ,  $2BT$  numbers can be used to specify these signals and the finite dimensional distribution function to specify their statistical structure, so the statistical information of the transmitted signal is given by:

$$P(x_1, \dots, x_n) = P(x) \quad (7)$$

The noise statistics are given by the conditional probability distribution:

$$P_{x_1, \dots, x_n}(y_1, \dots, y_n) = P_x(y) \quad (8)$$

A method similar to a discrete channel may be used to define the information transmission rate of continuous channels, namely:

$$R = H(x) - H_y(x) \quad (9)$$

Where  $H(x)$  is the entropy of the input,  $H_y(x)$  is the degree of doubt, and the channel capacity  $C$  is defined as the maximum value obtained by changing the input,  $R$ , across all possible selected sets, which means that with a finite dimensional approximation, Change  $P(x_1, \dots, x_n) = P(x)$ .

An important special case occurs when adding a signal-independent noise to the signal (in a probabilistic sense),  $P_x(y)$  is simply a function of the difference  $n = (y - x)$ ,

$$P_x(y) = Q(y - x) \quad (10)$$

Specify a certain entropy (independent of the statistical information of the signal) for the noise, that is, the entropy of distribution  $Q(n)$ , denoted by  $H(n)$ , and given by  $y = x + n$ , then

$$H(x, y) = H(x, n) \quad (11)$$

Expand the left-hand side and make use of the fact that  $x$  is independent of phase  $n$ :

$$H(y) + H_y(x) = H(x) + H(n) \quad (12)$$

hence,

$$R = H(x) - H_y(x) = H(y) - H(n) \quad (13)$$

That is, the entropy of the received signal minus the entropy of the noise. Since  $H(n)$  and  $P(x)$  are independent, to obtain the maximum value,  $H(y)$  (entropy of the received signal) needs to be maximized. Therefore, the channel capacity is:

$$C = \max_{P(x)} H(y) - H(n) \quad (14)$$

### GAUSSIAN WHITE NOISE DEFINITION

Gaussian white noise refers to the signal contains all the frequency components from negative infinity to positive infinity, and each frequency component in the signal of the same weight. The white light contains the light of each frequency component, the name of white noise comes from this. It is random in magnitude at any moment but satisfies the Gaussian distribution function as a whole.

Consider a set of band-limited functions with different states. The highest frequency is  $W$  cycles / second. Let the distribution density function of amplitude  $x_1, \dots, x_n$  of  $n$  continuous sampling points be:

$$p(x_1, \dots, x_n)$$

The Department of entropy (each degree of freedom) is defined as:

$$H' = -\lim_{n \rightarrow \infty} \frac{1}{n} \int \dots \int p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) dx_1 \dots dx_n \quad (15)$$

Then define an entropy  $H$  (per second), divided by the duration of  $n$  samples  $T$  (seconds), because of  $n = 2TW$ , so  $H = 2WH'$ . When the noise is Gaussian white thermal noise, there are:

$$H' = \log \sqrt{2\pi eN} \quad (16)$$

$$H = W \log 2\pi eN \quad (17)$$

For a given average power  $N$ , white noise has the largest entropy. It can be obtained by maximizing the Gaussian distribution property.

### FORMULA DERIVATION

When the noise is white thermal noise, the average power of the transmitted signal does not exceed  $P$ , the average power of the received signal is  $P + N$ , the noise is white thermal noise, the average power of the transmitted signal does not exceed  $P$ , and the average power of the received signal is  $P + N$ ,  $N$  is the average noise power, and when the received signal also constitutes a white noise set, the entropy of the received signal is maximized because this is the maximum possible entropy at power  $P + N$ , transmitted by the proper selection of the signal, that is, it is made into a power  $P$  For the white noise set, you can get this maximum value. Entropy of the accepted set (per second) is:

$$H(y) = W \log 2\pi e(P + N) \quad (18)$$

The noise entropy is:

$$H(n) = W \log 2\pi eN \quad (19)$$

Channel capacity is:

$$C = H(y) - H(n) = W \log \frac{P + N}{N} = W \log \left(1 + \frac{P}{n_0 W}\right) \quad (20)$$

This means that with a sufficiently sophisticated coding system, it is possible to transmit binary digits at the rate of  $W \log_2((P+N)/N)$  bits per second so that the error frequency is arbitrarily small. Any coding system that transmits information at a higher rate will make the error frequency higher than a certain positive value.

In order to approach this limit transfer rate, the statistical nature of the transmitted signal must be close to white noise.

## CONCLUSION

Intuitively, the channel capacity depends on the transmission bandwidth and the signal-to-noise ratio. Compared with narrow bandwidth and low power signals, the signals with wide bandwidth and high power have greater channel capacity. The capacity of the channel, in turn, reflects the ability of the communication system to transmit information without error under certain channel conditions. That is to say, when a given signal average power and noise average power are available, information that may be transmitted in a unit time in a channel with a certain bandwidth. This is just a theoretical limit, independent of the type of modulation and other channel parameters.

The channel capacity cannot be infinitely increased. When  $B$  is constant, increasing  $S$  or decreasing  $N$  can increase the channel capacity  $C$ . In the limit case,  $S \rightarrow \infty$  and  $N \rightarrow 0$  mean that the transmission power is infinite and noise-free, is unrealizable, that is, channel capacity  $C$  cannot be infinite; and when the bandwidth tends to infinity,  $\lim_{B \rightarrow \infty} C = S/n_0 \cdot \log_2 e \approx 1.44S/n_0$ . That is to increase the signal power and bandwidth can increase the channel capacity.

Now compare the effect of these two methods, under the conditions of large signal-to-noise ratio, the formula can be approximated as  $C \approx B \log_2 S/N$ , if the signal power does not change at this time, the channel capacity and bandwidth are approximately linear and rise faster; if the bandwidth is not variable, channel capacity and signal power approximately logarithmic relationship, the rise rate is slower. Under the conditions of small signal-to-noise ratio, the channel capacity and bandwidth are still approximately linear, while the signal power is logarithmically logarithmic, but the change slope is larger and similar to the change slope of the line. Thus, under the condition of large signal-to-noise ratio, large power can be saved as long as the smaller bandwidth is added, that is, the effect of changing power with bandwidth is better; and the two methods have the same effect under the condition of small signal-to-noise ratio.

According to this principle of bandwidth and power interchange, the signal transmission bandwidth should be expanded as much as possible to improve the output signal-to-noise ratio of the system, that is, the spread spectrum communication in communication.

In general, Shannon's theorem contains the following conclusions [2]:

- 1) Channel capacity  $C$  by the "three elements" -  $B$ ,  $S$ ,  $n_0$  constraints;
- 2) Improve the signal to noise ratio  $S/N$ , can increase the channel capacity  $C$ ;
- 3) If  $n_0 \rightarrow 0$ , then  $C \rightarrow \infty$ , the channel capacity of unscrambled channel is infinite;
- 4) If  $S \rightarrow \infty$ , then  $C \rightarrow \infty$ , that when the signal power is not limited, the channel capacity is infinite;
- 5) With the appropriate  $B$  increase and  $C$  increase, but not unlimited increase, namely:  
when  $B \rightarrow \infty$ ,  $C \rightarrow 1.44S/n_0$ .
- 6)  $C$  certain time,  $B$  and  $S/N$  can be interchangeable;
- 7) In general, the information rate  $R_b \leq C$  of the source can theoretically be transmitted without error (or any small error rate).

## REFERENCES

1. C. E. SHANNON. A Mathematical Theory of Communication [J]. The Bell System Technical Journal, 1948, 27: 379-423, 623-656.
2. Cao Lina. Fan changxin. the Principle of Communication [M]. Beijing: National Defense Industry Press, 2013.