Channel capacity

Channel capacity, in electrical engineering, computer science, and information theory, is the tight upper bound on the rate at which information can be reliably transmitted over a communication channel.

Following the terms of the noisy-channel coding theorem, the channel capacity of a given channel is the highest information rate (in units of information per unit time) that can be achieved with arbitrarily small error probability. [1][2]

Information theory, developed by Claude E. Shannon in 1948, defines the notion of channel capacity and provides a mathematical model by which one can compute it. The key result states that the capacity of the channel, as defined above, is given by the maximum of the mutual information between the input and output of the channel, where the maximization is with respect to the input distribution. [3]

The notion of channel capacity has been central to the development of modern wireline and wireless communication systems, with the advent of novel error correction coding mechanisms that have resulted in achieving performance very close to the limits promised by channel capacity.

Contents

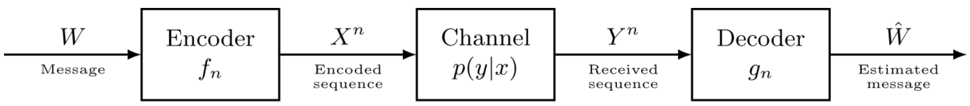
Formal definition
Additivity of channel capacity
Shannon capacity of a graph
Noisy-channel coding theorem
Example application
Channel capacity in wireless communications
- Bandlimited AWGN channel
- Frequency-selective AWGN channel
- Slow-fading channel
- Fast-fading channel

See also
- Advanced Communication Topics

External links
References

Formal definition

The basic mathematical model for a communication system is the following:



where:

- W is the message to be transmitted;
- X is the channel input symbol (X^n is a sequence of n symbols) taken in an alphabet $\mathcal{X}$;
- Y is the channel output symbol (Y^n is a sequence of n symbols) taken in an alphabet $\mathcal{Y}$;
- $\hat{W}$ is the estimate of the transmitted message;
- f_n is the encoding function for a block of length n;
- $p(y|x) = p_{Y|X}(y|x)$ is the noisy channel, which is modeled by a conditional probability distribution; and,
- g_n is the decoding function for a block of length n.

Let X and Y be modeled as random variables. Furthermore, let $p_{Y|X}(y|x)$ be the conditional probability distribution function of Y given X, which is an inherent fixed property of the communication channel. Then the choice of the marginal distribution $p_X(x)$ completely determines the joint distribution $p_{X,Y}(x,y)$ due to the identity

$$p_{X,Y}(x,y) = p_{Y|X}(y|x)\, p_X(x)$$

which, in turn, induces a mutual information $I(X;Y)$. The **channel capacity** is defined as

$$C = \sup_{p_X(x)} I(X;Y)$$

where the supremum is taken over all possible choices of $p_X(x)$.

Additivity of channel capacity

Channel capacity is additive over independent channels.[4] It means that using two independent channels in a combined manner provides the same theoretical capacity as using them independently. More formally, let p_1 and p_2 be two independent channels modelled as above; p_1 having an input alphabet $\mathcal{X}_1$ and an output alphabet $\mathcal{Y}_1$. Idem for p_2. We define the product channel $p_1 \times p_2$ as $\forall (x_1, x_2) \in (\mathcal{X}_1, \mathcal{X}_2),\ (y_1, y_2) \in (\mathcal{Y}_1, \mathcal{Y}_2),\ (p_1 \times p_2)((y_1, y_2)|(x_1, x_2)) = p_1(y_1|x_1)p_2(y_2|x_2)$

This theorem states:

$$C(p_1 \times p_2) = C(p_1) + C(p_2)$$

Proof

We first show that $C(p_1 \times p_2) \geq C(p_1) + C(p_2)$.

Let X_1 and X_2 be two independent random variables. Let Y_1 be a random variable corresponding to the output of X_1 through the channel p_1, and Y_2 for X_2 through p_2.

By definition $C(p_1 \times p_2) = \sup_{p_{X_1,X_2}} (I(X_1, X_2 : Y_1, Y_2))$.

Since X_1 and X_2 are independent, as well as p_1 and p_2, (X_1, Y_1) is independent of (X_2, Y_2). We can apply the following property of mutual information: $I(X_1, X_2 : Y_1, Y_2) = I(X_1 : Y_1) + I(X_2 : Y_2)$

For now we only need to find a distribution p_{X_1,X_2} such that $I(X_1, X_2 : Y_1, Y_2) \geq I(X_1 : Y_1) + I(X_2 : Y_2)$. In fact, π_1 and π_2, two probability distributions for X_1 and X_2 achieving $C(p_1)$ and $C(p_2)$, suffice:

$$C(p_1 \times p_2) \geq I(X_1, X_2 : Y_1, Y_2) = I(X_1 : Y_1) + I(X_2 : Y_2) = C(p_1) + C(p_2)$$

ie. $C(p_1 \times p_2) \geq C(p_1) + C(p_2)$

Now let us show that $C(p_1 \times p_2) \leq C(p_1) + C(p_2)$.

Let π_{12} be some distribution for the channel $p_1 \times p_2$ defining (X_1, X_2) and the corresponding output (Y_1, Y_2). Let $\mathcal{X}_1$ be the alphabet of X_1, $\mathcal{Y}_1$ for Y_1, and analogously $\mathcal{X}_2$ and $\mathcal{Y}_2$.

By definition of mutual information, we have

$$\begin{aligned} I(X_1, X_2 : Y_1, Y_2) &= H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2) \\ &\leq H(Y_1) + H(Y_2) - H(Y_1, Y_2|X_1, X_2) \end{aligned}$$

Let us rewrite the last term of entropy.

$$H(Y_1, Y_2|X_1, X_2) = \sum_{(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2} \mathbb{P}(X_1, X_2 = x_1, x_2) H(Y_1, Y_2|X_1, X_2 = x_1, x_2)$$

By definition of the product channel, $\mathbb{P}(Y_1, Y_2 = y_1, y_2|X_1, X_2 = x_1, x_2) = \mathbb{P}(Y_1 = y_1|X_1 = x_1)\mathbb{P}(Y_2 = y_2|X_2 = x_2)$. For a given pair (x_1, x_2), we can rewrite $H(Y_1, Y_2|X_1, X_2 = x_1, x_2)$ as:

$$\begin{aligned} H(Y_1, Y_2|X_1, X_2 = x_1, x_2) &= \sum_{(y_1,y_2)\in\mathcal{Y}_1\times\mathcal{Y}_2} \mathbb{P}(Y_1, Y_2 = y_1, y_2|X_1, X_2 = x_1, x_2) \log(\mathbb{P}(Y_1, Y_2 = y_1, y_2|X_1, X_2 = x_1, x_2)) \\ &= \sum_{(y_1,y_2)\in\mathcal{Y}_1\times\mathcal{Y}_2} \mathbb{P}(Y_1, Y_2 = y_1, y_2|X_1, X_2 = x_1, x_2)[\log(\mathbb{P}(Y_1 = y_1|X_1 = x_1)) + \log(\mathbb{P}(Y_2 = y_2|X_2 = x_2))] \\ &= H(Y_1|X_1 = x_1) + H(Y_2|X_2 = x_2) \end{aligned}$$

By summing this equality over all (x_1, x_2), we obtain $H(Y_1, Y_2|X_1, X_2) = H(Y_1|X_1) + H(Y_2|X_2)$.

We can now give an upper bound over mutual information:

$$\begin{aligned} I(X_1, X_2 : Y_1, Y_2) &\leq H(Y_1) + H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \\ &= I(X_1 : Y_1) + I(X_2 : Y_2) \end{aligned}$$

This relation is preserved at the supremum. Therefore

$$C(p_1 \times p_2) \leq C(p_1) + C(p_2)$$

Combining the two inequalities we proved, we obtain the result of the theorem:

$$C(p_1 \times p_2) = C(p_1) + C(p_2)$$

Shannon capacity of a graph

If G is an undirected graph, it can be used to define a communications channel in which the symbols are the graph vertices, and two codewords may be confused with each other if their symbols in each position are equal or adjacent. The computational complexity of finding the Shannon capacity of such a channel remains open, but it can be upper bounded by another important graph invariant, the Lovász number.[5]

Noisy-channel coding theorem

The noisy-channel coding theorem states that for any error probability $\varepsilon > 0$ and for any transmission rate R less than the channel capacity C, there is an encoding and decoding scheme transmitting data at rate R whose error probability is less than ε, for a sufficiently large block length. Also, for any rate greater than the channel capacity, the probability of error at the receiver goes to 0.5 as the block length goes to infinity.

Example application

An application of the channel capacity concept to an additive white Gaussian noise (AWGN) channel with B Hz bandwidth and signal-to-noise ratio S/N is the Shannon–Hartley theorem:

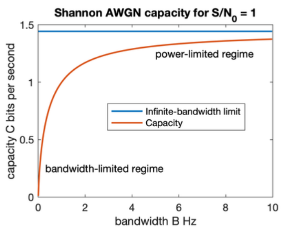
$$C = B\log_2\left(1+\frac{S}{N}\right)$$

C is measured in bits per second if the logarithm is taken in base 2, or nats per second if the natural logarithm is used, assuming B is in hertz; the signal and noise powers S and N are expressed in a linear power unit (like watts or volts^2). Since S/N figures are often cited in dB, a conversion may be needed. For example, a signal-to-noise ratio of 30 dB corresponds to a linear power ratio of $10^{30/10} = 10^3 = 1000$.

Channel capacity in wireless communications

This section[6] focuses on the single-antenna, point-to-point scenario. For channel capacity in systems with multiple antennas, see the article on MIMO.

Bandlimited AWGN channel



AWGN channel capacity with the power-limited regime and bandwidth-limited regime indicated. Here, $\frac{\bar{P}}{N_0} = 1$; B and C can be scaled proportionally for other values.

If the average received power is $\bar{P}$ [W] and the noise power spectral density is N_0 [W/Hz], the AWGN channel capacity is

$$C_{\mathrm{AWGN}} = W\log_2\left(1+\frac{\bar{P}}{N_0 W}\right) \text{ [bits/s]},$$

where $\frac{\bar{P}}{N_0 W}$ is the received signal-to-noise ratio (SNR). This result is known as the **Shannon–Hartley theorem**.[7]

When the SNR is large (SNR >> 0 dB), the capacity $C \approx W\log_2\frac{\bar{P}}{N_0 W}$ is logarithmic in power and approximately linear in bandwidth. This is called the *bandwidth-limited regime*.

When the SNR is small (SNR << 0 dB), the capacity $C \approx \frac{\bar{P}}{N_0 \ln 2}$ is linear in power but insensitive to bandwidth. This is called the *power-limited regime*.

The bandwidth-limited regime and power-limited regime are illustrated in the figure.

Frequency-selective AWGN channel

The capacity of the frequency-selective channel is given by so-called water filling power allocation,

$$C_{N_c} = \sum_{n=0}^{N_c-1}\log_2\left(1+\frac{P_n^*|\bar{h}_n|^2}{N_0}\right),$$

where $P_n^* = \max\left\{\left(\frac{1}{\lambda}-\frac{N_0}{|\bar{h}_n|^2}\right), 0\right\}$ and $|\bar{h}_n|^2$ is the gain of subchannel n, with λ chosen to meet the power constraint.

Slow-fading channel

In a slow-fading channel, where the coherence time is greater than the latency requirement, there is no definite capacity as the maximum rate of reliable communications supported by the channel, $\log_2(1+|h|^2 SNR)$, depends on the random channel gain $|h|^2$, which is unknown to the transmitter. If the transmitter encodes data at rate R [bits/s/Hz], there is a non-zero probability that the decoding error probability cannot be made arbitrarily small,

$$p_{out} = \mathbb{P}(\log(1+|h|^2 SNR) < R),$$

in which case the system is said to be in outage. With a non-zero probability that the channel is in deep fade, the capacity of the slow-fading channel in strict sense is zero. However, it is possible to determine the largest value of R such that the outage probability p_{out} is less than ϵ. This value is known as the ϵ-outage capacity.

Fast-fading channel

In a fast-fading channel, where the latency requirement is greater than the coherence time and the codeword length spans many coherence periods, one can average over many independent channel fades by coding over a large number of coherence time intervals. Thus, it is possible to achieve a reliable rate of communication of $\mathbb{E}(\log_2(1+|h|^2 SNR))$ [bits/s/Hz] and it is meaningful to speak of this value as the capacity of the fast-fading channel.

See also

- Bandwidth (computing)
- Bandwidth (signal processing)
- Bit rate
- Code rate
- Error exponent
- Nyquist rate
- Negentropy
- Redundancy
- Sender, Data compression, Receiver
- Shannon–Hartley theorem
- Spectral efficiency
- Throughput

Advanced Communication Topics

- MIMO
- Cooperative diversity

External links

- Hazewinkel, Michiel, ed. (2001) [1994], "Transmission rate of a channel" (https://www.encyclopediaofmath.org/index.php?title=p/t093890), *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer Academic Publishers, ISBN 978-1-55608-010-4
- AWGN Channel Capacity with various constraints on the channel input (interactive demonstration) (http://webdemo.inue.uni-stuttgart.de/webdemos/08_research/capacity/)

References

1. Saleem Bhatti. "Channel capacity" (https://web.archive.org/web/20070821212637/http://www.cs.ucl.ac.uk/staff/S.Bhatti/D51-notes/node31.html). *Lecture notes for M.Sc. Data Communication Networks and Distributed Systems D51 -- Basic Communications and Networks*. Archived from the original (http://www.cs.ucl.ac.uk/staff/S.Bhatti/D51-notes/node31.html) on 2007-08-21.
2. Jim Lesurf. "Signals look like noise!" (http://www.st-andrews.ac.uk/~www_pa/Scots_Guide/iandm/part8/page1.html). *Information and Measurement, 2nd ed.*
3. Thomas M. Cover, Joy A. Thomas (2006). *Elements of Information Theory* (https://books.google.com/books?id=VWq5GG6ycxMC&printsec=frontcover#v=snippet&q=%22channel%20capacity%22&f=false). John Wiley & Sons, New York.
4. Cover, Thomas M.; Thomas, Joy A. (2006). "Chapter 7: Channel Capacity". *Elements of Information Theory* (Second ed.). Wiley-Interscience. pp. 206–207. ISBN 978-0-471-24195-9.
5. Lovász, László (1979), "On the Shannon Capacity of a Graph", *IEEE Transactions on Information Theory*, IT-25 (1), doi:10.1109/tit.1979.1055985 (https://doi.org/10.1109%2Ftit.1979.1055985).
6. David Tse, Pramod Viswanath (2005), *Fundamentals of Wireless Communication* (https://books.google.com/books?id=66XBb5tZX6EC&printsec=frontcover#v=onepage&q=%22Channel%20capacity%22&f=false), Cambridge University Press, UK
7. *The Handbook of Electrical Engineering* (https://books.google.com/books?id=-WJS3VnvomIC&pg=RA1-SL4-PA41&dq=%22Shannon%E2%80%93Hartley+theorem%22&hl=en&sa=X&ei=7PVqUqaCNcPs2wXMmIGADw&ved=0CC4Q6AEwAA#v=onepage&q=%22Shannon%E2%80%93Hartley%20theorem%22&f=false). Research & Education Association. 1996. p. D-149. ISBN 9780878919819.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Channel_capacity&oldid=945585184"

This page was last edited on 14 March 2020, at 22:26 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.