Лекция №5

Информация и энтропия в телекоммуникационных технологиях

Понятие информации, информационная энтропия. Потеря информации в каналах связи. Пропускная способность каналов связи. Формула Шеннона

1. Информация и энтропия

Понятие информации является первичным понятием, поэтому ее определение будет неточным. Можно только описать ее свойства.

Информация трактует смысл полученного сообщения (сигнала), представляет его интерпретацию, результат реакции на него. Любое сообщение представляет собой совокупность сведений о некоторой физической системе. Если бы состояние физической системы было заранее известно, не было бы смысла передавать сообщение. Сообщение приобретает смысл только тогда, когда состояние системы заранее неизвестно, т.е. случайно. Степень неопределенности системы определяется числом ее возможных состояний и вероятностями появления этих состояний.

Количество информации естественно измерять величиной исчезнувшей неопределенности. Пусть источник сигнала имеет N элементов, каждой из которых может находиться в одном из M состояний. Тогда число различных сообщений

$$L = M^N. \qquad (1)$$

Однако это выражение для L не может быть выбрано за количество информации в такой форме: случаю L=1 соответствует отсутствие информации (N=0), не удовлетворяет условию аддитивности, согласно которому общее количество информации определяется суммой информаций от различных независимых источников. Поэтому за единицу количества информации I принимается мера Хартли-Шеннона:

$$I = log_b L = log_b M^N = N log_b M = -N log_b P, \qquad (2)$$

где P=1/M – вероятность реализации одинаковых M состояний.

Для различных значений основания логарифма b имеем различные единицы измерения информации: бит – b=2, байт-b=8, нат – b=e, дит - b=10.

Если вероятности каждого состояния различные, то нужно пользоваться средним значением информации $\langle I \rangle = S$:

$$S = -\frac{N}{N} \sum_i P_i log_b P_i = -\sum_i P_i log_b P_i. \qquad (3)$$

Величина S называется информационной энтропией. Если принять b=e=2.72; $P_i = 1/M$ из (3) мы получим формулу, совпадающую по виду с физической энтропией Больцмана:

$$S = k ln M, \; k = 1.38 \cdot 10^{-16} \text{эрг/град}, \qquad (4)$$

где M – термодинамическая вероятность. Физическая и информационная энтропии связаны друг с другом. Приобретение одного бита информации имеет цену устраненной энтропии:

$$1 \text{ бит} = k \ln 2 = 0.97 \cdot 10^{-23} \text{ Дж/град}. \quad (5)$$

Количественное равенство информации и энтропии не означает тождественность этих понятий. Если физическая и термодинамическая энтропии для замкнутых систем со временем растут, то информационная энтропия со временем может не увеличиваться, а уменьшаться для любых систем.

2. Условная энтропия статистически зависимых сообщений

В задачах телекоммуникаций необходимо учесть наличие нескольких источников, дающих зависимые сообщения. Рассмотрим два зависимых источника сигналов: $X = \{x_1 \dots, x_i \dots, x_M\}$, $Y = \{y_1 \dots, y_i \dots, y_N\}$, с вероятностями элементов $P(x_i)$, $P(y_j)$.

Частная условная энтропия $S(Y/x_k)$ источника Y при условии, что источник находится в состоянии x_k:

$$S(Y/x_k) = -\sum_{j=1}^{N} P(y_j/x_k) \ln P(y_j/x_k). \quad (6)$$

Полная условная энтропия источника Y:

$$S(Y/X) = -\sum_{k=1}^{M} P(x_k) S(Y/x_k) =$$

$$= -\sum_{k=1}^{M} \sum_{j=1}^{N} P(x_k) P(y_j/x_k) \ln P(y_j/x_k) = \sum_{k=1}^{M} \sum_{j=1}^{N} P(x_k, y_j) \ln P(y_j/x_k). \quad (7)$$

Если X и Y жестко связаны, то S(Y/X)=0. Это следует из $P(y_j/x_k) = 0$, либо $P(y_j/x_k) = 1$.

Можно убедиться, что для объединения зависимых сообщений справедливы равенства

$$S(Y,X) = S(X) + S(Y/X), \quad (8)$$

$$S(Y,X) = S(X,Y). \quad (9)$$

3. Потеря информации в каналах связи.

Необходимо определить изменение энтропии множества сигналов X если они недоступны для наблюдений, а наблюдаемым является Y, связанные с X. Такие ситуации могут иметь место, например, при воздействии помех на X.

Информация, переданная от X к Y, равна разности между начальной неопределенностью S(X) конечной неопределенностью S(X/Y):

$$I(X \to Y) = S(X) - S(Y/X) = \sum_{i=1}^{N} \sum_{j=1}^{M} P(x_i, y_j) ln \frac{P(x_i)}{P(x_i/y_j)}. \qquad (10)$$

Если источник, приемник и канал связи характеризуется непрерывными функциями плотности вероятности ρ(X), ρ(X/Y), то формула (10) имеет вид

$$I(X \to Y) = \int \int \rho(X, Y) ln \frac{\rho(X)}{\rho(X/Y)} dXdY . \qquad (11)$$

<u>Темы самостоятельных работ</u>

1. Общие свойства социальной, кибернетической, биологической информаций.
2. Доказательство свойств условных энтропий.
3. Энтропия иерархических уровней дерева с заданными вероятностями ветвей.

Литература ()