1.2 Информационная мера Шеннона

КОЛИЧЕСТВО ИНФОРМАЦИИ И ИЗБЫТОЧНОСТЬ

Дискретные системы связи - системы, в которых как реализации сообщения, так и реализации сигнала представляют собой последовательности символов алфавита, содержащего конечное число элементарных символов.

Пусть ξ и η - случайные величины с множествами возможных значений $X = \{x_1, x_2, \ldots, x_n\}$, $Y = \{y_1, y_2, \ldots, y_m\}$.

Количество информации $H(\xi)$ при наблюдении случайной величины $\xi \in X = \{x_1, x_2, \ldots, x_n\}$ с распределением вероятностей $P = \{p_1, p_2, \ldots, p_n\}$ задается формулой Шеннона:

$$H(\xi) = \sum_{i=1}^{n} p_i \log 1/p_i.$$

Единицей измерения количества информации является бит, который представляет собой количество информации, получаемое при наблюдении случайной величины, имеющей два равновероятных значения.

При равномерном распределении $p_1 = p_2 = \ldots = p_N$ количество информации задается формулой Хартли:

$$H(\xi) = \log_2 N.$$

Справедливы следующие соотношения:

1) $0 \le H(\xi) \le \log_2 N$;
2) $N = 2, p_1 = p_2 = 0{,}5,\ H(\xi) = 1$;
3) $H(\xi, \eta) = H(\xi) + H(\eta)$, если ξ и η - независимы.

Избыточностью называется $p = 1 - H(\xi)/\max H(\xi) = 1 - H(\xi)/\log_2 N$.

ЭНТРОПИЯ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Непрерывные системы передачи информации - системы, в которых как реализации сообщения, так и реализации сигнала на конечном временном интервале $(0, T)$ представляют собой некоторые непрерывные функции времени.

Пусть $x(t)$ - реализации непрерывного сообщения на входе какого-либо блока схемы связи, $y(t)$ - реализация выходного сообщения (сигнала), $W(x)$ - плотность вероятностей ансамбля входных сообщений, $W(y)$ - плотность вероятностей ансамбля выходных сообщений

Формулы для энтропии H непрерывных сообщений получаются путем обобщения формул для энтропии дискретных сообщений. Если Δx - интервал квантования (точность измерения), то при достаточно малом Δx ***энтропия непрерывных сообщений***

$$H(X) = -\int_{-\infty}^{\infty} W(x) \log W(x) dx - \log \Delta x \int_{-\infty}^{\infty} W(x) dx = H^*(X) - \log \Delta x =$$

$$= -\int_{-\infty}^{\infty} W(x) \log[W(x)\Delta x] dx,$$

где $H^*(X) = -\int_{-\infty}^{\infty} W(x) \log W(x) dx$ - приведенная энтропия.

$H(X)$ можно представить в компактной записи:
$H(X) = M[-\log(W(x)\Delta x)]$ и $H^*(X) = M[-log W(x)]$.

По аналогии $H(Y) = -\int_{-\infty}^{\infty} W(y) \log W(y) dy$.

Пример 1.

Источник сообщений выдает символы из алфавита $A = \{a_i\}$, $i = \overline{1,4}$ с вероятностями $p_1 = 0{,}2$; $p_2 = 0{,}3$; $p_3 = 0{,}4$; $p_4 = 0{,}1$. Найти количество информации и избыточность.

Решение.

По формуле Шеннона $H(\xi) = \sum_{i=1}^{n} p_i \log 1/p_i$.

$H(A) = -(0{,}2\cdot\log 0{,}2+0{,}3\cdot\log 0{,}3+0{,}4\cdot\log 0{,}4+0{,}1\cdot\log 0{,}1)=1{,}86$ бит

Избыточность $p = 1 - \frac{H(A)}{log_2 4} =1 - \frac{1{,}86}{2log_2 2} =0{,}07$

Пример 2.

Имеются два источника информации, алфавиты и распределения вероятностей которых заданы матрицами:

$$\begin{Vmatrix} X \\ P \end{Vmatrix} = \begin{Vmatrix} x_1 & x_2 \\ p_1 & p_2 \end{Vmatrix}, \quad \begin{Vmatrix} Y \\ Q \end{Vmatrix} = \begin{Vmatrix} y_1 & y_2 & y_3 \\ q_1 & q_2 & q_3 \end{Vmatrix}.$$

Определить, какой источник дает большее количество информации, если

1) $p_1 = p_2$; $q_1 = q_2 = q_3$;

2) $p_1 = q_1$; $\mathrm{p}_2 = q_2 + q_3$;

Решение.

2) Воспользуемся формулой Шеннона:

$$H(\xi) = \sum_{i=1}^{n} p_i \log 1/p_i.$$

с учетом условия задачи имеем $p_1 = q_1; p_2 = q_2 + q_3$;

$$H(\xi) = q_1 \log_2 1/q_1 + q_2 \log_2 1/(q_2 + q_3) + q_3 \log_2 1/(q_2 + q_3).$$

С другой стороны,

$$H(\eta) = q_1 \log_2 1/q_1 + q_2 \log_2 1/q_2 + q_3 \log_2 1/q_3.$$

Сравним слагаемые $H(\xi)$ и $H(\eta)$.

Поскольку

$$\frac{1}{q_2 + q_3} < \frac{1}{q_2}, \quad \frac{1}{q_2 + q_3} < \frac{1}{q_3}, \text{ то } H(\xi) < H(\eta).$$

Пример 4.

По линии связи передаются непрерывные амплитудно-модулированные сигналы $x(t)$, распределенные по нормальному закону с математическим ожиданием $m_x = 0$ и дисперсией $\sigma_x^2 = \sigma^2 = 8B^2$.

Определить энтропию $H(X)$ сигнала при точности его измерения $\Delta x = 0{,}2B.$

Решение.

По условию плотность вероятностей сигнала $x(t)$

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2},$$

$$H(X) = -\int_{-\infty}^{\infty} W(x) \log W(x) dx - \log \Delta x = -\frac{1}{\ln 2}\int_{-\infty}^{\infty} W(x) \ln W(x) dx - \log \Delta x =$$

$$-\frac{1}{\ln 2}\int_{-\infty}^{\infty} W(x)\left[\ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{x^2}{2\sigma^2}\right]dx - \log \Delta x = -\frac{1}{\ln 2}\left(\ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}\right) - \log \Delta x =$$

$$= \frac{1}{\ln 2}\ln\sqrt{2\pi e\sigma^2} - \log \Delta x = \log\frac{\sqrt{2\pi e\sigma^2}}{\Delta x}.$$

Подставляя числовые значения, получаем

$$H(X) = \log\frac{\sqrt{2\pi e 8}}{0{,}2} \cong 5{,}87 \text{ дв. ед.}$$

Пример 5.

Определить полную энтропию системы X, состояние которой имеет экспоненциальное распределение.

Решение.

Полная энтропия системы X

$$H(X) = H^*(X) - log_2\Delta x,$$

$$H^*(X) = M[-logW(x)] = M[-log\alpha \cdot e^{-\alpha x}] = \mathrm{M}[-log_2\alpha - log_2 e^{-\alpha x}] =$$

$$= -log_2\alpha + M[\alpha x \cdot log_2 e] = -log_2\alpha + \alpha \cdot log_2 e \cdot M[x] =$$

$$= -log\alpha + \alpha \cdot log_2 e \cdot \frac{1}{\alpha} = log_2\frac{e}{\alpha},$$

$$H(X) = H^*(X) - log_2\Delta x = log_2\frac{e}{\alpha} - log_2\Delta x = log_2\frac{e}{\alpha\Delta x}.$$

1.3 Условная энтропия и взаимная информация

ДИСКРЕТНЫЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Условной энтропией величины Y при наблюдении величины X называется $H(Y/X) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i y_j)\log_2\frac{1}{p(y_j/x_i)}$.

Справедливы соотношения:

$$H(X,Y) = H(X) + H(Y), \quad H(X,Y) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) log_2\frac{1}{p(x_i y_j)},$$

$$I(X,Y) = H(X) - H(X/Y),$$

$$I(X,Y) = I(Y,X) = H(Y) - H(Y/X)$$

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

Если X и Y независимы, то $I(X,Y) = 0$,

Если X и Y полностью зависимы (содержат одну и ту же информацию), то $I(X,Y) = H(X) = H(Y)$.

Справедливы следующие соотношения

$$I(X,Y) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) log_2\frac{p(x_i, y_j)}{p(x_i)p(y_j)},$$

$$0 \le I(X,Y) \le H(X), \quad 0 \le I(X,Y) \le H(Y).$$

Понятие взаимной информации широко используется в теор передачи информации. Требования к взаимной информации различны зависимости от того, с какой информацией работает потребите. Например, если X и Y - это сообщения, публикуемые различными газетам то для получения возможно большей суммарной (совместнс информации, взаимная (то есть одинаковая в данном случае) информац должна быть минимальной. Если же X и Y- это сообщения на входе и выходе канала связи с помехами, то для получения возможно больш информации ее получателем необходимо, чтобы взаимная информац была наибольшей. Тогда условная энтропия $H(X/Y)$ – это поте информации в канале связи (ненадежность канала). Условная энтроп $H(Y/X)$ – это информация о помехах (энтропия источника помех $H(n$ поступающая в канал извне или создаваемая внутренними помехами канале (рис 1.4).

При расчетах условной энтропии и взаимной информации удоб пользоваться следующими соотношениями теории вероятностей:

1) теорема умножения вероятностей
$p(x_i y_j) = p(x_i)p(y_j/x_i) = p(y_j)p(x_i/y_j)$;

2) формула полной вероятности $p(x_i) = \sum_{j=1}^{M} p(x_i, y_j)$;

$p(y_j) = \sum_{i=1}^{N} p(x_i, y_j)$;

3) формула Байеса $p(x_i/y_j) = \frac{p(x_i)p(y_j/x_i)}{p(y_j)} = \frac{p(x_i)p(y_j/x_i)}{\sum_{i=1}^{N} p(x_i)p(y_j/x_i)}$.

Пусть $x(t)$ - реализации непрерывного сообщения на входе какого-либо блока схемы связи, $y(t)$- реализация выходного сообщения (сигнала), $f_1(x)$- одномерная плотность вероятностей ансамбля входных сообщений, $f_2(y)$ - одномерная плотность вероятностей ансамбля выходных сообщений, $f(x,y)$ - совместная плотность вероятностей, $f(y/x)$- условная плотность вероятностей y при известном x. Тогда для количества информации I справедливы следующие соотношения:

$$I(x,y) = log_2 \frac{f(x,y)}{f_1(x) f_2(y)} = log_2 \frac{f(x/y)}{f_1(x)} = log_2 \frac{f(y/x)}{f_2(y)} = I(x,y),$$

$$I(X,y) = \int_X I(x,y)\, f(x/y)\, dx\,, \quad I(x,Y) = \int_Y I(x,y) f(y/x)\, dy,$$

Выражение для ***полной взаимной информации***, содержащейся в двух непрерывных системах X и Y будет определяться

$$I(X,Y) = \int_Y f_2(y) I(X,y)\, dy = \iint_{X\,Y} f(x,y)\, I(x,y)\, dx\, dy =$$

$$= \iint_{X\,Y} f(x,y) log_2 \frac{f(x,y)}{f_1(x) f_2(y)} dx\, dy = I(Y,X),$$

или, применяя знак математического ожидания

$$I_{Y \leftrightarrow X} = M\left[log_2 \frac{f(X,Y)}{f_1(X) f_2(Y)}\right]$$

$$I(X,y) \geq 0, \;\; I(x,Y) \geq 0, \qquad I(X,Y) \geq 0.$$

Здесь $I(x,y)$ - взаимная информация между каким-либо значением x входного и значением y выходного сообщений, $I(X,y)$ и $I(x,Y)$- средние значения условной информации, $I(X,Y)$- полная средняя взаимная информация.

Условная энтропия определяется по формуле:

$$H(X/Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) log\, f(x/y)\, dx\, dy,$$

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) log\, f(y/x)\, dx\, dy,$$

Можно представить в компактной записи

$$H(X/Y) = M[-log f(Y/X)] - log\Delta x$$

Когда X и Y статистически связаны между собой, то

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y).$$

При независимых X и Y

$$H(X,Y) = H(X) + H(Y).$$

Полная средняя взаимная информация определяется формулой:

$$I(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X).$$

Пример 1.

Дана матрица

$$P(X,Y)=\begin{Vmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & 0 & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{Vmatrix},\ x\in X,\ y\in Y$$

Определить: $H(X), H(Y), H(X/Y), H(Y/X), H(X,Y), I(X,Y)$.

Решение.

По формуле полной вероятности имеем:

$$P(x_1)=\sum_{j=1}^{3} p(x_1,y_j)=\frac{3}{8},\quad P(x_2)=\frac{1}{4},\ P(x_3)=\frac{3}{8},$$

$$P(y_1)=\sum_{i=1}^{3} p(x_i,y_1)=\frac{3}{8},\quad P(y_2)=\frac{1}{4},\ P(y_3)=\frac{3}{8}.$$

Следовательно,

$$H(X)=\sum_{i=1}^{3} p(x_i)\log_2\frac{1}{p(x_i)}=1{,}57;$$

$$H(Y)=\sum_{j=1}^{3} p(y_j)\log_2\frac{1}{p(y_j)}=1{,}57.$$

По теореме умножения $p(x_i/y_i)=\frac{p(x_i,y_i)}{p(y_i)}$

$$p(x_1/y_1)=\frac{1}{3},\ p(x_1/y_2)=\frac{1}{2},\ p(x_1/y_3)=\frac{1}{3},$$

$$p(x_2/y_1)=\frac{1}{3},\ p(x_2/y_2)=0,\ p(x_2/y_3)=\frac{1}{3},$$

$$p(x_3/y_1)=\frac{1}{3},\ p(x_3/y_2)=\frac{1}{2},\ p(x_3/y_3)=\frac{1}{2}.$$

Следовательно,

$$H(X/Y)=\sum_{j=1}^{3}\sum_{i=1}^{3} p(x_i y_j)\log_2\frac{1}{p(x_i/y_j)}=1{,}43.$$

Аналогично $H(Y/X)=1{,}43;$

Энтропия объединения: $H(X,Y)=H(X)+H(Y/X)=1{,}57+1{,}43=3;$

Взаимная информация величин X и Y:

$I(X,Y)=H(X)-H(X/Y)=0{,}14.$

Пример 2.

Канал связи описан следующей канальной матрицей

$$P(Y/X)=\begin{pmatrix} 0{,}98 & 0.1 & 0{,}2 \\ 0.01 & 0{,}75 & 0{,}3 \\ 0.01 & 0{,}15 & 0{,}5 \end{pmatrix}$$

Найти:

1) Среднее количество информации, которое переносится одним символом сообщения, если вероятности появления символов источника сообщений равны $p(x_1)=0{,}7,\ p(x_2)=0{,}2,\ p(x_3)=0{,}1$.
2) Чему равны информационные потери при передаче сообщения из 1000 символов алфавита x_1, x_2, x_3?
3) Чему равно количество принятой информации?

Решение.

1) Энтропия источника сообщений

$H(X)=-\sum_{i=1}^{n} p_i log\, p_i=$ −(0,7·log0,7+0,2·log0,2+0,1·log0,1)=1,16 бит

2) Общая условная энтропия

$$H(Y/X)=\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)log_2\frac{1}{p(y_j/x_i)}=$$

$$=\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i)\ p(y_j/x_i)log_2 p(y_j/x_i)$$

$H(Y/X)=-0{,}7[(0{,}98\cdot log_2 0.98+0{,}01\cdot log_2 0.01+0{,}01\cdot log_2 0.01)+$
$+0{,}2(0{,}1\cdot log_2 0.1+0{,}75\cdot log_2 0.75+0{,}15\cdot log_2 0.15)+0{,}1(0{,}2\cdot log_2 0.2+$
$0{,}3\cdot log_2 0.3+0{,}5\cdot log_2 0.5)]=0{,}473$ бит

Пример 3.

Найти энтропию шума $H(Y/X)$ в двоично-симметричном канале без памяти, если энтропия источника на входе канала $H(X) = 3400$ бит, энтропия ансамбля на выходе канала $H(Y) = 6800$ бит, ненадежность канала $H(X/Y) = 700$ бит.

Решение.

Взаимная информация источника сообщений

$$I(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X).$$

$$I(X,Y) = H(X) - H(X/Y) = 3400 - 700 = 2700 \text{ бит}.$$

Энтропия шума $H(Y/X) = H(Y) - I(X,Y)$=6800-2700=4100 бит.

Пример 4.

Принимаемый сигнал может иметь амплитуду A_1(событие x_1) или A_2(событие x_2), а также сдвиг фаз φ_1 (событие y_1) или φ_2 (событие y_2). Вероятности совместных событий имеют следующие значения: $p(x_1,y_1) = 0{,}73$, $p(x_1,y_2) = 0{,}21$, $p(x_2,y_1) = 0{,}02$, $p(x_2,y_2) = 0{,}04$. Вычислить количество информации, получаемой о фазовом сдвиге сигнала, если станет известной его амплитуда.

Решение.

Среднее количество информации о фазовом сдвиге при известной амплитуде $I(X,Y) = H(Y) - H(Y/X)$,

$$p(y_1) = p(x_1,y_1) + p(x_2,y_1) = 0{,}73 + 0{,}02 = 0{,}75$$

$$p(y_2) = p(x_1,y_2) + p(x_2,y_2) = 0{,}21 + 0{,}04 = 0{,}25,$$

$$H(Y) = -\sum_{j=1}^{M} p(y_i) log_2 p(y_j) = -(0{,}75\, log_2 0{,}75 + 0{,}25 log_2 0{,}25) =$$

$$= 0{,}81 \text{ бит}$$

$$p(x_1) = p(x_1,y_1) + p(x_1,y_2) = 0{,}73 + 0{,}21 = 0{,}94$$

$$p(x_2) = p(x_2,y_1) + p(x_2,y_2) = 0{,}02 + 0{,}04 = 0{,}06,$$

По теореме умножения $p(y_j/x_i) = \frac{p(x_i,y_j)}{p(x_i)}$

$$p(y_1/x_1) = \frac{0{,}73}{0{,}94} = 0{,}78, \quad p(y_2/x_1) = \frac{0{,}2}{0{,}94} = 0{,}22,$$

$$p(y_1/x_2) = \frac{0{,}02}{0{,}06} = 0{,}33, \quad p(y_2/x_2) = \frac{0{,}04}{0{,}06} = 0{,}67,$$

$$H(Y/X) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j) log_2 \frac{1}{p(y_j/x_i)} =$$

$$= 0{,}73 log_2 \frac{1}{0{,}78} + 0{,}21 log_2 \frac{1}{0{,}22} + 0{,}02 log_2 \frac{1}{0{,}33} + 0{,}04 log_2 \frac{1}{0{,}67}$$

$$= 0{,}77 \text{бит}$$

$$I(Y,X) = H(Y) - H(Y/X) = 0{,}81 - 0{,}77 = 0{,}04 \text{ бит}.$$