Передача информации

Александра Игоревна Кононова

НИУ МИЭТ

25 ноября 2019 г.

Информационный канал

Информационный канал — совокупность устройств, объединённых линиями связи, предназначенных для передачи информации от источника информации (начального устройства канала) до её приёмника (конечного устройства канала).

- достоверность передачи информации;
- надёжность работы устройств;
- скорость передачи информации;
- задержка сигнала во времени.

Матмодель канала

1. множество X допустимых сообщений на входе;
2. множество Y допустимых сообщений на выходе;
3. набор условных вероятностей $p(y|x)$ получения сигнала y на выходе при x на входе (статистические свойства шумов (помех)).

Канал без шумов: $X = Y$, $p(y|x) = \begin{cases} 1, & \text{при } y = x \\ 0, & \text{при } y \neq x \end{cases}$

- В **дискретных** каналах сигналы на входе и выходе — слова из одного или двух (вход и выход) **алфавитов.**
- В **непрерывных** каналах входной и выходной сигналы — функции от непрерывного параметра — времени.
- В смешанных или гибридных каналах рассматривают их дискретные и непрерывные компоненты раздельно.

Ёмкость канала

Способность канала передавать информацию характеризуется **пропускной способностью** или **ёмкостью** канала (C) — максимальным количеством информации, передаваемым в единицу времени.

$$C = \lim_{T\to\infty} \frac{\max\limits_X \left(I(X,Y)\right)}{T}$$

Для канала без шума:

$$C = \lim_{T\to\infty} \frac{\max\limits_X \left(I(X)\right)}{T} = \lim_{T\to\infty} \frac{\log_2 N(T)}{T},$$

где $N(T)$ — число всех возможных сигналов (сообщений) за время T.

Единицы измерения

Ёмкость — **бод:**

- «1 бод равен одному изменению информационного параметра в секунду. <...> скорость в бодах целиком определяется величиной такта» (В. Г. Олифер, Н. А. Олифер Компьютерные сети. Принципы, технологии, протоколы);
- «1 бод равно 0.8 бит/сек» (справочный портал `calc.ru`);
- «При пользовании нормальным компьютером бод эквивалентен количеству битов (bits) в секунду» (Словарь бизнес-терминов).

Информационная скорость — **бит в секунду** (Олифер, Олифер):

- Если сигнал $\nu > 2$ различимых состояний и нет избыточности, то каждое состояние несёт $\log_2 \nu$ бит, и 1 бод $= \log_2 \nu$ бит/с.
- Если сигнал имеет 2 состояния и для надёжности бит кодируется последовательностью из η символов, то 1 бит/с $= \eta$ бод.

Первая теорема Шеннона (для канала без помех)

1. При любой производительности источника сообщений, меньшей пропускной способности канала:

$$\frac{H(X)}{T} < C$$

существует способ кодирования, позволяющий передавать по каналу все сообщения, вырабатываемые источником.

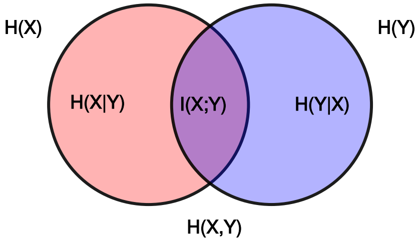
2. Не существует способа кодирования, обеспечивающего передачу сообщения без их неограниченного накопления, если

$$\frac{H(X)}{T} > C$$

Относительная информация

По каналу связи передаются сообщения из X. Из-за помех приёмником воспринимается Y.

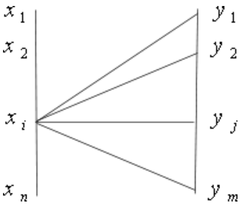
X, Y — источники сообщений:



$H(X) = I(X)$ — энтропия X
$H(Y) = I(Y)$ — энтропия Y
$I(X,Y) = I(Y,X)$ — относительная информация X и Y
$H(X,Y) = H(Y,X)$ — энтропия объединения X и Y

$H(Y|X)$ — условная энтропия Y относительно X (шум)
$H(X|Y)$ — условная энтропия X относительно Y (инф. потери)

Условная энтропия как инф. потери



$H(x_i|y_j)$ — неопределённость того, что, отправив x_i, получим y_j;
$H(y_j|x_i)$ — неуверенность, которая остаётся после получения y_j в том, что было отправлено именно x_i.

$$H(X|Y) = M[-\log_2 p(X|Y)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(x_i|y_j) \leqslant H(X)$$

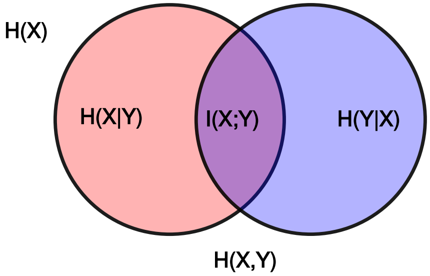
$$H(Y|X) = M[-\log_2 p(Y|X)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(y_j|x_i) \leqslant H(Y)$$

Если помехи отсутствуют ($|Y| = |X|$, посланному x_i соответствует принятый y_i), то $H(X) = H(Y)$, условная энтропия $H(Y|X) = 0$.
Если Y не зависит от X, то $H(Y|X) = H(Y)$.

Свойства меры информации и энтропии

1. $I(X,Y) \geqslant 0$,
   $I(X,Y) = 0 \Leftrightarrow X$ и Y независимы;
2. $I(X,Y) = I(Y,X)$;
3. $H(X) = 0 \Leftrightarrow X$ — константа;
4. $I(X,Y) = H(X) + H(Y) - H(X,Y)$;
5. $I(X,Y) \leqslant I(X,X) = I(X) = H(X)$.
   Если $I(X,Y) = I(X,X)$, то X — функция от Y (разные y при разных x, передача без потерь).

Взаимные энтропия и информация



К свойству 4:

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

$$I(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(x_i) \cdot p(y_j)}$$

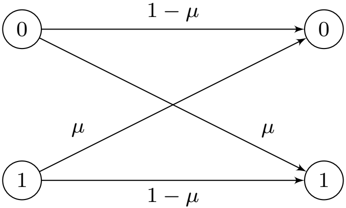
$$H(X|Y) = M[-\log_2 p(X|Y)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(x_i|y_j)$$

$$H(X,Y) = M[-\log_2 p(X,Y)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(x_i, y_j)$$

Двоичный симметричный канал

От X к Y передаются символы 0 и 1 (k символов в единицу времени).

Каждый символ, независимо от других, с вероятностью μ может быть искажён (т. е. заменён противоположным).



Пусть X производит $x_1 = 0$ и $x_2 = 1$ с вероятностями p и $1-p$, на выходе Y символы $y_1 = 0$ и $y_2 = 1$ с вероятностями r и $1-r$.

$$H(Y|X) = p \cdot H(Y|x_1) + (1-p) \cdot H(Y|x_2)$$

$$H(Y|x_1) = -\sum_{i=1}^{2} P(y_i|x_1)\log_2 P(y_i|x_1) = -(1-\mu)\log_2(1-\mu) - \mu\log_2\mu$$

$$H(Y|x_2) = -\sum_{i=1}^{2} P(y_i|x_2)\log_2 P(y_i|x_2) = -\mu\log_2\mu - (1-\mu)\log_2(1-\mu) = H(Y|x_1)$$

$$H(Y|X) = p \cdot H(Y|x_1) + (1-p) \cdot H(Y|x_1) = -\mu \cdot \log_2\mu - (1-\mu) \cdot \log_2(1-\mu)$$

Передаваемая информация

$$H(Y) = -r \cdot \log_2 r - (1-r) \cdot \log_2 (1-r)$$

Передаваемая информация на символ

$$I(X,Y) = H(Y) - H(Y|X) =$$
$$= \Big(-r \cdot \log_2 r - (1-r) \cdot \log_2 (1-r) \Big) - \Big(-\mu \cdot \log_2 \mu - (1-\mu) \cdot \log_2 (1-\mu) \Big)$$

Обозначим $\eta(x) = -x \cdot \log_2 x$: $I(X,Y) = (\eta(r) + \eta(1-r)) - (\eta(\mu) + \eta(1-\mu))$

Макс. передаваемая информация на символ

$$\max_X (I(X,Y)) = \max_X \Big((\eta(r) + \eta(1-r)) - (\eta(\mu) + \eta(1-\mu)) \Big) =$$
$$= \max_r \Big(\eta(r) + \eta(1-r) \Big) - \Big(\eta(\mu) + \eta(1-\mu) \Big) \quad = \quad 1 - \Big(\eta(\mu) + \eta(1-\mu) \Big)$$

Пропускная способность:

$$C = k \cdot \max_X (I(X,Y)) = k \cdot \Big(1 - (\eta(\mu) + \eta(1-\mu)) \Big)$$

Вторая теорема Шеннона (для канала с помехами)

1. При любой производительности источника сообщений, меньшей пропускной способности канала:

$$\frac{H(X)}{T} < C$$

существует способ кодирования, позволяющий обеспечить передачу всей информации со **сколь угодно малой вероятностью ошибки.**

2. Не существует способа кодирования, обеспечивающего передачу информации со сколь угодно малой вероятностью ошибки, если

$$\frac{H(X)}{T} > C$$

Простейшие помехозащитные коды

1. Обнаруживающий одиночную ошибку (здесь и далее — инверсию) в одном бите — двойное повторение каждого бита.
2. Обнаруживающий одиночную ошибку в блоке из N бит — **контроль чётности** (добавление к каждому блоку $N+1$-го бита так, чтобы дополнить количество единиц до заранее выбранного для кода чётного (even) или нечетного (odd) значения).

   Двойная ошибка в блоке не будет обнаружена.
3. Исправляющий одиночную ошибку в одном бите — тройное повторение каждого бита.
4. Исправляющий одиночную ошибку в блоке из N бит — код Хэмминга.

   Двойная ошибка в блоке будет принята за одиночную не в том месте.
5. Исправляющий одиночную ошибку и обнаруживающий двойную в блоке из N бит — **код Хэмминга с дополнительным битом чётности.**

Спасибо за внимание!

НИУ МИЭТ
http://miet.ru/

Александра Игоревна Кононова
illinc@mail.ru