Chương 4

Kênh truyền

Nguyễn Thanh Bình

Khoa CNTT&TT - Đại học Cần Thơ

02 - 2008

Nội dung

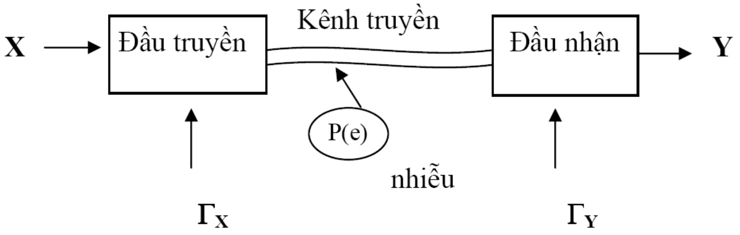
Khái niệm kênh truyền rời rạc không nhớ

Rời rạc

Truyền rời rạc từng ký tự và nhận rời rạc từng ký tự.

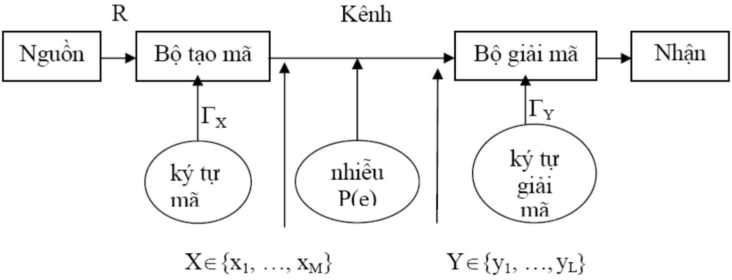
Không nhớ

Ký tự nhận sau không phụ thuộc vào ký tự nhận trước.



Gợi ý

Thử so sánh với mô hình cơ bản của hệ thống liên lạc truyền tin.



Gợi ý

Thử so sánh với mô hình vật lý.

Các ký hiệu quy ước

- $\Gamma_X = \{x_1, \ldots, x_M\}$: bộ ký tự sinh mã ở đầu truyền (input).
- $\Gamma_Y = y_1, \ldots, y_L\}$: bộ ký tự giải mã ở đầu nhận (output).
- X: biến ngẫu nhiên lấy giá trị đã mã hóa trên Γ_X và có phân phối $p(X = x_i) = p(x_i)$ với $i = 1, \ldots, M$.
- Y: biến ngẫu nhiên lấy giá trị nhận được trước khi giải mã trên Γ_Y và có phân phối xác suất có điều kiện
  $p(Y = y_j|X = x_i) = p(y_j|x_i) = p_{ij}$ với $j = 1, \ldots, L$.
- $A = \|p_{ij}\|$: ma trận truyền tin của kênh truyền rời rạc không nhớ.

Phân phối ở đầu nhận

- A_j là cột thứ j của ma trận A.
- $P'_X = [p(x_1) \;\; p(x_2) \;\; \dots \;\; p(x_M)]$

$$\begin{aligned} p(y_j) &= \sum_{i=1}^{M} p(x_i)p(y_j|x_j) = \sum_{i=1}^{M} p(x_i)p_{ij} \\ &= P'_X.A_j \end{aligned}$$

Tổng quát

$$P'_Y = P'_X.A$$

Bài toán

Xác suất truyền: $p(x_1) = 0.5$ và $p(x_2) = p(x_3) = 0.25$

$$A = \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{cccc} y_1 & y_2 & y_3 & \\ \left[\begin{array}{ccc} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{array}\right. & & \left.\right] & \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \end{array}$$

Phân phối xác suất ở đầu nhận

- Ta có: $P'_X = [0.5 \;\; 0.25 \;\; 0.25]$
- Phân phối của Y:

$$P'_Y = P'_X.A = [0.375 \;\; 0.3 \;\; 0.325]$$

Ví dụ

Lượng tin trên kênh truyền

$$\begin{aligned} H(Y) &= H(0.375, 0.3, 0.325) = 1.58(bit) \\ H(Y|X) &= H(0.5, 0.2, 0.3) = 1.49(bit) \\ I(X|Y) &= I(Y|X) = H(Y) - H(Y|X) = 0.09(bit) \end{aligned}$$

Dung lượng kênh truyền

Nhận xét

- Ta có: $I(X|Y) = I(Y|X) = H(Y) - H(Y|X)$
- $H(Y) = H(P'_X.A)$: phụ thuộc P_X.
- $H(Y|X) = \sum_i p(x_i)H(Y|X = x_i) = P'_X.H(Y|X = x)$: phụ thuộc P_X.

  $\Rightarrow$ $I(X|Y)$ phụ thuộc vào P_X.

  $\Rightarrow$ tồn tại một P_X xác định để $I(X|Y)$ đạt maximum.

Định nghĩa

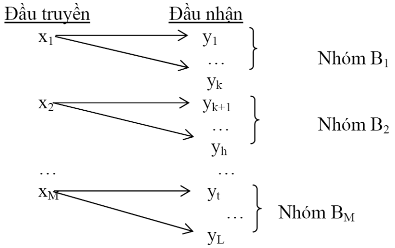
Dung lượng kênh truyền (*đơn vị tính: bit*)

$$C = Max_{\forall p(X)} I(X|Y)$$

Mô hình

- Phân Y thành M nhóm B_i ứng với các giá trị x_i ở đầu truyền.
- Xác suất truyền x_i với điều kiện đã nhận y_j là:

$$p(X = x_i | Y = y_j \in B_i) = 1$$



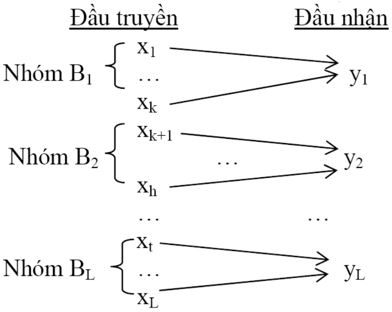
Đặc trưng

$$H(X|Y) = 0 \Leftrightarrow C = \log_2 M$$

Mô hình

- Phân X thành L nhóm B_j ứng với các giá trị y_j ở đầu nhận.
- Xác suất để nhận y_j với điều kiện đã truyền x_i là:

$$p(Y = y_j | X = x_i \in B_j) = 1$$



Đặc trưng

$$H(Y|X) = 0 \Leftrightarrow C = \log_2 L$$

Kênh truyền không nhiễu

Mô hình

- Là sự kết hợp của mô hình kênh truyền xác định và mô hình kênh truyền không mất thông tin.

$$
\begin{array}{ccc}
\underline{\text{Đầu truyền}} & & \underline{\text{Đầu nhận}} \\
x_1 & \longrightarrow & x_1 \\
x_2 & \longrightarrow & x_2 \\
\ldots & & \ldots \\
x_M & \longrightarrow & x_M
\end{array}
$$

Đặc trưng

$$H(X|Y) = H(Y|X) = 0 \Leftrightarrow C = \log_2 L = \log_2 M$$

Kênh truyền không sử dụng được

Mô hình

- Là kênh truyền mà khi truyền giá trị nào thì mất giá trị đó, hoặc kênh truyền mà xác xuất nhiễu trên kênh lớn hơn xác xuất nhận được.

Ví dụ

$$A = \begin{bmatrix} \varepsilon & 1-\varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

Đặc trưng

$$H(X|Y) = H(Y|X) = Max \Leftrightarrow C = 0$$

Mô hình

Là kênh truyền mà ma trận truyền tin A có đặc điểm:

- Mỗi dòng của A là một hoán vị của phân phối $P = p'_1, \ldots, p'_L\}$.
- Mỗi cột của A là một hoán vị của phân phối $Q = q'_1, \ldots, q'_M\}$.

Ví dụ

$$A = \begin{array}{cc} \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/2 \\ 1/6 & 1/2 & 1/3 \end{bmatrix} & \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \\ \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} & \end{array}$$

Xây dựng công thức

- Ta có:

$$C = Max\, I(X|Y) = Max\,(H(Y) - H(Y|X))$$

- Do kênh truyền đối xứng nên:

$$H(Y|X) = -\sum_{j=1}^{L} p'_j \log p'_j$$

- Do đó:

$$C = Max I(X|Y) = Max\, H(Y) + \sum_{j=1}^{L} p'_j \log p'_j$$

$$\Rightarrow C = \log L + \sum_{j=1}^{L} p'_j \log p'_j$$

- Dấu **=** xảy ra khi và chỉ khi:

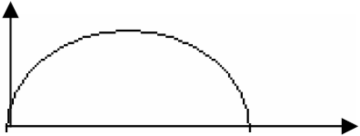
$$p(Y = y_j) = \frac{1}{L} \Leftrightarrow p(X = x_i) = \frac{1}{M}$$

Kênh truyền 1 bit với nhiễu β

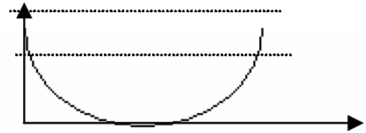
- Ma trận truyền tin:

$$A = \begin{bmatrix} 1-\beta & \beta \\ \beta & 1-\beta \end{bmatrix}$$

- Dung lượng: $C = \log_2 L - H(\beta, 1-\beta) = 1 - H(\beta, 1-\beta)$.



$H(\beta, 1-\beta)$



$1 - H(\beta, 1-\beta)$

Định lý

$$C = Max\ I(X|Y) = I(X^*|Y)$$

khi tín hiệu vào $X = X^*$ thỏa phân phối

$$P(X^* = x_k) = 2^{-C} d_k \text{ với } d_k > 0$$

Đặt vấn đề bài toán giải mã

Phân tích yêu cầu

- Khi truyền giá trị x_i, ta nhận được giá trị y_j nào đó.
- Đối với kênh truyền không nhiễu: $y_j \equiv x_i$.
- Đối với kênh truyền có nhiễu: y_j có thể khác $x_i \Rightarrow$ cần tìm cách để giải mã y_j về x_i.

Phân hoạch các giá trị ở đầu nhận

Phân chia chia tập Y thành các tập con B_i sao cho:

1. $\begin{cases} B_i \cap B_j = \emptyset & \text{với } \forall i \neq j \\ B_1 \cup \cdots \cup B_M = Y \end{cases}$

2. Khi nhận $y_j \in B_i$ thì giải mã về x_i.

Ví dụ bài toán giải mã

- Cho tập các từ mã truyền X và nhận Y như sau:

$$X = \{0000, 0101, 1110, 1011\}$$
$$Y = \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}$$

- Giả sử tập Y được phân hoạch như sau:

$$B_1 = \{0000, 1000, 0001, 0010\}$$
$$B_2 = \{0101, 1101, 0100, 0111\}$$
$$B_3 = \{1110, 0110, 1111, 1100\}$$
$$B_4 = \{1011, 0011, 1010, 1001\}$$

- Giả sử nhận được $y_j = 0011$. Kiểm tra thấy $y_j \in B_4 \Rightarrow$ giải mã y_j về $x_4 = 1011$.

Khái niệm

- **Từ mã**: là dãy n ký tự truyền hay dãy n ký tự nhận đúng.
- **Bộ mã** (S, n): là tập hợp gồm S từ mã, mỗi từ mã dài n ký tự, ký hiệu là: $x^{(1)}, \ldots, x^{(S)}$.
- **Lược đồ giải mã**: là một hàm gán cho dãy n ký tự nhận được. Ký hiệu: $g(y_j) = w_i$.
- **Lược đồ giải mã tối ưu**: là lược đồ giải mã sao cho tổng xác suất truyền sai là nhỏ nhất (tổng xác suất truyền đúng là lớn nhất).

Ví dụ minh họa

- Kênh truyền từng bit ($C = 1$), nguồn phát thông báo với tốc độ $R = 2/5$ bit/giây ($R < C$).
- Xét từng khoảng thời gian 5 giây, ta có:
  - Số bit được phát ra: $nR = 2$
  - Tập hợp các tín hiệu khác nhau: $2^{nR} = 4$, ký hiệu: m_1, m_2, m_3, m_4.
- Có 2 cách mã hóa tín hiệu:

	Cách 1	**Cách 2**
m_1	00000	00
m_2	01101	01
m_3	11010	10
m_4	10111	11

Các khái niệm cơ bản

Phân tích

Cách 1

- Trong 5 bit mã hóa có thể hiểu là có 2 bit thông tin cần truyền và 3 bit bổ sung giúp phát hiện nhiễu (theo một phương pháp nào đó, sẽ được đề cập sau).
- Với cách mã hóa này, ta có nhiều khả năng phát hiện và sửa sai do nhiễu.

Cách 2

- Chỉ có 2 bit thông tin, không có bit bổ sung.
- Nếu có nhiễu (có 1 bit truyền sai) $\Rightarrow$ trùng lặp sang một trong các tín hiệu khác $\Rightarrow$ không thể phát hiện được có nhiễu hay không.

Phương pháp

Bước 0: Khởi tạo các $B_i = \emptyset$ với $\forall i = \overline{1, M}$.

Bước lặp $j = \overline{1, L}$: xét $\forall y_j \in Y$

- $\forall i = \overline{1, M}$ tính các giá trị $p(w_i).p(y_j|w_i)$
- Chọn giá trị $w^*{}_i$ sao cho $p(w^*{}_i).p(y_j|w^*{}_i) = \max$.
- Đặt $B_i = B_i + \{y_j\}$ và $g(y_j) = w^*{}_i$

Ví dụ minh họa

$$A = \begin{array}{cc} \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/2 \\ 1/6 & 1/2 & 1/3 \end{bmatrix} & \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \\ \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} & \end{array}$$

$$p(x_1) = 1/2 \qquad p(x_2) = p(x_3) = 1/4$$

Bước 0: $B_1 = B_2 = B_3 = \emptyset$

Bước 1: Xét giá trị y_1

- $p(x_1).p(y_1|x_1) = 1/2 \times 1/2 = 1/4$ (max)
- $p(x_2).p(y_1|x_2) = 1/4 \times 1/3 = 1/12$
- $p(x_3).p(y_1|x_3) = 1/4 \times 1/6 = 1/24$

$\Rightarrow$ liệt kê y_1 vào tập $B_1 \Rightarrow B_1 = \{y_1\}$.

Bước 2: Xét giá trị y_2

- $p(x_1).p(y_2|x_1) = 1/2 \times 1/3 = 1/6$ (max)
- $p(x_2).p(y_2|x_2) = 1/4 \times 1/6 = 1/24$
- $p(x_3).p(y_2|x_3) = 1/4 \times 1/2 = 1/8$

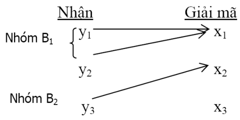
$\Rightarrow$ liệt kê y_2 vào tập $B_1 \Rightarrow B_1 = \{y_1, y_2\}$.

Bước 3: Xét giá trị y_3

- $p(x_1).p(y_3|x_1) = 1/2 \times 1/6 = 1/12$
- $p(x_2).p(y_3|x_2) = 1/4 \times 1/2 = 1/8$ (max)
- $p(x_3).p(y_3|x_3) = 1/4 \times 1/3 = 1/12$

$\Rightarrow$ liệt kê y_3 vào tập $B_2 \Rightarrow B_2 = \{y_3\}$.

Kết quả: $B_1 = \{y_1, y_2\}$ $B_2 = \{y_3\}$ $B_3 = \emptyset$



Các dạng sai số cơ bản

Định nghĩa

- Xác xuất truyền sai từ mã x_i

$$p(e|x_i) = \sum p(Y = y_j \notin B_i | X = x_i)$$

- Xác suất truyền sai trung bình:

$$p(e) = \sum_{i=1}^{M} p(X = x_i) p(e|x_i)$$

- Xác suất truyền sai lớn nhất:

$$p_m(e) = \max_{i=\overline{1,M}} p(e|x_i)$$

Ví dụ

Xét ví dụ ở phần xây dựng lược đồ giải mã tối ưu.

- Xác xuất truyền sai một từ mã:

  $p(e|x_1) = \sum p(Y = y_j \notin B_1 | X = x_1) = p(y_3|x_1) = 1/6$

  $p(e|x_2) = \sum p(Y = y_j \notin B_2 | X = x_2) = p(y_1|x_2) + p(y_2|x_2) = 1/3 + 1/6 = 1/2$

  $p(e|x_3) = \sum p(Y = y_j \notin B_3 | X = x_3) = 1$
- Xác suất truyền sai trung bình:

  $p(e) = \sum_{i=1}^{M} p(X = x_i) p(e|x_i) = 1/2 \times 1/6 + 1/4 \times 1/2 + 1/4 \times 1 = 11/24 < 1/2$
- Xác suất truyền sai lớn nhất:

  $p_m(e) = \max_{i=\overline{1,M}} p(e|x_i) = \max_{i=\overline{1,M}} (p(e|x_1), p(e|x_2), p(e|x_3)) = 1$