C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1: Mã khối tuyến tính)

Lý thuyết thông tin

Biên soạn: Phạm Văn Sự

Bộ môn Xử lý tín hiệu và Truyền thông
Khoa Kỹ thuật Điện tử I
Học viện Công nghệ Bưu chính Viễn thông

ver. 22a

Notes

Mục tiêu của bài học

- Trang bị một số khái niệm cơ bản về mã hóa kênh
- Mã khối tuyến tính
- Mã vòng tuyến tính

Notes

Các câu hỏi cần trả lời

- Các tham số đánh giá mã hóa kênh?
- Khoảng cách mã Hamming tối thiểu? Có vai trò gì trong việc đánh giá khả năng phát hiện lỗi và sửa lỗi của bộ mã?
- Mã khối tuyến tính? Ma trận sinh và ma trận kiểm tra của mã khối tuyến tính? Mã khối tuyến tính hệ thống?
- Bài toán thiết kế mã khối tuyến tính?
- Mã vòng (mã cyclic, mã xyclic) tuyến tính? Đa thức sinh và đa thức kiểm tra của mã vòng tuyến tính? Mã vòng tuyến tính hệ thống?

Notes

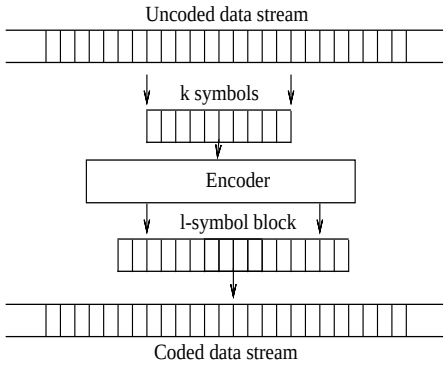
C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Một số định nghĩa và khái niệm cơ bản

Mã hóa khối



Hình: Quá trình mã hóa khối

Notes

Một số định nghĩa và khái niệm cơ bản

Véc-tơ mã

Định nghĩa (Véc-tơ mã)

Một bộ mã $\mathfrak{C} = \{c_0, c_1, \ldots, c_{M-1}\}$ *chứa các từ mã có độ dài* l*, mỗi từ mã* $c_k = (c_{k,0}, c_{k,1}, \ldots, c_{k,l-1})$ *với các dấu mã* $c_{k,i} \in GF(q)$ $(i = \overline{0, l-1})$.

- $\mathfrak{C}$*: bộ mã cơ số* q
- c_k *được gọi là từ mã, véc-tơ mã*
- M *là số từ mã của bộ mã* $\mathfrak{C}$.

Khối thông tin đầu vào là tập $\{m_i\}$, trong đó $m_i = (m_{i,0}, m_{i,1}, \ldots, m_{i,k-1})$ với $m_{i,j} \in GF(q)$. Tập $\{m_i\}$ tạo thành một không gian véc-tơ trên $GF(q)$.

- Nếu các khối thông tin có cùng độ dài k thì số từ mã của bộ mã $\mathfrak{C}$ phải thỏa mãn $M = q^k$.
- Nếu các khối tin có độ dài thay đổi thì M không có dạng trên.
  - Các bộ mã hóa loại này khó thực thi hơn.

Notes

Một số định nghĩa và khái niệm cơ bản

Độ dư thừa mã, Tỷ số mã, Trọng số mã

Định nghĩa (Độ dư thừa của bộ mã)

Độ dư thừa của bộ mã $\mathfrak{C}$ được định nghĩa là $r = l - \log_q(M)$.

- *Nếu* $M = 2^k$ *thì* $r = l - k$.

Định nghĩa (Tỷ số mã hóa)

Tỷ số mã hóa R được định nghĩa: $R = \frac{\log_q(M)}{l}$

- *Nếu* $M = 2^k$ *thì* $R = k/l$

Định nghĩa (Trọng số của từ mã/cấu trúc lỗi)

Trọng số của một từ mã c *hoặc của một cấu trúc lỗi* e *là số dấu mã khác* 0 *trong* c *hoặc* e. *Kí hiệu là* $w(\mathrm{c})$ *hoặc* $w(\mathrm{e})$

- $0 \leq w(\mathrm{c}) \leq l$

Notes

Một số định nghĩa và khái niệm cơ bản

Khoảng cách mã Hamming

Định nghĩa (Khoảng cách mã Hamming)

Khoảng cách Hamming giữa hai từ mã c_1 *và* c_2 *là tổng số vị trí tương ứng trong hai từ mã mà dấu mã khác nhau.*

$$d_{Hamming}(\mathrm{c}_1, \mathrm{c}_2) = d(\mathrm{c}_1, \mathrm{c}_2) = |\{i | c_{1,i} \neq c_{2,i}, i = 0, 1, ..., l-1\}|$$

- $d(\mathrm{c}_1, \mathrm{c}_2) = d(\mathrm{c}_2, \mathrm{c}_1)$.
- $0 \leq d(\mathrm{c}_1, \mathrm{c}_2) \leq l$.
- $d(\mathrm{c}_1, \mathrm{c}_2) + d(\mathrm{c}_2, \mathrm{c}_3) \geq d(\mathrm{c}_1, \mathrm{c}_3)$ (Bất đẳng thức tam giác).

Định nghĩa (Khoảng cách Hamming tối thiểu)

Khoảng cách mã tối thiểu, hay khoảng cách Hamming tối thiểu của một bộ mã khối $\mathfrak{C}$ *là khoảng cách Hamming tối thiểu giữa tất cả các cặp từ mã phân biệt trong bộ mã.*

$$d_{min} = d_0 = \min_{\forall \mathrm{c}_1, \mathrm{c}_2 \in \mathfrak{C}, \mathrm{c}_1 \neq \mathrm{c}_2} d(\mathrm{c}_1, \mathrm{c}_2)$$

Notes

Một số định nghĩa và khái niệm cơ bản

Khả năng phát hiện và sửa lỗi của mã

Định lý (Khả năng phát hiện lỗi của bộ mã)

Một bộ mã có khoảng cách mã tối thiểu d_{min} có khả năng phát hiện tất cả các cấu trúc lỗi có trọng nhỏ hơn hoặc bằng $(d_{min} - 1)$.

- **Chú ý**: Một số bộ mã có thể phát hiện được các cấu trúc lỗi có trọng $\geq d_{min}$

Định lý (Khả năng sửa lỗi của bộ mã)

Một bộ mã có khoảng cách mã tối thiểu d_{min} có khả năng sửa được tất cả các cấu trúc lỗi có trọng nhỏ hơn hoặc bằng $\lfloor \frac{d_{min}-1}{2} \rfloor$ [a].

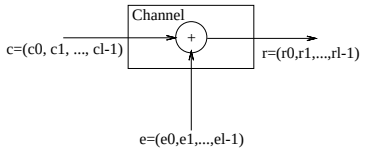
[a] $\lfloor x \rfloor$ là phần nguyên lớn nhất nhỏ hơn hoặc bằng x

- **Chú ý**: Một số bộ mã có thể sửa được các cấu trúc lỗi có trọng $\lfloor \frac{d_{min}-1}{2} \rfloor + 1$ hoặc lớn hơn.

Notes

Một số định nghĩa và khái niệm cơ bản

Mô hình mã truyền dẫn trong kênh có nhiễu



Hình: Mô hình kênh nhiễu cộng

- c: từ mã phát, e: cấu trúc lỗi, r = c + e: véc-tơ thu.
  - Nếu không có lỗi thì véc-tơ thu là một từ mã hợp lệ.
- Định dạng điều chế, mức công suất phát, và mức nhiễu trên kênh quyết định xảy ra một cấu trúc lỗi trong q^l cấu trúc lỗi có thể.

- Máy thu thực hiện việc xem xét véc-tơ thu có phải là từ mã hợp lệ hay không: quá trình phát hiện lỗi.
- Khi máy thu phát hiện lỗi:
  1. Yêu cầu phát lại: thông qua ARQ
  2. HOẶC Đánh dấu từ mã lỗi: với các ứng dụng real-time (voice, video,...)
  3. HOẶC Sửa lỗi: FEC.

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Mã khối tuyến tính

Định nghĩa

Định nghĩa (Mã khối tuyến tính)

Xét một bộ mã khối $\mathfrak{C}$ gồm các từ mã độ dài l $\{c_k = (c_{k,0}, c_{k,1}, \ldots, c_{k,l-1})\}$ với các dấu mã thuộc $GF(q)$. Bộ mã $\mathfrak{C}$ là một bộ mã khối tuyến tính cơ số q nếu và chỉ nếu $\mathfrak{C}$ tạo thành một không gian véc-tơ con trên $GF(q)$.

Định nghĩa (Chiều của một bộ mã khối)

Chiều của một bộ mã khối là chiều của không gian véc-tơ tương ứng.

- Ký hiệu: $\mathfrak{C}(l, k)$ hoặc $\mathfrak{C}(l, k, d_0)$.

1. Tổ hợp tuyến tính của một tập các từ mã bất kỳ là một từ mã $\Rightarrow \mathfrak{C}$ luôn chứa từ mã toàn 0
2. Khoảng cách mã tối thiểu của bộ mã khối tuyến tính bằng trọng số của một từ mã có trọng số nhỏ nhất khác từ mã toàn không.
3. Các cấu trúc lỗi không thể phát hiện được của bộ mã độc lập với từ mã phát và luôn chứa tập tất cả các từ mã không toàn 0.

Notes

Mã khối tuyến tính

Ma trận sinh của mã khối tuyến tính

Gọi $\{g_0, g_1, \ldots, g_{k-1}\}$ là cơ sở của các từ mã trong bộ mã $\mathfrak{C}(l, k)$.

Ma trận sinh G$(k \times l)$của bộ mã được thành lập như sau:

$$G = \begin{pmatrix} g_0 \\ g_1 \\ \cdots \\ g_{k-1} \end{pmatrix} = \begin{pmatrix} g_{0,0} & g_{0,1} & \cdots & g_{0,l-1} \\ g_{1,0} & g_{1,1} & \cdots & g_{1,l-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,l-1} \end{pmatrix}$$

Gọi $a = (a_0, a_1, \ldots, a_{k-1})$ là khối dữ liệu đầu vào (bản tin) cần mã hóa.

Từ mã thu được từ phép mã hóa:

$$\begin{aligned} c &= aG = [a_0, a_1, \ldots, a_{k-1}]G \\ &= a_0 g_0 + a_1 g_1 + \cdots + a_{k-1} g_{k-1} \end{aligned}$$

Notes

Mã khối tuyến tính

Ma trận kiểm tra tính chẵn lẻ

Với $\mathfrak{C}$, tồn tại $\mathfrak{C}^{\perp}$ là không gian véc-tơ đối ngẫu $(l-k)$ chiều.

Gọi $\{h_0, h_1, \ldots, h_{l-k-1}\}$ là cơ sở của $\mathfrak{C}^{\perp}$. $\Rightarrow$ Ma trận sinh $H(l-k \times l)$ của $\mathfrak{C}^{\perp}$:

$$H = \begin{pmatrix} h_0 \\ h_1 \\ \ldots \\ h_{l-k-1} \end{pmatrix} = \begin{pmatrix} h_{0,0} & h_{0,1} & \ldots & h_{0,l-1} \\ h_{1,0} & h_{1,1} & \ldots & h_{1,l-1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{l-k-1,0} & h_{l-k-1,1} & \ldots & h_{l-k-1,l-1} \end{pmatrix}$$

- H là ma trận kiểm tra chẵn lẻ của mã $\mathfrak{C}$
- $GH^T = 0$.

Định lý

Một véc-tơ c *là một từ mã thuộc* $\mathfrak{C}$ *nếu và chỉ nếu* $cH^T = 0$

- $cH^T = 0$ gọi là biểu thức kiểm tra chẵn lẻ.

Notes

Mã khối tuyến tính

Ma trận kiểm tra tính chẵn lẻ và khoảng cách mã

Định lý

Giả sử bộ mã $\mathfrak{C}$ *có ma trận kiểm tra tính chẵn lẻ* H. *Khoảng cách mã tối thiểu của bộ mã* $\mathfrak{C}$ *bằng số cột tối thiểu khác* 0 *của* H *mà tổ hợp tuyến tính không tầm thường của chúng bằng* 0.

Định lý (Giới hạn Singleton)

Với bộ mã khối tuyến tính $\mathfrak{C}(l,k)$, *khoảng cách mã tối thiểu thỏa mãn bất đẳng thức:*

$$d_{min} \leq l - k + 1$$

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Mã khối tuyến tính

Mã khối tuyến tính hệ thống

Định nghĩa (Mã khối tuyến tính hệ thống)

Mã khối tuyến tính hệ thống $\mathfrak{C}(l,k)$ thực hiện việc ánh xạ bản tin (khối dữ liệu) độ dài k thành một véc-tơ/từ mã độ dài l sao cho trong số l bít có thể chỉ ra k bít bản tin và số còn lại $l-k$ bít kiểm tra tính chẵn lẻ.

Giả sử từ mã xây dựng mã có dạng $c = [p_1 \quad | \quad a]$

- a: khối thông tin (bản tin) độ dài k; p_1: khối bít kiểm tra độ dài $l-k$

$G \xrightarrow{\text{phương pháp khử Gausse}} G = [P \quad | \quad I_k]$

- $P_{(k\times l-k)}$: ma trận tạo dấu kiểm tra
- $I_{(k\times k)}$: ma trận đơn vị.

$\Rightarrow H = [I_{l-k} \quad | \quad -P^T]$

- $\Rightarrow GH^T = 0$

Chý ý: Nếu xét $c = [a \quad | \quad p_1]$

- $G = [I_k \quad | \quad P]$
- $\Rightarrow H = [-P^T \quad | \quad I_{l-k}]$

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Ví dụ

Ví dụ

Xét bộ mã nhị phân đều chiều dài l (ví dụ bộ mã nhị phân đều chiều dài 2: $\mathfrak{C} = (00), (01), (11), (10)$). Giả sử kết quả mã hóa được truyền qua kênh nhị phân rời rạc đối xứng không nhớ (BSC) có xác suất thu sai p_0, các bít được phát đi độc lập nhau, và xác suất phát đi bít 0 và bít 1 tương đương nhau.

1. *Tính xác suất thu được một từ mã đúng.*
2. *Giả sử xác suất sai cho phép đối với việc thu các từ mã là p_a, tìm điều kiện đối với p_0 để có thể sử dụng được bộ mã cho việc thông tin qua kênh.*

Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Đánh giá khả năng phát hiện lỗi

Cho $\mathfrak{C}(l, k, d_{min})$ truyền qua kênh BSC có xác suất chuyển sai p.

- $P_u(E)$: xác suất véc-tơ thu có lỗi mà không phát hiện được.
- $P_e(E)$: xác suất véc-tơ thu có lỗi.
- $P_d(E)$: xác suất véc-tơ thu có lỗi được phát hiện.

$$
\begin{aligned}
P_u(E) &\leq \sum_{j=d_{min}}^{l} \binom{l}{j} p^j (1-p)^{l-j} = 1 - \sum_{j=0}^{d_{min}-1} \binom{l}{j} p^j (1-p)^{l-j} \\
P_u(E) &= \sum_{j=d_{min}}^{l} A_j p^j (1-p)^{l-j} \\
P_e(E) &= \sum_{j=1}^{l} \binom{l}{j} p^j (1-p)^{l-j} = 1 - (1-p)^l \\
P_d(E) &= P_e(E) - P_u(E) = 1 - (1-p)^l - P_u(E)
\end{aligned}
$$

Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Đánh giá khả năng phát hiện lỗi (cont.)

- P_{u_b}: tỷ lệ bít lỗi không được phát hiện
  - $\triangleq$ xác suất bít thông tin nhận được bị lỗi trong một từ mã bị tác động bởi cấu trúc lỗi không phát hiện được
  - $P_u(E) \geq P_{u_b}(E) \geq \frac{1}{k} P_u(E)$
- P_{d_b}: tỷ lệ bít lỗi được phát hiện
  - $\triangleq$ xác suất bít thông tin nhận được bị lỗi trong một từ mã bị tác động bởi cấu trúc lỗi có thể phát hiện được.
  - $P_d(E) \geq P_{d_b}(E) \geq \frac{1}{k} P_d(E)$
- Nếu biết phân bố trọng của bộ mã, P_{u_b} có thể tính một cách chính xác:

$$P_{u_b} = \sum_{j=d_{min}}^{l} \frac{B_j}{k} p^j (1-p)^{l-j}$$

trong đó B_j là tổng trọng của các khối tin tương ứng với tất cả các từ mã có trọng là i.

Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Đánh giá khả năng sửa lỗi

Cho $\mathfrak{C}(l, k, d_{min})$ truyền qua kênh BSC có xác suất chuyển sai p.

Xét bộ giải mã có độ dài giới hạn.

- $P(E)$: xác suất giải mã sai

$$P(E) \leq \sum_{j=\lfloor \frac{d_{min}-1}{2} \rfloor + 1}^{l} \binom{l}{j} p^j (1-p)^{l-j} = 1 - \sum_{j=0}^{\lfloor \frac{d_{min}-1}{2} \rfloor} \binom{l}{j} p^j (1-p)^{l-j}$$

Đẳng thức xảy ra chỉ khi mã là hoàn hảo.

- $P(F)$: xác suất giải mã thất bại

$$P(F) \leq 1 - \sum_{j=0}^{\lfloor \frac{d_{min}-1}{2} \rfloor} \binom{l}{j} p^j (1-p)^{l-j}$$

Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Đánh giá khả năng sửa lỗi (cont')

Xét $\mathfrak{C}(l, k, d_{min})$ với phân bố trọng số đã biết $\{A_i\}$

$P_k^j \triangleq$ xác suất một véc-tơ thu có khoảng cách Hamming chính xác là k so với một từ mã có trọng là j.

$$P_k^j = \sum_{r=0}^{k} \binom{j}{k-r} \binom{l-j}{r} p^{j-k+2r} (1-p)^{l-j+k-2r}$$

$$\begin{aligned} P(E) &= \sum_{j=d_{min}}^{l} A_j \sum_{k=0}^{\lfloor \frac{d_{min}-1}{2} \rfloor} P_k^j \\ P(F) &= 1 - \sum_{j=0}^{\lfloor \frac{d_{min}-1}{2} \rfloor} \binom{l}{j} p^j (1-p)^{l-j} - P(E) \end{aligned}$$

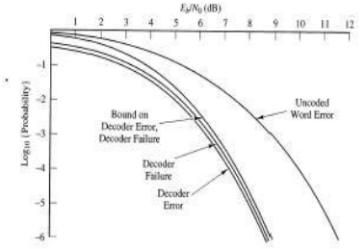
Notes

Đánh giá mã khối nhị phân tuyến tính trên kênh BSC

Đánh giá khả năng sửa lỗi (cont')

- Nếu biết được mối quan hệ giữa trọng số của các khối tin và trọng số các từ mã tương ứng
  - $\Rightarrow B_j$
- $\Rightarrow$

$$BER = P_b(E) = \frac{1}{k}\sum_{j=d_{min}}^{l} B_j \sum_{k=0}^{\lfloor \frac{d_{min}-1}{2} \rfloor} P_k^j$$



Chú ý: Thường, thông tin $\{B_j\}$ không khả thi.

- $\Rightarrow$ Chủ yếu dựa vào các đánh giá biên

$$P(E) \geq P_b(E) \geq \frac{1}{k}P(E)$$

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Các vấn đề khi thiết kế mã khối tuyến tính

Thiết kế mã khối tuyến tính tối ưu

Khi thiết kế, ta mong muốn có được bộ mã có độ dư thừa nhỏ nhất có thể, nhưng lại có khả năng phát hiện và sửa lỗi lớn nhất có thể.

Trường hợp 1

Với k và d_{min} cho trước, xây dựng bộ mã có độ dư thừa tối thiểu: $\min\{l\}$.
Độ dài từ mã của bộ mã thỏa mãn giới hạn Griesmer:

$$l \geq \sum_{i=0}^{k-1} \lceil \frac{d_{min}}{2^i} \rceil$$

$\lceil x \rceil$: phần nguyên nhỏ nhất lớn hơn hoặc bằng x.

Notes

Các vấn đề khi thiết kế mã khối tuyến tính

Thiết kế mã khối tuyến tính tối ưu (cont')

Trường hợp 2

Với l và k cho trước, xây dựng bộ mã có khả năng phát hiện và sửa sai lớn nhất: $\max\{d_{min}\}$.
Khoảng cách Hamming tối thiểu của bộ mã thỏa mãn giới hạn Plotkin:

$$d_{min} \leq \frac{l \times 2^{k-1}}{2^k - 1}$$

Trường hợp 3

Với l và khả năng sửa sai t cho trước, xây dựng bộ mã có độ dư thừa nhỏ nhất: $\max\{k\}$.
Mối liên hệ giữa l, k và t thỏa mãn giới hạn Hamming:

$$2^{l-k} \geq \sum_{i=0}^{t} \binom{l}{i}$$

Notes

C4: Mã hóa kênh - Truyền dẫn dữ liệu (Part 1)

Nội dung chính

Notes

Kết thúc phần mã khối tuyến tính

Notes