Shannonova věta o kanálovém kódování

Informační kanál $\mathcal{K}$ lze opakovaně použít

$$\mathcal{K} : X \longrightarrow Y \; : \; x \rightarrow y$$

opakování:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Budeme předpokládat nezávislost akcí kanálu $\mathcal{K}$.

<u>Definice</u> <u>Bezpaměťové rozšíření</u> kanálu $\mathcal{K}$ (n-tého řádu) je kanál

$$\mathcal{K}^n = \langle X^n, p^n(\vec{y}|\vec{x}), Y^n \rangle, \text{ kde}$$

$$\forall \; \vec{y} \in Y^n, \; \vec{x} \in X^n$$

$$p^n(\vec{y}|\vec{x}) = p(y_1|x_1) \cdot p(y_2|x_2) \cdots p(y_n|x_n)$$

· v reálných situacích (rádiové, optické spojení) je třeba předpokládat markovskost.

· Přenášíme zprávy W (zprávy) $W = \{1, 2, \dots, M\}$ s distribucí $p(w)$

· Kodér (podle zadání) $k : W \to X^n$ (slova z abecedy X)

$$\left.\begin{aligned} k(1) &= \vec{x}(1) \\ &\vdots \\ k(M) &= \vec{x}(M) \end{aligned}\right\} \text{kódová kniha}$$

· dekodér: $\delta : Y^n \to W$

schéma přenosu informace

$$w \to \boxed{\text{kodér } k} \xrightarrow[\text{něco z } X^n]{} \boxed{\mathcal{X} = \langle X^n, p^n(y|x), Y^n \rangle} \xrightarrow[\text{něco z } Y^n]{} \boxed{\delta} \to \hat{w}$$

Definice: ~~Pst~~ (M, n)-kód je dvojice kodér k a dekodér δ.

Podmíněná pravděpodobnost chyby, $1 \le i \le M$

$$\lambda_i = \Pr\left(\delta(Y) \neq i \mid X = \vec{x}(i)\right) =$$

$$= \sum_{\{\vec{y} \in Y^n \mid \delta(\vec{y}) \neq i\}} p^n(\vec{y} \mid \vec{x}(i)).$$

Maximální pravděp. chyby

$$\lambda^{(n)} = \max_{i=1,\dots,M} \lambda_i$$

Průměrná pravděpodobnost chyby

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i$$

Má-li W rovnoměrné rozdělení, pak $\Pr[W=i] = 1/M$ $\forall i$ a tedy

$$P_e^{(n)} = \sum_{i=1}^{M} \Pr(\hat{W} \neq W \mid W=i)\, P[W=i]$$

$$= \Pr(\hat{W} \neq W).$$

Definice: Rychlost přenosu (M,n) kódem v kanálu $\mathcal{K}$ je

$$R = \frac{\log M}{n}.$$

Rychlost $R \geq 0$ je v kanálu $\mathcal{K}$ dosažitelná, pokud existuje posloupnost $(2^{nR}, n)$-kódů s vlastností

$$\lim_{n\to\infty} \lambda^{(n)} = 0.$$

BOBO

Operační kapacita, $C_{oper}(\mathcal{K})$ je supremum rychlostí, R, které jsou v kanálu dosažitelné.

Příklad $\mathcal{K} = BSK(p)$

zdroj $W = \{1, 2, \dots, 8\}$ $\quad M = 8$

$C = \{\vec{c}(1), \vec{c}(2), \dots, \vec{c}(8)\} \in \{0,1\}^8$

$\vec{c}(i) = i_0 i_1 i_2 i_3 i_0 i_1 i_2 i_3$, $\quad n = 8$

kde $i - 1 = 2^2 i_0 + i_1 2^1 + i_2 2^0$ $\quad$ a $i_3 = i_0 + i_1 + i_2 \pmod 2$

$\vec{c}(1) = 0000\,0000$

$\vec{c}(2) = 0011\,0011$

$\vec{c}(3)\ 0101\,0101$

$\vec{c}(4) = 0110\,0110$

$\vec{c}(5) = 1001\,1001$

$\vec{c}(6) = 1010\,1010$

$\vec{c}(7) = 1100\,1100$

$\vec{c}(8) = 1111\,1111$

Vyslanému kódovému slovu $\vec{u}(i)$ odpovídá

$$Y = (i_0 + z_1, i_1 + z_2, i_2 + z_3, i_3 + z_4, i_0 + z_5, i_1 + z_6, i_2 + z_7, i_3 + z_8)$$

$\vec{z} = (z_1, z_2, \dots, z_8)$ realizace kopie $(1-p, p)$

Počet chyb $\bar{u}(z) = \sum_{i=1}^{8} z_i$

s binomickým rozdělením s parametry 8 a p

$$P[\bar{u}(z) = k] = \binom{8}{k} p^k (1-p)^{8-k}, \quad 0 \le k \le 8$$

Pro nejvýše jedné chyby

$$P[\bar{u}(z) = 0] + P[\bar{u}(z) = 1] = (1-p)^8 + 8p(1-p)^7$$

Nechť $B(i)$ je množina přijatých zpráv odpovídajících vyslané zprávě $\vec{u}(i)$ a nejvýše jedné chyba

$$|B(i)| = 9$$

Definujme dekodér $\delta : \{0,1\}^8 \to W$

$$\delta(y) = i \quad \text{jestliže } y \in B(i), \ 1 \le i \le 8$$

a libovolné pro ostatní prvky

$C(Y) \neq i$ za podmínky $x \cdot \alpha^{-1}(i)$ je možný opravit při přenosu s ~~nejvýše~~ tj. víc než jednou chybou.
Pravděp. takového je

$$\cancel{\langle 0,1 \rangle}\ \lambda = 1 - (1-p)^8 - 8p(1-p)^7$$

$$\lambda^{(8)} = \lambda$$

$$P_e^{(8)} = \frac{1}{8}\sum_{i=1}^{8} \lambda_i$$

Např. pro $p = 10^{-3}$ platí

$\lambda = 2{,}78 \cdot 10^{-5}$

"3 ze 100 000 zpráv jsou špatně"

Rychlost $R = \dfrac{\log_2 8}{8} = \dfrac{3}{8}\ \dfrac{\text{bit}}{\text{znak}}$

<u>Shannonova věta</u>: Operační kapacita kanálu $K = \langle X, p(y|x), Y \rangle$ je totožná s jeho informační kapacitou, tj.

$$C_{oper}(K) = C(K)$$

BOBO

Věta: Je-li $(\mathcal{U}_1, \mathcal{U}_2, \ldots)$ stacionární proces pro který platí, že

$$-\frac{1}{n} \log p(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_n) \to H(\mathcal{U}_1)$$

ve smyslu kvadratických průměrů. Pak je tento zdroj přenositelný kanálem $\mathcal{K} = \langle X, p(y|x), Y \rangle$ když $H(\mathcal{U}_1) < C(\mathcal{K})$ a není přenositelný když $H(\mathcal{U}_1) > C(\mathcal{K})$.

Příklad: $BSK(10^{-3})$ má

$C(\mathcal{K}) = 1 - h(10^{-3}) = 0{,}9886$ bitů/znak

lze tedy přenášet informaci rychlostí 0,9886 bitů/znak s pravděpodobností chyby libovolně malou.

měli jsme $R = 3/8 = 0{,}375$ při chybě menší než $3 \cdot 10^{-5}$.