Optymalne przesyłanie kanałem
(notatki do wykładu z Teorii Informacji)

Bogusław Kluge, Marta Łuksza

Definicja 1. Niech $\mathcal{A}$, $\mathcal{B}$ będą sygnałami, Γ kanałem, a Δ regułą decyzyjną.

$$\mathcal{A} \longrightarrow \boxed{\Gamma} \longrightarrow \mathcal{B} \xrightarrow{\Delta} \mathcal{A}$$

Definiujemy średnie prawdopodobieństwo trafnego odczytu

$$\mathrm{Pr_c}(\Delta, \mathcal{A}) := \sum_{b \in \mathcal{B}} \mathrm{p}(\mathcal{B} = b)\, \mathrm{p}(\mathcal{A} = \Delta(b) | \mathcal{B} = b)$$

oraz średnie prawdopodobieństwo błędnego odczytu

$$\mathrm{Pr_e}(\Delta, \mathcal{A}) := 1 - \mathrm{Pr_c}(\Delta, \mathcal{A}).$$

Przykład 1. Dla symetrycznego binarnego kanału $\left(\begin{smallmatrix} \pi & \bar{\pi} \\ \bar{\pi} & \pi \end{smallmatrix}\right)$ i reguły identycznościowej $\Delta(i) = i$ ($\Delta \equiv \mathrm{id}$), mamy

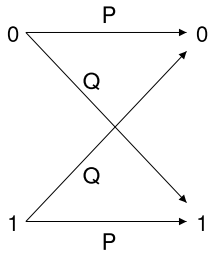
$$\begin{aligned} \mathrm{Pr_c}(\Delta, \mathcal{A}) &= \sum_{i \in \{0,1\}} \mathrm{p}(\mathcal{B} = i)\, \mathrm{p}(\mathcal{A} = \mathrm{id}(i) | \mathcal{B} = i) \\ &= \sum_{i \in \{0,1\}} \mathrm{p}(\mathcal{A} = \mathrm{id}(i))\, \mathrm{p}(\mathcal{B} = i | \mathcal{A} = \mathrm{id}(i)) \\ &= \sum_{i \in \{0,1\}} \mathrm{p}(\mathcal{A} = \mathrm{id}(i))\pi \\ &= \pi. \end{aligned}$$

Chcemy pomimo błędów kanału komunikować się możliwie wiernie.

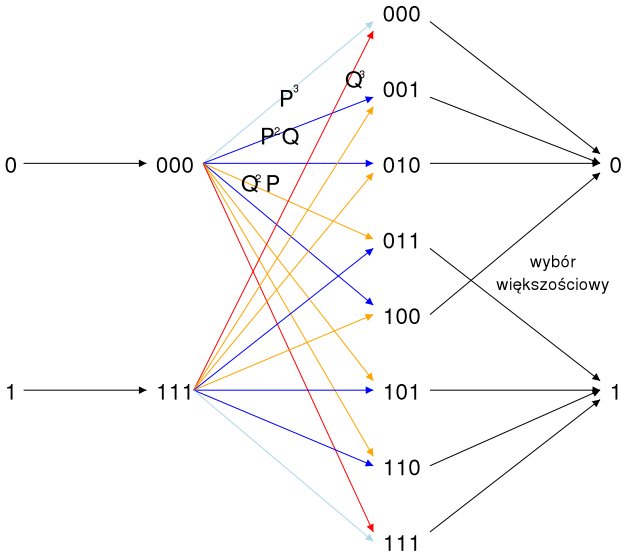
Załóżmy, że dysponujemy kanałem Γ_1 (rysunek 1) o macierzy $\left(\begin{smallmatrix} P & Q \\ Q & P \end{smallmatrix}\right)$. Możemy wówczas, poprzez powielanie przesyłanych bitów skonstruować kanał Γ_3 (rysunek 2) o macierzy $\left(\begin{smallmatrix} P^3+3P^2Q & Q^3+3Q^2P \\ Q^3+3Q^2P & P^3+3P^2Q \end{smallmatrix}\right)$.

Jeśli $1 > P = 1 - Q > \frac{1}{2}$, to różnica prawdopodobieństw błędnego przekazu w kanałach Γ_1 oraz Γ_3 wynosi $Q - (Q^3 + 3Q^2P) = Q(2Q-1)(Q-1) > 0$. Zatem w kanale Γ_3 prawdopodobieństwo błędnego przekazu jest mniejsze niż w kanale Γ_1.

Ogólnie zamiast 3-krotnie możemy powielać bity n-krotnie (gdzie n jest nieparzyste, żeby wybór większościowy miał sens). Wtedy, przy założeniu $1 > P =$



Rysunek 1: Kanał Γ_1



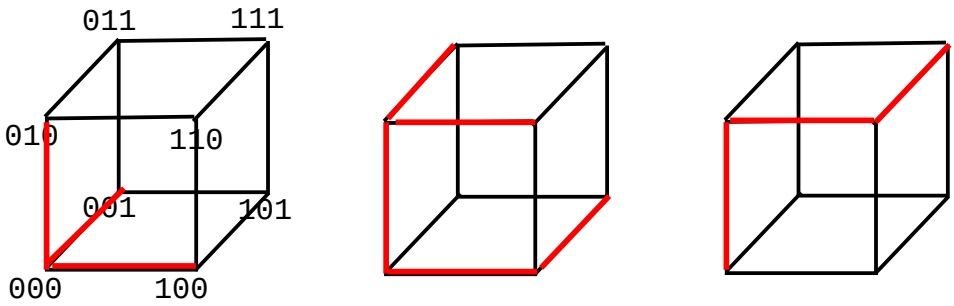
Rysunek 2: Kanał Γ_3

$1 - Q > \frac{1}{2}$ prawdopodobieństwo błędnego przekazu wynosi

$$\begin{aligned}
\sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} P^i Q^{n-i} &\leqslant \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} P^{\lfloor \frac{n}{2} \rfloor} Q^{\lceil \frac{n}{2} \rceil} \\
&= \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} P^{\lceil \frac{n}{2} \rceil - 1} Q^{\lceil \frac{n}{2} \rceil} \quad \text{(bo } n \text{ jest nieparzyste)} \\
&< \frac{1}{P} (PQ)^{\lceil \frac{n}{2} \rceil} \sum_{i=0}^{n} \binom{n}{i} \\
&< \frac{1}{P} (PQ)^{\frac{n}{2}} 2^n = \frac{1}{P} (\underbrace{4PQ}_{<1})^{\frac{n}{2}} \longrightarrow 0.
\end{aligned}$$

Definicja 2 (Odległość Hamminga). Niech $u, v \in \mathrm{A}^n$. Definiujemy odległość Hamminga

$$\delta(u, v) := |\{\, i \mid u_i \neq v_i \}|.$$



Rysunek 3: Na czerwono ścieżki od punktu 000 do punktów w odległości Hamminga 1, 2 i 3 od niego.

Fakt 1 (Odległość Hamminga jest metryką). *Niech $u, v, w \in \mathrm{A}^n$. Odległość Hamminga spełnia:*

$$\begin{aligned}
&\delta(u, v) = 0 \iff u = v,\\
&\delta(u, v) \geqslant 0,\\
&\delta(u, v) = \delta(v, u),\\
&\delta(u, w) \leqslant \delta(u, v) + \delta(v, w).
\end{aligned}$$

Zajmiemy się teraz ogólniejszym wykorzystaniem binarnego symetrycznego kanału.

Dana jest zmienna losowa $\mathcal{A}$ o wartościach $\mathrm{A} = \{a_1, a_2, \ldots, a_n\}$ oraz kod

$$\mathcal{C} : \mathrm{A} \longrightarrow \{0,1\}^k \subseteq \{0,1\}^*.$$

Używając kanału $\left(\begin{smallmatrix} P & Q \\ Q & P \end{smallmatrix}\right)$ produkujemy ciągi w $\{0,1\}^k$. Następnie stosujemy regułę maksymalnego podobieństwa $\Delta\colon \{0,1\}^k \longrightarrow \overrightarrow{\mathcal{C}}(\mathrm{A})$. Wtedy

$$\begin{aligned}
\mathrm{Pr_c}(\Delta, \mathcal{CA}) &= \sum_{u \in \{0,1\}^k} \mathrm{p}(\mathcal{B} = u)\, \mathrm{p}(\mathcal{CA} = \Delta(u) | \mathcal{B} = u)\\
&= \sum_{u \in \{0,1\}^k} \mathrm{p}(\mathcal{CA} = \Delta(u))\, \mathrm{p}(\mathcal{B} = u | \mathcal{CA} = \Delta(u)),
\end{aligned}$$

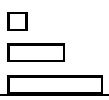
	Chcemy przesłać	Faktycznie przesyłamy	Wydłużenie (krotność)	Błąd
Kanał wierny	0 101 01101	0 101 01101	1	0
Metoda powielania	0 101 01101	0^n $1^n0^n1^n$ $0^n1^n1^n0^n1^n$	$n \nearrow \infty$	$\searrow 0$
Twierdzenie Shannona	0 101 01101		$\nearrow C_\Gamma$	$\searrow 0$

Tabela 1: Podsumowanie

gdzie

$$\mathrm{p}(\mathcal{B} = u) = \sum_{i=1}^{m} \mathrm{p}(\mathcal{A} = a_i)\, \mathrm{p}(\mathcal{C}(a_i) \longrightarrow u),$$
$$\mathrm{p}(v \longrightarrow u) = \prod_{i=1}^{k} \mathrm{p}(v_i \longrightarrow u_i) = P^{k-\delta(u,v)} Q^{\delta(u,v)},$$

bo

$$\mathrm{p}(v_i \longrightarrow u_i) = \begin{cases} P & \text{gdy } v_i = u_i \\ Q & \text{gdy } v_i \neq u_i. \end{cases}$$

Przypuśćmy, że $P > Q$. Wtedy reguła maksymalnego podobieństwa sprowadza się do reguły minimalnej odległości Hamminga

$$\Delta_H(v) = u \in \overrightarrow{\mathcal{C}}(\mathrm{A}), \text{takie że } \delta(v,u) \text{ jest minimalne}$$

(*reguła dobrosąsiedzka*).

Definicja 3 (Stopa kodu). Niech $\mathcal{A}$ będzie zmienną losową o wartościach $\mathrm{A} = \{a_1, a_2, \ldots, a_m\}$. Dla kodu $\mathcal{C}\colon \mathrm{A} \longrightarrow \{0,1\}^n$ definiujemy stopę kodu (ang. *rate*)

$$\mathrm{R}(\mathcal{C}) := \frac{\log_2 m}{n}$$

(zachodzi $1 \geqslant \mathrm{R}(\mathcal{C}) > 0$).

Dla danego kanału Γ nad $\{0,1\}$ o macierzy $\left(\begin{smallmatrix} P & Q \\ Q & P \end{smallmatrix}\right)$, gdzie $P > \frac{1}{2} > Q$, zmiennej losowej $\mathcal{A}$ o wartościach $\mathrm{A} = \{a_1, a_2, \ldots, a_m\}$ oraz kodu $\mathcal{C}\colon \mathrm{A} \longrightarrow \{0,1\}^n$, wygodnie jest utożsamiać $\overrightarrow{\mathcal{C}}(\mathrm{A}) = \mathcal{C}$ i A (a zatem także $\mathcal{A}$ i $\mathcal{CA}$).

Wykorzystujemy Γ jako kanał z $\mathcal{C}$ do $\{0,1\}^n$ o macierzy

$$\begin{pmatrix} \mathrm{p}(c_1 \longrightarrow u_1) & \cdots & \mathrm{p}(c_1 \longrightarrow u_{2^n}) \\ \vdots & \ddots & \vdots \\ \mathrm{p}(c_m \longrightarrow u_1) & \cdots & \mathrm{p}(c_m \longrightarrow u_{2^n}) \end{pmatrix},$$

gdzie $\mathrm{p}(v \longrightarrow u) = P^{n-\delta(u,v)}Q^{\delta(u,v)}$, $\mathcal{C} = \{c_1, c_2, \ldots, c_m\}$. Wtedy

$$\Delta(u) = v, \text{takie że } \mathrm{p}(v \longrightarrow u) \text{ największe } (\iff \delta(u,v) \text{ najmniejsze})$$

(reguła dobrosąsiedzka).

Uwaga 1 *(Nieformalnie).* Znajdujemy słowo, które najmniej (w sensie Hamminga) różni się od słowa wychodzącego z kanału. Na przykład, gdy dostaniemy słowo *barcara* i mamy do wyboru *bar**b**ara* oraz *bar**n**a**b**a*, to wybierzemy pierwsze z nich.

Fakt 2 (Użyteczna transformacja). *Dla kanału z $\mathcal{A}$ do $\mathcal{B}$, mamy*

$$\begin{aligned} \mathrm{Pr_e}(\Delta, \mathcal{A}) &= 1 - \mathrm{Pr_c}(\Delta, \mathcal{A}) \\ &= 1 - \sum_{b \in \mathcal{B}} \mathrm{p}(\mathcal{B} = b)\, \mathrm{p}(\mathcal{A} = \Delta(b) | \mathcal{B} = b) \\ &= \sum_{b \in \mathcal{B}} \sum_{a \in \mathcal{A} \setminus \{\Delta(b)\}} \mathrm{p}(\mathcal{B} = b \wedge \mathcal{A} = a) \\ &= \sum_{a \in \mathcal{A}} \sum_{b \notin \Delta^{-1}\{a\}} \mathrm{p}(\mathcal{B} = b \wedge \mathcal{A} = a) \\ &= \sum_{a \in \mathcal{A}} \mathrm{p}(\mathcal{A} = a \wedge \Delta\mathcal{B} \neq a) \\ &= \sum_{a \in \mathcal{A}} \mathrm{p}(a)\, \mathrm{p}(\Delta\mathcal{B} \neq a | \mathcal{A} = a) \end{aligned}$$

Twierdzenie Shannona. *Niech Γ będzie binarnym symetrycznym kanałem o macierzy $\left(\begin{smallmatrix} P & Q \\ Q & P \end{smallmatrix}\right)$, gdzie $P > \frac{1}{2} > Q$. Wtedy przepustowość $\mathrm{C}_\Gamma = 1 - \mathrm{H}_2(P)$. Niech $\delta, \varepsilon > 0$.*

Wówczas dla każdego dostatecznie dużego n, istnieje kod $\mathcal{C}$ o zbiorze wartości $\mathcal{C} \subseteq \{0,1\}^n$, taki że

$$\mathrm{Pr_e}(\Delta, \mathcal{C}) < \delta$$

oraz

$$\mathrm{C}_\Gamma - \varepsilon \leqslant \mathrm{R}(\mathcal{C}) < \mathrm{C}_\Gamma,$$

gdzie Δ jest regułą dobrosąsiedzką.

Uwaga 2. Z tego wynika, że można wskazać ciąg kodów $\mathcal{C}_n$, taki że $\mathrm{Pr_e}(\Delta, \mathcal{C}_n) \longrightarrow 0$ i $\mathrm{R}(\mathcal{C}_n) \longrightarrow \mathrm{C}_\Gamma$.

Uwaga 3 *(Bardzo nieformalnie).* Dla dużych n, co najmniej $\approx 2^{\mathrm{C}_\Gamma n}$ słów można przesłać w miarę wiernie.