

Dowód twierdzenia Shannona

Łukasz Chmielewski, Marzena Szałas

05 maj 2004

Teraz przeprowadzimy **dowód twierdzenia Shannona**:

Wybieramy $\eta > 0$ (w zależności od $\delta, \varepsilon > 0$, „bardzo małe”).

$$\rho = n(Q + \eta),$$

gdzie $Q = E(\chi_i)$ dla $i \in \{0, \dots, n\}$ (wartość oczekiwana zmiennych wspomnianych wcześniej – jest taka sama dla wszystkich, gdyż mają one ten sam rozkład).

Gdy $\delta(u, u \oplus e) \leq \rho$ (gdzie δ jest odległością Hamminga) i $\forall_{v \in C - \{u\}} \delta(v, u \oplus e) > \rho$, to $\Delta(u \oplus e) = u$.

Zatem jeśli $\Delta(u \oplus e) \neq u$ to $(\delta(u, u \oplus e) > \rho$ lub $\exists_{v \in C - \{u\}} \delta(v, u \oplus e) \leq \rho)$.

Stąd wynika:

$$p(\Delta(u \oplus e) \neq u) \leq p(\delta(u, u \oplus e) > \rho) + \sum_{u' \in C \setminus \{u\}} p(\delta(u', u \oplus e) \leq \rho)$$

Sumując stronami powyższą nierówność osiągamy (po wszystkich e):

$$p(\Delta(u \oplus \mathcal{E}) \neq u) \leq p(\delta(u, u \oplus \mathcal{E}) > \rho) + \sum_{u' \in C \setminus \{u\}} p(\delta(u', u \oplus \mathcal{E}) \leq \rho)$$

Następnie będziemy szacować obie części prawej strony powyższej nierówności.

Słabe prawo wielkich liczb

Mamy ciąg zmiennych losowych: $\chi_1, \chi_2, \dots$ o identycznym rozkładzie. Elementy każdego skończonego podciągu (którego elementami są początkowe elementy tego ciągu) $= \chi_1, \chi_2, \dots, \chi_n$, są niezależne, czyli:

$$p(\chi_1 = a_1 \cap \chi_2 = a_2 \cap \dots \cap \chi_n = a_n) = p(\chi_1 = a_1) \times p(\chi_2 = a_2) \times \dots \times p(\chi_n = a_n)$$

Wtedy:

$$p(|\frac{1}{n} \times \sum_{i=1}^n \chi_i - \mu| > \alpha) \rightarrow 0, \text{ gdy } n \rightarrow \infty$$

gdzie $\mu = E(\chi_i)$.

Ponieważ $\sum_{i=1}^n \chi_i = \delta(u, u \oplus \mathcal{E})$ (pamiętajmy, że $\mathcal{E} = (\chi_1, \chi_2, \dots, \chi_n)$) i ze słabego prawa wielkich liczb, dla dostatecznie dużych n wynika, że

$$p(\delta(u, u \oplus \mathcal{E}) > \rho) = p(\frac{1}{n} \times \sum_{i=1}^n e_i > Q + \eta) \leq p(|\frac{1}{n} \times \sum_{i=1}^n e_i - Q| > \eta) \leq \frac{\delta}{2}$$

Ostatnia linijka wynika z faktu, że powyższe operacje (wzięcie wartości modułu nie zmniejsza wartości wyrażenia) nie zmniejsza prawdopodobieństwa (zatem nierówność nadal zachodzi).

Uwaga: Jeśli A ma rozkład jednostajny, to:

$$\Pr_E(\Delta, \mathcal{C} \circ A) = \sum_{w \in \mathcal{C}} p(\mathcal{C} \circ A = w) \times p(\Delta(w \oplus \mathcal{E}) \neq w) = \frac{1}{m} \times \sum_{w \in \mathcal{C}} p(\Delta(w \oplus \mathcal{E}) \neq w)$$

gdzie $\mathcal{C}$ przyjmuje wartości $u_i^C, \dots, u_m^C$. $\Pr_E$ nie zależy od A , będziemy więc pisać $\Pr_E(\Delta, \mathcal{C})$.

Do tej pory szacowanie przebiegło po wszystkich kodach $\mathcal{C}$. Kolejne szacowanie będzie inaczej przeprowadzane, mianowicie oszacujemy uśrednione $\Pr_E(\Delta, \mathcal{C})$, gdzie $\mathcal{C}$ przebiega wszystkie n -elementowe kody o rozkładzie jednostajnym.

Wszystkich m -elementowych kodów $\mathcal{C} : A \rightarrow \{0,1\}^n$ jest

$$\binom{2^n}{m} m! = N.$$

$$\begin{aligned} \frac{1}{N} \times \sum_{\mathcal{C}} \Pr_E(\Delta, \mathcal{C}) &\leq \frac{1}{N} \times \sum_{\mathcal{C}} \left(\frac{\delta}{2} + \frac{1}{m} \times \sum_{i=1}^m \sum_{j \neq i} p(\delta(u_j^{\mathcal{C}}, u_i^{\mathcal{C}} + \mathcal{E}) \leq \rho) \right) \\ &= \frac{\delta}{2} + \frac{1}{m} \sum_{i=1}^m \sum_{j \neq i} \underbrace{\frac{1}{N} \sum_{\mathcal{C}} p(\delta(u_j^{\mathcal{C}}, u_i^{\mathcal{C}} + \mathcal{E}) \leq \rho)}_{*} \end{aligned}$$

Teraz zajmiemy się oszacowaniem $*$ dla ustalonej pary indeksów $i \neq j$ (X jest funkcją boolowską; $X(\text{true}) = 1$ i $X(\text{false}) = 0$):

$$\begin{aligned} \frac{1}{N} \sum_{\mathcal{C}} p(\delta(u_j^{\mathcal{C}}, u_i^{\mathcal{C}} + \mathcal{E}) \leq \rho) &= \frac{1}{N} \sum_{\mathcal{C}} p(u_j^{\mathcal{C}} - u_i^{\mathcal{C}} \in S_{\rho}(\mathcal{E})) = \\ \frac{1}{N} \sum_{\mathcal{C}} \sum_{e \in \{0,1\}^n} p(\mathcal{E} = e) \cdot X(u_j^{\mathcal{C}} - u_i^{\mathcal{C}} \in S_{\rho}(e)) &= \\ = \sum_{e \in \{0,1\}^n} p(\mathcal{E} = e) \cdot \frac{1}{N} \sum_{\mathcal{C}} X(\underbrace{u_j^{\mathcal{C}} - u_i^{\mathcal{C}}}_{\in S_{\rho}(e)}) & \end{aligned}$$

każdy taki wektor pojawi się w $\frac{N}{2^n - 1}$ kodów

Zatem:

$$\sum_{\mathcal{C}} X(u_j^{\mathcal{C}} - u_i^{\mathcal{C}} \in S_{\rho}(e)) = \frac{N}{2^n - 1} \cdot |S_{\rho}(e) \setminus \{0^n\}|$$

Więc:

$$\frac{1}{N} \sum_{\mathcal{C}} X(\underbrace{u_j^{\mathcal{C}} - u_i^{\mathcal{C}}}_{\in S_{\rho}(e)}) \leq \frac{|S_{\rho}(e)|}{2^n - 1} = \frac{1}{2^n - 1} \sum_{r \leq \rho} \binom{n}{r}, \text{ niezależnie od } e.$$

Ponieważ $\sum p(\mathcal{E} = e) = 1$, więc:

$$\sum_{e \in \{0,1\}^n} p(\mathcal{E} = e) \cdot \frac{1}{N} \sum_e X(u_j^e - u_i^e \in S_\rho(e)) \leq \frac{1}{2^n - 1} \sum_{r \leq \rho} \binom{n}{r}$$

Kontynuując dowód:

$$\leq \frac{\delta}{2} + \frac{1}{m} \cdot m(m-1) \cdot \frac{1}{2^n - 1} \cdot \sum_{r \leq \rho} \binom{n}{r} \leq \frac{\delta}{2} + \frac{m}{2^n} \sum_{r \leq \rho} \binom{n}{r} \leq$$

Nierówność dla dowolnego $\lambda < \frac{1}{2}$: $\sum_{r \leq \lambda \cdot n} \binom{n}{r} \leq 2^{n \cdot H_2(\lambda)}$ (udowodniono na ćwiczeniach, można też wykazać elementarnym rachunkiem). Stąd dla $\lambda = Q + \eta < \frac{1}{2}$ i $\rho = \lambda n$: $\sum_{r \leq \rho} \binom{n}{r} \leq 2^{n \cdot H_2(\lambda)}$, gdzie H jest binarną entropią funkcji H_2 .

Zatem:

$$\leq \frac{\delta}{2} + \underbrace{\frac{m}{2^n} \cdot 2^{n \cdot H(\lambda)}}_{2^{n(\frac{\log_2 m}{n} + H(\lambda) - 1)}}$$

Aby dokończyć dowód pozostaje dobrać η i m . Musimy pokazać, że mając dane $Q < \frac{1}{2}$ i $\varepsilon > 0$ możemy znaleźć $\eta > 0$ (niezależne od n) i $m < n$ takie, że jeśli n jest wystarczająco duże to:

$$(i) \quad Q + \eta < \frac{1}{2}$$

$$(ii) \quad C_\Gamma - \varepsilon \leq \frac{\log_2 m}{n} < C_\Gamma \quad (\text{gdzie } C_\Gamma = 1 - H(Q))$$

$$(iii) \quad \frac{\log_2 m}{n} - 1 + H(Q + \eta) \leq -\frac{\varepsilon}{3}$$

Bez straty ogólności, ponieważ $C > 0$ możemy założyć, że $C_\Gamma - \varepsilon \geq 0$. Ponieważ $C_\Gamma = 1 - H(Q)$, funkcja H jest ciągła i $Q < \frac{1}{2}$ więc możemy wybrać $\eta > 0$ wystarczająco małe tak, by

$$Q + \eta < \frac{1}{2} \quad \text{i} \quad 1 - H(Q + \eta) \geq C_\Gamma - \frac{\varepsilon}{3}.$$

Niech $n_0' = \left\lceil \frac{3}{\varepsilon} \right\rceil$. Dla każdego $n \geq n_0'$ możemy tak dobrać k , że

$$C_{\Gamma} - \varepsilon \leq \frac{k}{n} \leq C_{\Gamma} - \frac{2}{3} \cdot \varepsilon$$

Wybieramy $M = 2^k$, zatem $\frac{1}{n} \log_2 M = \frac{k}{n}$ (warunek (ii) spełniony). Skoro

$\frac{k}{n} \leq C_{\Gamma}$ i $1 - H(Q + \eta) \geq C_{\Gamma} - \frac{\varepsilon}{3}$ to mamy:

$$\frac{\log_2 M}{n} + H(Q + \eta) - 1 \leq (C_{\Gamma} - \frac{2}{3} \varepsilon) - (C_{\Gamma} - \frac{\varepsilon}{3}) = -\frac{\varepsilon}{3}$$

czyli warunek (iii) z $m = -\frac{\varepsilon}{3} < 0$.

I teraz wynik końcowy:

$$\leq \frac{\delta}{2} + \underbrace{\left(2^{-\frac{\varepsilon}{3}}\right)^n}_{\leq \frac{\delta}{2} \text{ dla dostatecznie dużych } n}$$

□.