Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

O Teorema de Shannon para Codificação Com Ruído

Phellippe Aprigio[1]
Centro de Engenharias e Ciências Exatas, UNIOESTE, Foz do Iguaçu, PR
Luciano Panek[2]
Centro de Engenharias e Ciências Exatas, UNIOESTE, Foz do Iguaçu, PR

Os resultados básicos da *Teoria da Informação* foram estabelecidos em 1948 no trabalho intitulado "*A Mathematical Theory of Communication*" de C. E. Shannon [2]. Shannon fundamentou sua teoria em problemas práticos de engenharia de comunicação: *compactação* (codificação de fonte) e *transmissão de dados* (codificação de canal).

Em nosso trabalho estudamos o *problema da transmissão de dados.* Mais especificamente, dado um canal com ruído, o problema consiste em determinar a máxima taxa de transmissão $\mathcal{C}$ por bit de entrada tal que: se R é menor ou igual a $\mathcal{C}$ e ϵ é um número real positivo, então existe um código de canal com taxa de transmissão de informações R com probabilidade de erro de decodificação menor ou igual a ϵ. O número $\mathcal{C}$ é chamado de *capacidade do canal.* A resposta para o problema é conhecido na literatura como *Teorema de Shannon para Codificação Com Ruído* [1].

Considere um *canal discreto sem memória* $(\mathcal{X}, P(\mathcal{Y} \mid \mathcal{X}), \mathcal{Y})$: $\mathcal{X} = \{x_1, \ldots, x_s\}$ é o *alfabeto de entrada*, $\mathcal{Y} = \{y_1, \ldots, y_t\}$ é o *alfabeto de saída*, $p(y_j \mid x_i)$ são as *probabilidades do canal*, que satisfazem a condição

$$\sum_{j=1}^{t} p(y_j \mid x_i) = 1 \tag{1}$$

para todo $1 \leq i \leq s$, e

$$p(y_{j_1}, \ldots, y_{j_n} \mid x_{i_1}, \ldots, x_{i_n}) = \prod_{k} p(y_{j_k} \mid x_{i_k}) \tag{2}$$

para todo $x_{i_k} \in \mathcal{X}$ e $y_{j_k} \in \mathcal{Y}$.

Fixada uma distribuição de probabilidades $P(\mathcal{X} = x_i) = p(x_i)$ para o alfabeto de entrada $\mathcal{X}$, as probabilidades do canal induzem uma distribuição de probabilidades sobre o alfabeto de saída $\mathcal{Y}$:

$$P(\mathcal{Y} = y_j) = \sum_{i=1}^{s} p(y_j \mid x_i) p(x_i). \tag{3}$$

[1]phellippe58@gmail.com
[2]luciano.panek@unioeste.br

A partir destas distribuições consideramos as entropias do alfabeto de entrada $H(\mathcal{X})$ e condicional $H(\mathcal{X} \mid \mathcal{Y})$:

$$H(\mathcal{X}) = \sum_i p(x_i) \log_2 \frac{1}{p(x_i)}, \tag{4}$$

$$H(\mathcal{X} \mid \mathcal{Y}) = \sum_{i,j} p(x_i \mid y_j) p(y_j) \log_2 \frac{1}{p(x_i \mid y_j)}. \tag{5}$$

Pondo $I(X,Y) = H(X) - H(X \mid Y)$, a *capacidade do canal* é o número

$$\mathcal{C} = \max_{p(x_i)} I(X,Y). \tag{6}$$

Seja C em $\mathcal{X}^n$ um código de canal contendo M palavras-código $\mathbf{c}_1, \ldots, \mathbf{c}_M$. Para um decodificador de máxima verossimilhança, seja P_i a probabilidade de cometermos um erro de decodificação ao transmitirmos $\mathbf{c}_i$. Seja

$$P_C = \frac{1}{M} \sum_i P_i \tag{7}$$

e $P^*(M,n) = \min_C P_C$ o mínimo entre todos os códigos C com parâmetros M e n.

Teorema 0.1 (Teorema de Shannon)**.** *Seja* $\mathcal{C} = \max_{p(x_i)} I(X,Y)$*,* $0 < R < \mathcal{C}$ *e* $M_n = s^{[nR]}$*. Então* $P^*(M_n, n) \to 0$ *se* $n \to +\infty$*. Mais ainda,* $\mathcal{C}$ *é máximo com esta propriedade.*

Para um canal binário simétrico, não é difícil verificar que a probabilidade de erro de decodificação do código de repetição tende a zero quando o seu comprimento tende ao infinito. O problema neste caso é que sua taxa também tende a zero. É neste fato que reside a contribuição do Teorema de Shannon: mesmo fixando a taxa, desde que ela seja menor do que a capacidade do canal, é possível construir códigos com probabilidades de erro de decodificação tão pequenas quanto se queira.

Referências

[1] S. Roman. *Coding and Information Theory.* Graduate Texts in Mathematics, Springer, 1992.

[2] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, 27:379-423 and 623-656, 1948.