Elemente de teoria codurilor

Fie $\mathbb{F}_2 = \{0, 1\}$ un corp cu 2 elemente.

Procesul de codare: transformarea unui bloc $a_1 a_2 \dots a_k$ de k simboluri din $\mathbb{F}_2$ într-un *cuvânt codat* $x = x_1 \cdots x_n \in \mathbb{F}_2^n$, $n \geq k$.

Considerăm cazul $x_i = a_i \ \forall i \in \{1, \dots, k\}$. În aceste condiţii simbolurile $x_{k+1}, \dots, x_n$ se numesc *simboluri de control.*

Observaţia 0.1. Simbolurile de control au rolul de a determina şi corecta eventualele erori de transmisie.

Definiţia 0.2. Prin *cod de tipul* (n, k) înţelegem o funcţie injectivă $\gamma : \mathbb{F}_2^k \to \mathbb{F}_2^n$. Elementele mulţimii $\gamma(\mathbb{F}_2^k)$ s.n. *cuvinte codate.*

Definiţia 0.3. a) Dacă $x = x_1 \dots x_n, \ y = y_1 \dots y_n \in \mathbb{F}_2^n$, atunci cardinalul mulţimii $\{i \in \{1, \dots, n\} \mid x_i \neq y_i\}$ s.n. *distanţa Hamming* între x şi y. Acesta se notează cu $d(x, y)$.

b) Dacă $x = x_1 \dots x_n, \in \mathbb{F}_2^n$, atunci cardinalul mulţimii $\{i \in \{1, \dots, n\} \mid x_i \neq 0\}$ s.n. *norma Hamming* a lui x şi se notează cu $w(x)$.

Teorema 0.4. *Fie γ un cod cu mulţimea cuvintelor codate C.*

a) Codul γ determină existenţa oricărei mulţimi de erori cu cel mult t elemente dacă şi numai dacă

$$\min\{d(x, y) \mid x, y \in C, \ x \neq y\} \geq t + 1.$$

a) Codul γ corectează orice mulţime de erori cu cel mult t elemente dacă şi numai dacă

$$\min\{d(x, y) \mid x, y \in C, \ x \neq y\} \geq 2t + 1.$$

Demonstraţie.

□

DEFINIŢIA 0.5. a) Spunem că un cod $\gamma : \mathbb{F}_2^k \to \mathbb{F}_2^n$ este liniar dacă γ este o aplicaţie liniară de $\mathbb{F}_2$-spaţii vectoriale.

b) Fie γ un cod liniar de tip (n, k), $\mathbf{e}$ baza canonică din $\mathbb{F}_2^k$ şi $\mathbf{e}'$ baza canonică din $\mathbb{F}_2^n$. Matricea $G = [\gamma]_{\mathbf{ee}'} \in \mathcal{M}_{kn}(\mathbb{F}_2)$ s.n. *matricea generatoare a codului* γ sau *matricea de codare*.

OBSERVAŢIA 0.6. $G = [I_k \ P]$ (pe primele k-poziţii avem chiar cuvântul transmis).

TEOREMA 0.7. *Fie $\gamma : \mathbb{F}_2^k \to \mathbb{F}_2^n$ un cod liniar cu matricea generatoare $G = [I_k \ P]$, $P \in \mathcal{M}_{k,n-k}(\mathbb{F}_2)$. Atunci aplicaţia liniată $\nu : \mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$ cu matricea sa în bazele canonice $H = \begin{bmatrix} P \\ I_{n-k} \end{bmatrix}$ are proprietăţile:*

(i) $\mathrm{Ker}(\nu) = \mathrm{Im}(\gamma)$;

(ii) $u \in \mathrm{Im}(\gamma) = C \Leftrightarrow uH = 0$.

DEMONSTRAŢIE.

□

DEFINIŢIA 0.8. Matricea H s.n. *matricea de verificare*.

Decodarea

Fie $\gamma : \mathbb{F}_2^k \to \mathbb{F}_2^n$ un cod şi $C = \mathrm{Im}(\gamma)$ mulţimea cuvintelor codate. Dacă $v \in C$ este cuvântul (codat) trimis şi $u \in \mathbb{F}_2^n$ este cuvântul recepţionat, atunci $e = v - u = v + u$ s.n. *eroarea de tansmisie*.

OBSERVAŢIA 0.9. $e \in u + C$; $u + C$ s.n. *clasa erorilor ataşate lui* u. Elementul $u + c_0$ cu proprietatea $d(u, u + c_0) = \min\{d(u + c) \mid c \in C\}$ s.n. *element principal*.

DEFINIŢIA 0.10. Dacă H este matricea de verificare şi $u \in \mathbb{F}_2^n$, atunci uH s.n. *sindromul* lui u.

TEOREMA 0.11. *Doi vectori din $\mathbb{F}_2^n$ aparţin aceleiaşi clase de erori dacă şi numai dacă ei au acelaşi sindrom.*

Procedeu de decodare:

1) Calculăm sindromul cuvântului recepţionat u;

2) Determinăm elementul principal e din clasa de erori a sindromului găsit (a lui u);

3) Calculăm $u-e$ (este cuvântul codat posibil transmis, cu cea mai mare proabilitate, i.e. numaăr minim de erori);

4) Determinăm cuvântul necodat transmis.

Example 0.12. Cod de tipul $(3,6)$ cu matrcea de codare

$$G = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

$$H = \left[\begin{array}{ccc} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

Observăm că $u \in \mathbb{F}_2^6$ este cuvânt codat $\Leftrightarrow uH = 0$, adică

$$u = u_1 \dots u_6 \in C \Leftrightarrow \begin{cases} u_1 + u_2 + u_4 = 0 \\ u_2 + u_3 + u_5 = 0 \\ u_1 + u_2 + u_3 + u_6 = 0 \end{cases}$$

Din $\min\{d(x,y) \mid x, y \in C\} = 3$ rezultă că *acest cod determină cel mult 2 erori şi repară cel mult o eroare.*

Dacă se recepţionează vectorul $u = 110101 \in \mathbb{F}_2^6$, atunci $uH = 111$.

Se rezolvă sistemul $x_1 \dots x_6 H = 111$ şi găsim

$$\begin{array}{r} x_4 = 1 - x_1 - x_2 = 1 + x_1 + x_2 \\ x_5 = 1 - x_2 - x_3 = 1 + x_2 + x_3 \\ x_6 = 1 - x_1 - x_2 - x_3 = 1 + x_1 + x_2 + x_3 \end{array},$$

de unde găsim

$u+C = \{000111, 001100, 010000, 011001, 100010, 101001, 110101, 111110\}$

De unde rezultă că eroare cea mai probabilă este 010000 şi cuvantul transmis este 100101.