Mały Projekt nr 5.

Temat: *Kody Hamminga*

Powstające w trakcie transmisji błędy mogą drogo kosztować i dlatego ważne są starania o to, aby błędy te były jak najmniej prawdopodobne. Zastosowanie kodowania w procesie przesyłania informacji umożliwia zwiększenie niezawodności przekazu. Kody korekcyjne są jedyną metodą poprawienia wierności transmisji tam, gdzie retransmisja błędnego sygnału jest niemożliwa, np. w łączności satelitarnej. Za ich pomocą można również zabezpieczać dane przechowywane na nośnikach elektronicznych, które z czasem ulegają degradacji. Okazuje się, że można je także z powodzeniem wykorzystywać do tworzenia skutecznych systemów kryptograficznych.

Zasadnicza idea korekcyjnego kodowania nadmiarowego polega na przesyłaniu wraz z oryginalną wiadomością pewnej informacji "nadmiarowej", nie wnoszącej nic do treści samej wiadomości. Odebrana, wydłużona w ten sposób wiadomość, odwzorowywana jest za pomocą przekształcenia dekodującego na ciąg pierwotnej długości.

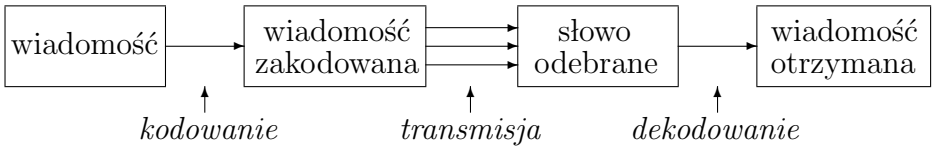

FIGURE 1. Schemat procesu kodowania

Funkcje kodująca i dekodująca powinny być tak określone, aby prawdopodobieństwo odczytania wiadomości z błędem było minimalne. W procesie dekodowania należy błędy wykryć, zlokalizować a następnie poprawić. Analiza informacji zawartych w dodatkowo przesyłanych znakach powinna umożliwić bezbłędne odtworzenie wysłanego słowa, jeśli w czasie transmisji wystąpiło stosunkowo mało błędów.

Powszechnie stosowanymi w praktyce są *kody blokowe*, których słowa kodowe tworzą (dla ustalonego $n \in \mathbb{N}$) niepusty podzbiór $\mathcal{C}$ n-wymiarowej przestrzeni wektorowej nad ciałem $GF(q)$. Mówimy wówczas, że kod $\mathcal{C}$ jest długości n nad ciałem $GF(q)$.

Załóżmy, że przesyłaną informację dzielimy na skończone ciągi zawierające ustaloną liczbę $k \in \mathbb{N}$ tzw. ***symboli informacji***, które mogą być kodowane i dekodowane niezależnie od innych ciągów. W tym przypadku kod $\mathcal{C}$ jest obrazem pewnego różnowartościowego przekształcenia

$$\xi : GF(q)^k \to GF(q)^n,$$

zwanego ***funkcją kodującą*** (*algorytmem kodowania*). Odwzorowanie ξ każdemu wektorowi $\boldsymbol{u} \in GF(q)^k$ długości k, w jednoznaczny sposób przyporządkowuje słowo kodowe $\boldsymbol{c} \in \mathcal{C} = \xi(GF(q)^k)$, w którym pewne $n-k$ symboli dodatkowych to tzw. ***symbole sprawdzające***.

Przekształcenie

$$\eta : GF(q)^n \to GF(q)^k$$

takie, że dla każdego $\boldsymbol{u} \in GF(q)^k$, $\eta(\xi(\boldsymbol{u})) = \boldsymbol{u}$, nazywamy ***funkcją dekodującą*** (*algorytmem dekodowania*).

W przypadku, gdy wysłanym wektorem jest słowo kodowe $\boldsymbol{c} \in \mathcal{C}$, zaś odebranym po transmisji jest wektor $\boldsymbol{f} \in GF(q)^n$, to wektor $\boldsymbol{e} := \boldsymbol{f} - \boldsymbol{c} \in GF(q)^n$ nazywamy ***wektorem błędu***.

Niech $\mathcal{C}$ będzie kodem blokowym długości n nad ciałem $GF(q)$. Jeśli $q = 2$, to $\mathcal{C}$ jest ***kodem binarnym***.

[2] Jeżeli funkcja kodująca jest liniowa to kod $\mathcal{C}$ nazywamy (n,k)-***kodem liniowym*** nad ciałem $GF(q)$, Wówczas $\mathcal{C}$ jest k-wymiarową podprzestrzenią n-wymiarowej przestrzeni $GF(q)^n$.

Przykład 1. Kod powtórzeniowy długości n nad ciałem $GF(q)$ powstaje przez $(n-1)$-krotne powtórzenie pojedynczego symbolu $c \in GF(q)$. Funkcja kodująca $\xi : GF(q) \to GF(q)^n$ ma postać: $c \mapsto c \dots c$.

Przykład 2. Kod kontroli parzystości długości n nad ciałem $GF(q)$ powstaje przez dodanie na końcu każdej wysyłanej wiadomości $c_1 \dots c_{n-1} \in GF(q)^{n-1}$, elementu przeciwnego do sumy $\sum_{i=1}^{n-1} c_i$.

Funkcja kodująca $\xi : GF(q)^{n-1} \to GF(q)^n$ jest wtedy postaci:

$$c_1 \dots c_{n-1} \mapsto c_1 \dots c_{n-1} - \sum_{i=1}^{n-1} c_i.$$

Algorytm kodowania. Dla (n,k)-kodów liniowych warunki na symbole nadmiarowe w słowach kodowych można określić za pomocą odpowiedniego układu równań liniowych.

Macierz $H \in \mathfrak{M}_{n-k}^n(GF(q))$, dla której $\boldsymbol{c} \in \mathcal{C}$ wtedy i tylko wtedy, gdy spełniony jest następujący warunek:

$$H\boldsymbol{c}^T = \mathbf{0}_{n-k}^T \tag{1}$$

będziemy nazywać ***macierzą kontroli parzystości*** (***sprawdzającą***) kodu $\mathcal{C}$. Układ $n-k$ równości (1) nosi nazwę *równości kontroli parzystości.*

Kody Hamminga. Niech $m \in \mathbb{N}$ i niech $H_m \in \mathfrak{M}_m^{2^m-1}(Z_2)$ będzie macierzą, której każda i-ta kolumna jest binarną reprezentacją liczby $1 \le i \le 2^m - 1$ zapisaną "z dołu do góry".

Definicja 3. ***Binarnym kodem Hamminga*** $\mathcal{H}_m$ *rzędu* m *nazywamy* $(2^m-1, 2^m-m-1)$*-kod liniowy, dla którego* H_m *jest macierzą kontroli parzystości.*

Zadania do wykonania:

(a) Skonstruować macierz kontroli parzystości dla kodu $\mathcal{H}_3$. Zakodować wiadomość $\boldsymbol{u} = 1010$. Zakładając, że w trakcie transmisji został popełniony jeden błąd odkodować słowo $\boldsymbol{f} = 0011111$.

(b) Skonstruować macierz kontroli parzystości dla kodu $\mathcal{H}_4$. Zakodować wiadomość $\boldsymbol{u} = 11101110000$. Zakładając, że w trakcie transmisji został popełniony jeden błąd odkodować słowo $\boldsymbol{f} = 111010100000000$.