# Laboratorium ochrony danych

**Ćwiczenie nr 3**

## Temat ćwiczenia: Kod BCH

*Cel dydaktyczny*: Zapoznanie się z metodami detekcji i korekcji błędów transmisyjnych za pomocą binarnych kodów cyklicznych, na przykładzie kodu Bose-Chaudhuri-Hocquenghema.

## Wprowadzenie teoretyczne

### Kody cykliczne

Kody cykliczne są podklasą kodów liniowych i znalazły największe zastosowania praktyczne. Popularność kodów cyklicznych wynika z następujących ich zalet:

- istnieją efektywne algebraiczne metody konstrukcji kodów cyklicznych o wymaganych właściwościach,
- realizacja koderów i dekoderów kodów cyklicznych za pomocą rejestrów przesuwnych ze sprzężeniem zwrotnym jest stosunkowo prosta.

W algebrze kodów cyklicznych ciągi informacyjne i kodowe zapisuje się w postaci wielomianów, a właściwości kodów opisuje się za pomocą pojęć z zakresu pierścieni wielomianów i ciał Galois.

Nazwa kodów cyklicznych pochodzi od właściwości przesunięcia cyklicznego, którą spełniają wektory kodowe. Stąd wywodzi się też definicja kodu cyklicznego.

Kod $(n,k)$ jest kodem cyklicznym, jeśli każdy wektor kodowy

$$\boldsymbol{c} = \left[a_{n-1}, a_{n-2}, \ldots, a_1, a_0\right]$$

po *i*-tym przesunięciu cyklicznym daje wektor

$$\boldsymbol{c}_i = \left[a_{n-1-i}, a_{n-2-i}, \ldots, a_1, a_0, a_{n-1}, a_{n-2}, \ldots, a_{n-i}\right]$$

będący również wektorem kodowym tego kodu.

Wektor kodowy $\boldsymbol{c}$ można zapisać w postaci wielomianu

$$c(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1 x + a_0 x.$$

W teorii blokowych kodów cyklicznych wykorzystuje się pojęcie pierścienia klasy reszt modulo $x^n - 1$. W dalszych rozważaniach ograniczymy się do kodów nad $GF(2)$. Pierścień klas reszt modulo $x^n + 1$ jest pierścieniem wielomianów stopnia nie większego niż $n-1$, które odpowiadają ciągom binarnym o długości *n*.

### Wielomiany generujące kody BCH

Ideałem jest podzbiór wielomianów pierścienia generowany przez pewien wielomian $g(x)$, który jest dzielnikiem $x^n + 1$. Ideał ten stanowi kod, a wielomian $g(x)$ nazywamy *wielomianem generującym kod*. Wielomian $g(x)$ dzieli bez reszty każdy wielomian odpowiadający wektorowi kodowemu. Stopień wielomianu generującego kod określa liczbę elementów kontrolnych wektora kodowego.

Z powyższych rozważań wynika, że wielomianem generującym kod cykliczny może być każdy wielomian, który jest podzielnikiem $x^n + 1$, gdzie $n = q^m - 1$, a *m* jest liczbą naturalną.

Kody Bose-Chaudhuri-Hocquenghema (BCH) należą do kodów korygujących błędy losowe i mają duże znaczenie praktyczne. Zostały one niezależnie skonstruowane przez Hocquenghema w 1959 r. oraz przez Bose z Chaudhurim w 1960 r. Kody BCH swoją popularność zawdzięczają następującym zaletom:

- Istnieją efektywne metody konstruowania kodów BCH o zadanych właściwościach detekcyjnych i korekcyjnych.
- Konstrukcja koderów i dekoderów kodów BCH jest prostsza niż dla innych kodów.

Kody BCH można konstruować nad ciałem binarnym i ciałami rozszerzonymi. Największe znaczenie mają kody binarne. Udowodniono, że dla każdej liczby całkowitej $m$ i $t < 2^{m-1}$ istnieje kod BCH o długości $n = 2^m - 1$. Może on korygować do $t$ błędów i ma nie więcej niż $mt$ elementów kontrolnych. Kody te mają następujące parametry:

- długość wektora kodowego $n = 2^m - 1$,
- liczba pozycji kontrolnych $n - k \le mt$,
- odległość minimalna $d \ge 2t + 1$.

Wielomiany generujące kody BCH wyznacza się w następujący sposób. Niech $\alpha$ będzie elementem pierwotnym ciała $GF(2^m)$. Zbiór $\{f(x)\}$ jest zbiorem ciągów kodowych kodu BCH, jeśli pierwiastkami dowolnie wybranego wielomianu $f(x)$ są elementy ciała

$$\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{2t}.$$

Każdy element ciała o parzystym wykładniku ma w tej sekwencji taką samą funkcję minimalną jak któryś z poprzedzających go elementów o wykładniku nieparzystym. Na przykład $\alpha^2$ i $\alpha^4$ są pierwiastkami $m_1(x)$, $\alpha^6$ jest pierwiastkiem $m_3(x)$ itd. Uwzględniając ten fakt podczas wyznaczania wielomianu generującego kod BCH, wystarczy wziąć pod uwagę elementy ciała z wykładnikami nieparzystymi.

Wielomian generujący kod BCH o zdolności korekcyjnej $t$ jest najmniejszą wspólną wielokrotnością funkcji minimalnych $m_1(x), m_3(x), \ldots, m_{2t-1}(x)$

$$g(x) = NWW\left(m_1(x), m_3(x), \ldots, m_{2t-1}(x)\right).$$

P r z y k ł a d

Wyznaczanie wielomianów generujących kody BCH.

Dla $m = 4$ dwumian $x^{q^m - 1} - 1$ ma następujący rozkład

$$x^{15} + 1 = (x+1)\left(x^4 + x + 1\right)\left(x^4 + x^3 + x^2 + x + 1\right)\left(x^2 + x + 1\right)\left(x^4 + x^3 + 1\right).$$

Wielomiany nierozkładalne z prawej strony znaku równości są wielomianami minimalnymi elementów ciała $GF(2^4)$. Podstawiając symbole wielomianów minimalnych, otrzymamy

$$x^{15} + 1 = m_0(x)\, m_1(x)\, m_3(x)\, m_5(x)\, m_7(x).$$

Korzystając z tego wyrażenia, dla zadanych wartości $t$ można wyznaczyć wielomiany generujące kody BCH.

| | | |
|---|---|---|
| $t = 1,$ | $g(x) = m_1(x),$ | kod Hamminga (15,11); |
| $t = 2,$ | $g(x) = m_1(x)\, m_3(x),$ | kod (15,7); |
| $t = 3,$ | $g(x) = m_1(x)\, m_3(x)\, m_5(x),$ | kod (15,5). |

Po prawej stronie wielomianów generujących podano parametry kodów $(n,k)$. Na przykład dla $t = 2,$ $(n,k) = (15,7)$. Aby obliczyć liczbę pozycji informacyjnych $k$ wektora kodowe-

go, należy wyznaczyć stopień wielomianu generującego. Dla kodu $(n,k)=(15,7)$ otrzymamy wielomian generujący ósmego stopnia

$$g(x)=\left(x^4+x+1\right)\left(x^4+x^3+x^2+x+1\right)=x^8+x^7+x^6+x^4+1.$$

Wektor kodowy będzie zatem zawierał osiem pozycji kontrolnych i siedem informacyjnych. Odległość minimalna tego kodu jest $d\geq 5$ i może on korygować dwa błędy.

Dla $t=1$ kod BCH ma wielomian generujący

$$g(x)=m_1(x)=x^4+x+1\,.$$

Kod BCH korygujący jeden błąd jest jednocześnie kodem Hamminga. Generalnie kody Hamminga są podzbiorem kodów BCH.

## Algorytm kodowania

Do kodowanie informacji za pomocą kodów cyklicznych można wykorzystać wielomian generujący kod lub macierz generującą kod. Tutaj rozważymy kodowanie za pomocą wielomianu generującego.

Wektor kodowy cyklicznego kodu systematycznego ma formę:

$$\boldsymbol{c}_X=\left[m_{n-1},\ldots,m_{n-k},r_{n-k-1},\ldots,r_0\right],$$

gdzie współrzędne $m_i$ są elementami informacyjnymi, a współrzędne $r_i$ – elementami kontrolnymi.

Gdy mamy wielomian generujący $g(x)$ stopnia $n-k$, to aby obliczyć wektor kodowy systematycznego kodu cyklicznego $(n,k)$, należy wykonać następujące czynności:

1. Wielomian odpowiadający informacji $m(x)$ pomnożyć przez $x^{n-k}$

$$x^{n-k}m(x).$$

2. Otrzymany iloczyn $x^{n-k}m(x)$ podzielić przez wielomian generujący kod $g(x)$ i wyznaczyć resztę $r(x)$ z tego dzielenia

$$x^{n-k}m(x)=q(x)\,g(x)+r(x).$$

3. Obliczyć wielomian $c_X(x)$, odpowiadający wektorowi kodowemu, dodając $x^{n-k}m(x)$ i resztę $r(x)$

$$c_X(x)=x^{n-k}m_N(x)+r(x).$$

Napiszmy ciąg informacyjny w postaci wielomianu

$$m(x)=m_{k-1}x^{k-1}+m_{k-2}x^{k-2}+\ldots+m_1x+m_0\,.$$

Pomnożenie tego wielomianu przez $x^{n-k}$ jest równoważne z przesunięciem wektora kodowego w lewo o $n-k$ pozycji

$$m(x)\,x^{n-k}=m_{k-1}x^{n-1}+m_{k-2}x^{n-2}+\ldots+m_1x^{n-k+1}+m_0x^{n-k}\,.$$

Dzielenie uzyskanego wyrażenia przez $g(x)$ można zapisać w formie algorytmu Euklidesa

$$m(x)\,x^{n-k}=q(x)\,g(x)+r(x),$$

gdzie $q(x)$ jest częścią całkowitą, a $r(x)$ resztą z dzielenia w postaci

$$r(x)=r_{n-k-1}x^{n-k-1}+\ldots r_1x+r_0\,.$$

Wzór umożliwiający obliczenia części kontrolnej wektora kodowego możemy również zapisać w formie kongruencji

$$\sum_{i=1}^{k} m_{n-i} x^{n-i} \equiv \sum_{i=1}^{n-k} r_{n-k-i} x^{n-k-i} \left(\mathrm{mod}\ g(x)\right).$$

Wielomian z lewej strony kongruencji odpowiada części informacyjnej wektora kodowego, a wielomian z prawej strony – części kontrolnej wektora kodowego, która jest równa reszcie z dzielenia wielomianu informacyjnego przez wielomian generujący kod $g(x)$.

## Uproszczony algorytm dekodowania

W czasie transmisji wektorów kodowych kanałem transmisyjnym powstają błędy transmisyjne. Zadaniem dekodera jest wykrycie lub wykrycie i usunięcie tych błędów.

Każdy kod cykliczny ma swój algorytm dekodowania, który pozwala skorygować wszystkie błędy korygowalne przez dany kod. W praktyce często używa się algorytmu uproszczonego, wspólnego dla wszystkich kodów cyklicznych. Algorytm ten umożliwia wykrycie i korekcję wszystkich błędów znajdujących się na $n-k$ pozycjach wektora kodowego. Algorytm ten omówimy szczegółowo.

W procesie dekodowania oblicza się syndrom wektora odebranego. Dla kodów cyklicznych syndrom oblicza się, dzieląc wielomian $c_Y(x)$, odpowiadający wektorowi odebranemu $\boldsymbol{c_Y}$, przez wielomian generujący kod $g(x)$. Syndrom $s(x)$ jest równy reszcie z tego dzielenia

$$c_Y(x) = q(x)\, g(x) + s(x).$$

Syndrom $s(x)$ jest wielomianem stopnia $\le n-k-1$. Jeśli syndrom ma wartość zerową, oznacza to, że wektor odebrany jest wektorem kodowym i w czasie transmisji nie wystąpiły żadne błędy wykrywalne przez kod. Niezerowa wartość syndromu świadczy o tym, że odebrany wektor nie jest wektorem kodowym i zostały wykryte błędy transmisyjne.

Wektor odebrany $\boldsymbol{c_Y}$, jest sumą wektora nadanego $\boldsymbol{c_X}$ i wektora błędów $\boldsymbol{e}$. Wzór ten zapisujemy w postaci wielomianów

$$c_Y(x) = c_X(x) + e(x).$$

Wielomian odpowiadający wektorowi kodowemu $\boldsymbol{c_X}$ dzieli się bez reszty przez wielomian generujący kod $g(x)$, można zatem napisać

$$c_X(x) = m(x)\, g(x).$$

Podstawiając tę zależność do wzoru poprzedniego, otrzymamy

$$c_Y(x) = m(x)\, g(x) + e(x).$$

Porównujemy prawą stronę tego wzoru z prawą stroną wzoru pierwszego. Po przekształceniach mamy

$$e(x) = \left(m(x) + q(x)\right) g(x) + s(x).$$

Syndrom jest resztą z dzielenia wielomianu odpowiadającego wektorowi błędów $e(x)$ przez wielomian generujący kod $g(x)$. Syndrom zawiera informację o położeniu błędów transmisyjnych, co jest wykorzystywane w trakcie korekcji błędów.

Schemat blokowy algorytmu dekodowania z korekcją błędów pokazano na rysunku. Na rysunku tym zastosowano takie same oznaczenia jak w opisie algorytmu. Proces dekodowania ma następujący przebieg.

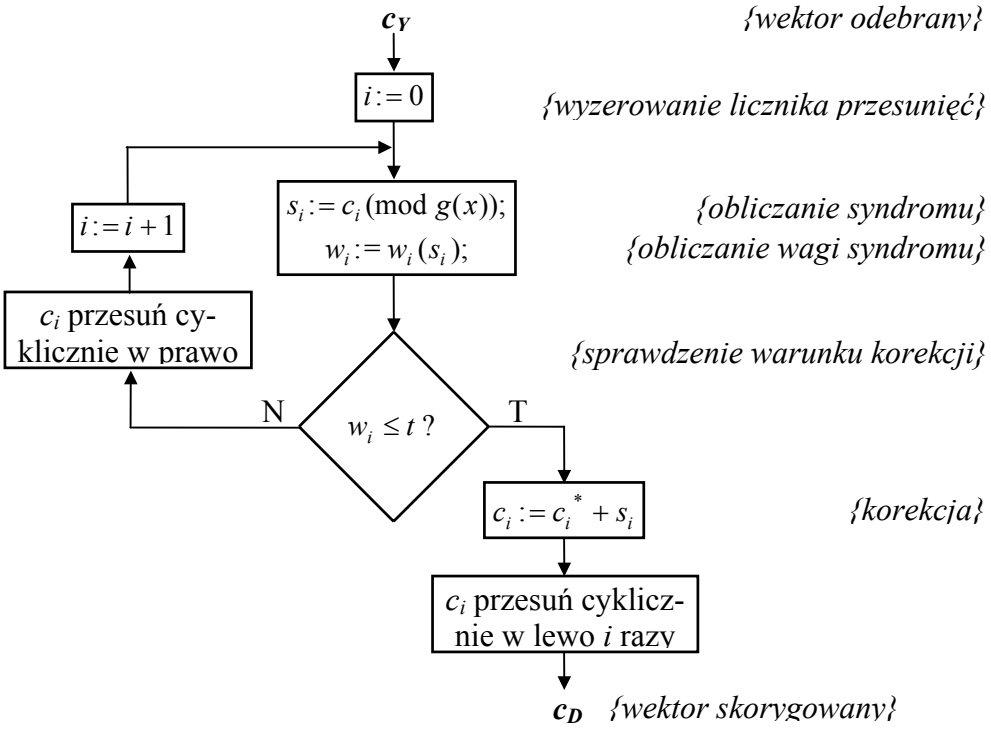
Po wyznaczeniu syndromu oblicza się jego wagę Hamminga $w(s)$. Mogą wówczas wystąpić następujące przypadki:

1. Waga syndromu jest mniejsza lub równa zdolności korekcyjnej kodu, $w(s) \le t$. Oznacza to, że błędy są położone w części kontrolnej wektora kodowego (poprawną część informacyjną można otrzymać bezpośrednio z wektora odebranego bez korygowania). Wektor odebrany

$\boldsymbol{c_Y}$ może być wtedy skorygowany dzięki dodaniu syndromu do wektora odebranego. W wyniku tego działania otrzymamy wektor wyjściowy dekodera $\boldsymbol{c_D}$

$$\boldsymbol{c}_D = \boldsymbol{c}_Y + \boldsymbol{s}.$$

Na podstawie tego wektora można wyznaczyć informację odebraną $m^*$. Będzie ona równa części informacyjnej wektora $\boldsymbol{c}_D$.




Schemat blokowy uproszczonego algorytmu dekodowania

2. Waga syndromu jest większa od zdolności korekcyjnej kodu, $w(s) > t$. Przypadek ten oznacza, że błędy obejmują część informacyjną wektora kodowego. Należy wówczas przesunąć cyklicznie wektor odebrany tak, aby błędy znalazły się w części kontrolnej, a potem go skorygować. W tym celu wykonujemy następujące czynności. Przesuwamy wektor odebrany cyklicznie o jedną pozycję w dowolnym kierunku (np. w prawo), obliczamy syndrom i jego wagę oraz sprawdzamy, czy został spełniony warunek podany w p. 1, czy też warunek podany w p. 2.

- Jeżeli $w(s) \le t$, należy skorygować wektor odebrany zgodnie z p. 1, a następnie przesunąć go cyklicznie w odwrotną stronę (w lewo), aby odtworzyć jego pierwotną postać.
- Jeżeli $w(s) > t$, trzeba ponownie przesuwać cyklicznie wektor odebrany w tę samą stronę, obliczając po każdym przesunięciu syndrom i jego wagę aż do momentu, kiedy $w(s) \le t$. Wtedy należy skorygować wektor odebrany i przesunąć go w odwrotną stronę o taką samą liczbę pozycji.

Algorytm kodowania i obliczania syndromu wykorzystują dzielenie wielomianów. W realizacji programowej tych algorytmów można zastosować symulacje układu do dzielenia wielomianów.

## Opis oprogramowania

Dla zademonstrowania właściwości kodu BCH opracowano program komputerowy w języku Pascal. Składa się on z następujących części:

- program główny BCHCODE.PAS, realizujący kod BCH i udostępniony w postaci źródłowej;
- moduł biblioteczny BCHUNIT.PAS, zawierający procedurę obliczania syndromu i dostępny w postaci skompilowanej. Fragmenty tego programu, bez procedury obliczania syndromu, podano w postaci źródłowej.

Analogiczny program BCH.CPP zrealizowano w języku C++.

Dodatkowo w zbiorze BCCGPOL.TXT podano zestaw wielomianów generujących kody cykliczne (w tym kody BCH). Kolejne kolumny zawierają:

n – długość wektora kodowego,
k – liczba symboli informacyjnych,
r – liczba symboli kontrolnych,
d – minimalna odległość Hamminga,
g(x) – wielomian generujący kod, zapisany w systemie ósemkowym.

## Przebieg ćwiczenia

1. Uruchomić program BCHCODE (można skompilować BCHCODE.PAS w systemie BP7.0), lub BCH, i prześledzić jego działanie. Zwrócić uwagę na możliwości korekcyjne kodu.
2. W programie BCHCODE.PAS do obliczania syndromu używa się procedury Calculate_Syndrom zdefiniowanej w module BCHUNIT.TPU. Napisać własną procedurę obliczania syndromu wektora kodowego i dołączyć ją do programu w miejsce modułu BCHUNIT.TPU. Podobnie w przypadku korzystania z programu BCH.CPP dopisać kod funkcji umożliwiającej obliczanie syndromu. Uruchomić program.
3. W zbiorze BCCGPOL.TXT znaleźć wielomian generujący kod BCH (7,4). Zmodyfikować program tak aby realizował ten kod.
4. W zbiorze BCCGPOL.TXT znaleźć wielomian generujący kod cykliczny (51,24). Zmodyfikować program tak aby realizował ten kod.