Semerenko Vasyl Petrovych, Vinnytsia National Technical University, UKRAINE
Ph.D, lecturer, Department of Computer Technique,
E-mail: vpsemerenko@ukr.net

Decoding of Cyclic Codes for the Erasure Channel

Семеренко Василь Петрович, Вінницький національний технічний університет,
к.т.н., доцент, кафедра обчислювальної техніки
E-mail: vpsemerenko@ukr.net

Декодування циклічних кодів для каналів зв'язку зі стиранням

1. Вступ

По традиції завадостійке кодування в системах передачі даних орієнтовано на ідеалізовану двійково-симетричну модель каналу та інверсну модель помилок, згідно якої одиничні символи в кодових словах можуть замінюватись лише нульовими символами, або навпаки. Однак, спотворення в каналах безпровідного мобільного зв'язку, іоносферних та тропосферних каналах викликають також стирання окремих розрядів в повідомленнях, випадіння символів та інші нестандартні помилки [1].

Відомі методи декодування інверсних помилок непридатні для виправлення стирань. Звичайно, декодувати стирання при відсутності інших типів помилок можна і простим перебором варіантів і вгадуванням стертих символів. Однак, такий спосіб вимагає багато часу для декодування.

Тому актуальною є створення більш широкої математичної моделі помилок і розробки відповідних методів декодування. Розглянемо можливі способи вирішення цієї задачі для класу циклічних кодів.

2. Математичні основи декодування стирань

Розряди кодового слова Z, який формується кодером на боці джерела даних, можуть приймати значення з множини $M=\{0,1\}$. Розряди кодового слова Z_{erase}, який поступає з виходу демодулятора на боці приймача даних, можуть приймати значення з множини $M_x=\{0,1,x,\bar{x}\}$. Символи x та $\bar{x}$ будуть використовуватись для позначення стирань в кодовому слові Z_{erase}, в тих випадках, коли неможливо точно встановити значення прийнятого символу.

Для кожної пари елементів з множини M_x визначимо операцію додавання (табл. 1) и операцію множення (табл. 2). Оскільки кожний стертий символ повинен бути декодований окремо, тому будемо також використовувати індексацію для кожного стертого символу.

Таблиця 1. – Операція додавання (+)

	0	**1**	x	$\bar{x}$
0	0	1	x	$\bar{x}$
1	1	0	$\bar{x}$	x
x	x	$\bar{x}$	0	1
$\bar{x}$	$\bar{x}$	x	1	0

Таблиця 2. – Операція множення $(\times)$

	0	**1**	x	$\bar{x}$
0	0	0	0	0
1	0	1	x	$\bar{x}$
x	0	x	x	0
$\bar{x}$	0	$\bar{x}$	0	$\bar{x}$

Множина елементів M з визначеними над ними операціями додавання по модулю 2 и множення по модулю 2 утворює традиційне поле Галуа $GF(2)$. Безпосередньою перевіркою можна показати, що алгебраїчна структура для множини елементів M_x із зазначеними операціями додавання та множення відповідає всім ознакам комутативного кільця R [2]. Введемо автоматну модель циклічних кодів над комутативним кільцем R на основі теорії лінійних послідовнісних схем (ЛПС) [3].

ЛПС Λ з l входами, m виходами і r елементами пам'яті є кінцевим автоматом лінійного типу, який над комутативним кільцем R задається функцією станів (переходів)

$$S(t+1) = A \times S(t) + B \times U(t),$$

(1)

і функцією виходів

$$Y(t) = C \times S(t) + D \times U(t),$$

(2)

де t – дискретний час,

$A = \|a_{ij}\|_{r \times r}$, $B = \|b_{ij}\|_{r \times l}$, $C = \|c_{ij}\|_{m \times r}$, $D = \|d_{ij}\|_{m \times l}$ – характеристичні матриці ЛПС,

$S(t) = \|s_i\|_r$, $U(t) = \|u_i\|_l$, $Y(t) = \|y_i\|_m$ – слова стану, вхідне і вихідне.

Елементи матриць A, B, C, D в формулах (1) і (2) приймають значення з множини M, а елементи слів $S(t), U(t), Y(t)$ – з множини M_x. Як і для полів Галуа, вибір характеристичних матриць A і B в даному випадку також визначається вимогою r-керованості ЛПС, тобто можливості переходу з будь-якого стану $S(i)$ в стан $S(j)$ не більше, ніж за r тактів роботи автомата. Зазвичай використовують такі характеристичні матриці ЛПС:

$$A=\begin{vmatrix} 0 & 0 & 0 & \dots & g_0 \\ 1 & 0 & 0 & \dots & g_1 \\ 0 & 1 & 0 & \dots & g_2 \\ \dots & \dots & \dots & \cdots & \dots \\ 0 & 0 & 0 & 1 & g_{r-1} \end{vmatrix}, \quad B=\begin{vmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{vmatrix}, \quad C=\begin{vmatrix} 0 & 0 & \dots & 0 & 1 \end{vmatrix}, \quad D=\begin{vmatrix} 0 \end{vmatrix} . \tag{3}$$

Елементи останнього стовпця матриці A в (3) являють собою коефіцієнти породжувального багаточлена циклічного кода:

$$g(x)=g_0+g_1x+g_2x^2+\ldots+g_{r-1}x^{x-1}+g_rx^r .$$

3. Декодування випадкових стирань

В задачу декодування циклічного (n,k)-кода входять виявлення та виправлення помилок двох типів: інверсій и стирань. Будемо розглядати інверсії і випадкові стирання значень не більше, ніж в r $(r=n-k)$ розрядах кодового слова. Виконання задачі декодування складається з двох етапів: визначення типу виявленої помилки в отриманому від демодулятора кодовому слові та виправлення цієї помилки.

З використанням математичного апарату ЛПС сенс першого етапу декодування полягає в наступному. При подачі на вхід ЛПС кодового слова Z, в якому відсутні помилки, відбудеться перехід ЛПС з початкового нульового стану $S(0)$ знову в стан $S(0)$. При наявності стирань в кодовому слові Z_{erase}, ЛПС перейде в деякий ненульовий стан $S_{erase}(n)$, який будемо йменувати “синдромом стирань”.

З врахуванням можливості наявності в кодовому слові $S_{erase}(n)$ помилок різного типу, можна отримати два варіанти синдрому помилок:

а) Синдром стирань $S_{erase}(n)$, який містить хоча б в одному розряді невизначене значення x або $\bar{x}$, а в інших розрядах нулі – відповідає наявності в кодовому слові Z_{erase} тільки стирань;

б) Синдром стирань $S_{erase}(n)$, який містить хоча б в одному розряді одиницю, і хоча б в одному розряді невизначене значення x або $\bar{x}$ – відповідає наявності в кодовому слові Z_{erase} як помилок типу інверсії, так и стирань.

Розглянемо загальний випадок, коли в кодовому слові Z_{erase} можливі m помилок типу стирань. Для їх взаємного розрізнення введемо розширену множину M_{ext} = {0,1, x_1, $\bar{x}_1$, x_2, $\bar{x}_2$, … , x_m, $\bar{x}_m$ }. Будемо йменувати елемент x_j ($\bar{x}_j$) “стертим”, якщо він відповідає стертому j-му розряду слова Z_{erase} $(j=1\div m)$. Для виконання перетворень над елементами множини M_{ext} не будемо вводити спеціальні операції, а лише визначимо відповідність між елементами множин M_{ext} і M_x (табл. 3).

Якщо виникає необхідність виконання операцій між двома стертими елементами $\tilde{x}_l$ і $\tilde{x}_h$ з множини M_{ext}, тоді перехід до елементів множини M_x не здійснюється, а результатом такої операції буде вираз $\tilde{x}_i+\tilde{x}_j$ або $\tilde{x}_i \times \tilde{x}_j$ (символ ~ позначає як наявність, так і відсутність інверсії).

Таблиця 3 – Відповідність між елементами множин M_{ext} і M_x

Елементи множини M_{ext}	**Елементи множини** M_x
0	0
1	1
x_j, $(j = 1 \div m)$	x
$\bar{x}_j$, $(j = 1 \div m)$	$\bar{x}$

Розглянутий нижче алгоритм декодування передбачає спочатку виправлення стирань, а потім, при потребі, знаходженні і виправлення інверсних помилок.

АЛГОРИТМ 1.

Початкові дані:

– довжина n коду і характеристичні матриці ЛПС,

– кодове слово Z_{erase} зі стираннями та інверсними помилками.

1. Обчислити синдром стирань $S_{erase}(n)$ над комутативним кільцем R.

1.1 Присвоїти $S(0) = 0$ (0 -нульове слово).

1.2 Для i от 0 до $n-1$ виконати наступне:

$$S(i+1) = A \times S(i) + B \times z(i), \qquad z(i) \in Z_{erase}.$$

1.3 Присвоїти $S_{erase}(n) = S(n)$.

2. Скласти систему з m ($m \le r$) логічних рівнянь над комутативним кільцем R:

$$\begin{cases} s_1^n = \gamma_1 \\ s_2^n = \gamma_2 \\ \quad \dots \\ s_m^n = \gamma_m \end{cases}, \qquad (4)$$

де s_j^n – j-й розряд синдрому $S_{erase}(n)$ із символами x або $\bar{x}$, $j = 1 \div m$.

3. Розв'язати над комутативним кільцем R систему рівнянь (4) відносно невідомих значень стертих розрядів кодового слова Z_{erase} при наступних варіантах.

3.1 Якщо існує розв'язок системи рівнянь (4) при нульовому варіанті розрядів синдрому $S_{erase}(n)$ ($\gamma_j = 0$ для всіх $j = 1 \div m$), тоді перейти до п. 4.

3.2 Якщо існує розв'язок системи рівнянь (4) при ненульових варіантах розрядів синдрому $S_{erase}(n)$ ($\gamma_i = 0$ і $\gamma_j = 1$, $i \ne j$), тоді перейти до п. 5.

4. В слові Z_{erase} існують лише стирання. Здійснити підстановку знайдених значень стертих розрядів у слові Z_{erase} і отримати правильне кодове слово Z. Перейти до п. 6.

5. В слові Z_{erase} існують також інверсні помилки. Знайти їх за допомогою алгоритмів пошуку інверсних помилок [4].

5.1 Вибрати підстановку значень s_j^n ($j = 1 \div m$) стертих розрядів синдрому $S_{erase}(n)$ символами 0 або 1 і отримати синдром інверсних τ помилок $S_{err}^{(\tau)}(n)$.

5.2. Для синдрому $S_{err}^{(\tau)}(n)$ обчислити над полем Галуа $GF(2)$ слово інверсної помилки $\mathbf{E}_{err}^{(\tau)}$ з мінімальною кратністю τ інверсних помилок. Виправити кодове слово Z_{erase} відповідно зі знайденими помилками типу інверсій та стирань.

6. Кінець.

При m стертих розрядах може знадобитись 2^m варіантів виконання п. 5 Алгоритму 1 при знаходженні інверсних помилок. Тому можна обмежитись лише пошуком одиничних інверсних помилок при наявності до $(r-1)$ стирань.

4 Декодування пакетів стирань

Можливі два типи циклічних пакетів стирань: розріджені й суцільні.

Циклічний розріджений пакет стирань представляє собою набір з кількох стирань, але зосереджених в циклічному інтервалі довжини b. Декодування такого пакета стирань нічим не відрізняється від декодування окремих стирань, тому далі він не розглядається.

Циклічний суцільний пакет стирань довжини b з початком в позиції v ($v=1\div n$), складається тільки з b стертих символів. Позначимо кодове слово з пакетом суцільних стирань як $Z_{ers}^{b,v}$. Такому пакету стирань буде відповідати слово помилок:

$$\mathbf{E}_{ers}^{b,v}=Z+Z_{ers}^{b,v}.$$

Як і у випадку з пакетами інверсних помилок, пакети стирань також дозволяють збільшити кратність виправлених стирань в порівнянні з окремими стираннями по всій довжині кодового слова. При цьому довжина максимального виправленого пакета стирань буде вдвічі більшою максимального виправленого пакета інверсних помилок.

В багатьох роботах, зокрема в [5], доведено, що двійкові циклічні (n,k)-коди і MDS-коди (наприклад, коди Ріда-Соломона) дозволяють виправляти пакети стирань довжини $(n-k)$. Задачами, на яких доцільно зосередити увагу, є розробка алгоритмів мінімальної складності для декодування стирань в поєднанні з іншими типами помилок

Як вже відзначалось, для виконання дій з символами стирань необхідно використати комутативне кільце R, що ускладнює обчислення. Однак цю проблему легко обійти, якщо символи стирань позначити якимось одним символом поля Галуа $GF(2)$, наприклад, нулем. В результаті перейдемо від слова $Z_{ers}^{b,v}$ з пакетом стирань та інверсними помилками до кодового слова $Z_{sol}^{b,v}$ з суцільним пакетом інверсних помилок довжини ($b\le(n-k)$) [6].

ТЕОРЕМА. Якщо в циклічному (n,k)-коді суцільний пакет стирань довжини b починається з розряду v ($v=1\div n$), тоді після подачі на вхід ЛПС, яка знаходиться в нульовому початковому стані $S(0)$, кодового слова $Z_{sol}^{b,v}$, а потім нульового слова довжини ϑ ($\vartheta=(b+v-1)\bmod n$), значення b старших розрядів слова стану $S(n+\vartheta)$ ЛПС будуть дорівнювати b стертим позиціям кодового слова $Z_{ers}^{b,v}$.

Доведення. Після подачі на вхід ЛПС, яка знаходиться в нульовому початковому стані $S(0)$, кодового слова Z без помилок, знову отримаємо нульовий стан ЛПС: $S(n)=S(0)$.

Після заміни стертих позицій кодового слова $Z_{ers}^{b,\nu}$ нулями, отримаємо кодове слово $Z_{sol}^{b,\nu}$, в якому частина розрядів буде помилковою. При наявності інверсних помилок в останніх r розрядах кодового слова $Z_{sol}^{b,\nu}$ їх значення будуть міститись в розрядах слова стану $S(n)$, тобто синдрому інверсної помилки $S_{err}^{(\tau)}(n)$, і тому вони легко піддаються виправленню.

Для виправлення довільного пакету помилок довжини b, його необхідно циклічно зсунути в сторону молодших розрядів (вліво) на ϑ позицій таким чином, щоб вони розмістились в останніх r (контрольних) розрядах кодового слова. Це еквівалентно наступній подачі на вхід ЛПС нульового слова довжини ϑ ($\vartheta=(b+\nu-1) \bmod n$), аналітично це відповідає множенню синдрома $S_{err}^{(\tau)}(n)$ на характеристичну матрицю A^{ϑ}. Виправлення пакета інверсних помилок довжини b в слові $Z_{sol}^{b,\nu}$ дозволяє виправити і пакет стирань такої ж довжини в слові $Z_{ers}^{b,\nu}$.

$\perp$

Розглянемо алгоритм виправлення пакетів стирань.

АЛГОРИТМ 2.

Початкові дані:

– довжина n коду і характеристичні матриці ЛПС,

– кодове слово $Z_{ers}^{b,\nu}$ з циклічним пакетом стирань.

1. Отримати слово $Z_{ers}^{r,\nu}$, в якому довжина суцільного пакета стирань доповнена до r.
2. Зсунути слово $Z_{ers}^{r,\nu}$ на ϑ розрядів в сторону молодших розрядів (вліво) так, щоб стерті розряди займали всі контрольні розряди (r старших розрядів).
3. Перейти від слова $Z_{ers}^{r,\nu}$ до кодового слова $Z_{sol}^{r,\nu}$ заміною символів стирань нулями.
4. Обчислити синдром помилки $S_{err}^{(\tau)}(n)$:
   4.1 Присвоїти $S(0)=0$.
   4.2 Для i от 0 до $n-1$ виконати наступне:
   $$S(i+1)=A\times S(i)+B\times z(i), \quad GF(2), \qquad \text{де } z(i)\in Z_{sol}^{r,\nu}.$$
   4.3 Присвоїти $S_{err}^{(\tau)}(n)=S(n)$.
5. Отримати слово Z'' заміною в слові $Z_{ers}^{r,\nu}$ всіх стертих розрядів значеннями синдрому $S_{err}^{(\tau)}(n)$.
6. Отримати виправлене кодове слово Z циклічним зсувом на ϑ розрядів в сторону старших розрядів (вправо) кодового слова Z''.
7. Кінець.

Зі всіх типів помилок найбільш просто декодуються и виправляються суцільні пакети стирань. Якщо кількість виправлених окремих стирань, як і інверсних помилок, обмежена величиною мінімальної кодової відстані $d_{\min}$, то довжина виправленого пакету стирань обмежена тільки значенням r. Тому в розріджених пакетах стирань доцільно всі розряди помічати як стерті і розглядати їх як суцільний пакет стирань.

5 Висновки

Особливості функціонування реальних каналів передачі даних такі, що на помилки часто групуються в пакети, а також трапляються і нестандартні спотворення: стирання окремих розрядів, випадіння символів тощо. Традиційна модель випадкових помилок в такій ситуації є неадекватною, тому запропонована нова модель з врахуванням особливостей різних типів помилок.

Введена автоматна модель циклічних (n, k)-кодів над комутативним кільцем R на основі теорії лінійних послідовнісних схем (ЛПС). Розглянуті алгоритми лінійної складності для виправлення різних типів помилок: випадкових стирань в поєднанні з інверсними помилками та суцільних пакетів стирань довжини b $(b \le (n-k))$. При наявності в кодовому слові лише стирань відсутня процедура перебору всіх можливих варіантів помилок.

Література

1. Sklar B. Digital Communications: Fundamentals and Applications, (2nd Edition) / – Prentice Hall PTR, 2001. – 1070 p.
2. Фрид Э. Элементарное введение в абстрактную алгебру: Пер. с венг. / – М.: Мир, 1979. – 260 с.
3. Gill A. Linear Sequential Circuits. Analysis, Synthesis and Application / – McGraw-Hill Book Company, New York, London, 1967.
4. Семеренко В. П. Высокопроизводительные алгоритмы для исправления независимых ошибок в циклических кодах // Системи обробки інформації: збірник наук. праць – Харків: ХУПС, 2010. – Вип. 3(84). – С. 80-89.
5. Fossorier M. Universal burst error correction // Proc. IEEE Int. Symp. Information Theory – Seatle, WA, Jul. 2006. – pp. 1969-1973.
6. Semerenko V. P. Burst-Error Correction for Cyclic Codes. Proceeding of International IEEE Conference EUROCON2009, S. Petersburg, Russia – pp.1646-1651.