УДК 681.32

КРИТЕРІЇ ВИБОРУ ПОРОДЖУВАЛЬНИХ ПОЛІНОМІВ ДЛЯ КОДІВ CRC

Семеренко Василь

Вінницький національний технічний університет

Анотація

Розглянуто теоретичний базис кодів CRC (Cyclic Redundancy Code) за допомогою математичного апарату лінійних послідовнісних схем (ЛПС). Показано властивості кодів CRC з позицій їх автоматно-графової моделі і дані рекомендації по їх вибору.

Abstract

The theoretical basis of CRC codes (Cyclic Redundancy Code) with the help of the theory of linear finite-state machine (LFSM) is considered. The properties of CRC codes based on its automatongraphical model are analysed and the recommendations about the choice of the generator polynomials for the CRC are done.

Вступ

Найпопулярнішим сучасним методом контролю передачі даних в комп'ютерних мережах та перевірки цілісності збереження даних на різних носіях є метод CRC [1]. Ці два основних напрямки використання CRC відображаються в двох розшифровках його абревіатури: *Cyclic Redundancy Code* – циклічний надлишковий код, і *Cyclic Redundancy Check* – циклічний надлишковий контроль. Суть першої інтерпретації CRC полягає у перевірці правильності даних за відомими правилами завадостійкого кодування, а другої інтерпретації – перевірка даних за допомогою контрольної суми [2]. Далі будемо розглядати CRC лише як код.

Ще з часу своєї появи CRC вважаються найпростішим різновидом циклічних кодів. Традиційний спосіб обчислення CRC-суми (або синдрому помилки) способом ділення поліномів стовпчиком, на перший погляд, не вимагає грунтовної математичної підготовки. Мабуть тому інженери легко впровадили CRC-контроль в більшість протоколів зв'язку і затвердили його в міжнародні стандарти. І лише після цього почали з'являтися запитання. А як вибрати оптимальний породжувальний поліном коду? А чи може CRC виправляти помилки? А як прискорити обчислення CRC?

В Інтернеті з'явились численні переліки "дуже хороших" поліномів, на зміну яким прийшли "найкращі" [3]. Правда, автори чомусь не змогли надати переконливі докази свого вибору: єдиний критерій – експериментальний. В [4] пропонується робити вибір на основі дослідження вагового спектру коду, але виявляється що це NP-складна задача навіть для невеликих кодів [5].

Тому актуальною є розробка теорії кодів CRC на основе сучасного рівня розвитку завадостійкого кодування та вирішення проблем практичного використання CRC.

Теоретичний базис CRC-кодів

Для розуміння суті CRC-кодів необхідно повернутись до основ циклічних кодів. Традиційні способи представлення циклічних кодів (матричне, поліноміальне, алгебраїчне) відіграли важливу роль в становленні цього класу кодів. Однак практика формує нові проблеми і для їх розв'язання найбільш придатним математичним аппаратом є теорія лінійних послідовнісних схем (ЛПС).

Згідно з [6] ЛПС з r елементами пам'яті, l входами і m виходами є це кінцевий автомат лінійного типу (лінійний автомат), який над полем Галуа $GF(2)$ описується функцією станів (переходів)

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

і функцією виходів

$$Y(t) = C \times S(t) + D \times U(t), GF(2)$$

де t – дискретний час; A, B ,C , D – характеристичні матриці ЛПС;

$S(t)$ – слово стану, $U(t)$ – вхідне слово, $Y(t)$ – вихідне слово.

Автоматне представлення циклічних кодів об'єднує в собі всі попередні способи представлення і додає нові можливості для аналізу властивостей циклічних кодів на основі теорії кінцевих автоматів. Це представлення має дві моделі: автоматно-аналітичну і автоматно-графову.

Автоматно-аналітична модель базується на характеристичних матрицях ЛПС. Множина всіх двійкових послідовностей L довжини n, які переводять ЛПС із будь-якого початкового стану $S_{beg}(t)$ знову в стан $S_{beg}(t)$, утворює циклічний (n, k)-код Ω над полем Галуа $GF(2)$. Кожна така послідовність L є кодовим словом Z циклічного (n, k)-коду. Розмірності матриць ЛПС і параметри циклічного (n, k)-коду Ω пов'язані через коефіцієнт r , який для коду дорівнює числу контрольних розрядів кодового слова Z при систематичному кодуванні ($r = n - k$). Одна з характеристичних матриць ЛПС містить в собі коефіцієнти породжувального поліному $g(x)$ коду.

Практика вибору CRC-кодів з позицій їх автоматно-графової моделі

Автоматно-графовою моделлю циклічного коду може служити граф переходіввиходів ЛПС, який складається із сукупності нульових циклів (НЦ) [6]. Множина всіх кодових шляхів довжини n, які починаються і закінчуються в початковій вершині графу, утворює циклічний (n, k)-код Ω над полем Галуа $GF(2)$.

Автоматно-графові моделі дають дуже багато корисної інформації про коректувальні властивості циклічних кодів, дозволяють дати наочну інтерпретацію їх процедур кодування і декодування. Тісний зв'язок породжувального поліному коду з матрицями ЛПС дозволяє провести дослідження поліномів на основі аналізу графової структури ЛПС. В теорії ЛПС доведено, що збільшення числа співмножників породжувального поліному призводить до збільшення числа НЦ з пропорційним зменшенням їх довжини (тобто, довжини самого коду). А зміна кількості НЦ відповідно впливає на здатність коду виявляти та виправляти помилки [7].

Наприклад, (n, k)-код Хемінга з незвідним примітивним породжувальним поліномом має лише один НЦ довжини n. Цей код дозволяє виправляти всі поодинокі помилки, а також виявляти всі подвійні і велику кількість помилок більшої кратності. Якщо ж породжувальний поліном буде складатись із двох співмножників виду

$$g(x) = (1+x)\,p(x),\ GF(2),$$

де $p(x)$ – примітивний поліном степені $r-1$,

тоді у такого коду, який іменується кодом Абрамсона, збільшується вдвічі кількість НЦ і зменшується вдвічі їх довжина. Безперечною перевагою цього коду є можливість виявлення максимальної кількості помилок навіть далеко за межами традиційної кодової відстані. Для контрольованих повідомлень довжини не більшої за n код Абрамсона є кращим кодом CRC, якщо необхідно лише встановити сам факт наявності помилки.

В деяких ситуаціях може бути виправданою операція виправлення помилок.

Для такої задачі необхідні зовсім інші поліноми: з великою кількістю НЦ малої довжини. Збільшення коректуючої здатності коду призводить одночасно до зменшення здатності виявляти помилки та суттєвому зменшенню довжини контрольованих повідомлень. При відсутності графової моделі коду важко оцінити такі залежності, що і призводить до постановки нездійсненної задачі: знайти "найкращий" поліном для всіх випадків [3, 4].

Цікаво проаналізувати деякі стандартизовані 8-розрядні поліноми з позицій автоматної моделі циклічних кодів (Табл. 1). Легко помітити, що DARK-поліном здатний

виправити помилки підвищеної кратності, але для дуже коротких кодів, довжина коду з ATM-поліномом недостатня для контролю 48 байтів ATM-пакету даних, а лише для його заголовку, а ETSI-поліном зі всіх позицій є надзвичайно невдалим вибором.

Таблиця 1 – Характеристика стандартизованих поліномів CRC-кодів

Породжувальний поліном (назва стандарту)	**Тип породжувального поліному**	**Тип коду**	**Довжина коду** n	**Структура графової моделі**
$1+x+x^2+x^8$ (ATM)	$g(x)=(1+x)p(x), p(x)$ -примітивний	Абрамсона	127	Два НЦ довжини 127
$1+x^3+x^4+x^5+x^8$ (DARK)	незвідний, непримітивний	Квадратично-лишковий	17	15 НЦ довжини 17
$1+x^4+x^5+x^8$ (1-Wire bus)	$g(x)=(1+x)p(x)$, $p(x)$-примітивний	Абрамсона	127	Два НЦ довжини 127
$1+x^2+x^4+x^6+x^7+x^8$ (ETSI)	непримітивний		93	Два НЦ довжини 93
$1+x^2+x^3+x^4+x^8$ (SAE J1850)	примітивний	Хемінга	255	Один НЦ довжини 255

Список використаних джерел:

1. Столлингс, В. Компьютерные системы передачи данных / В. Столлингс; изд. 6е; пер. с англ. – М., Издательский дом «Вильямс», 2002. – 928 с. – ISBN 979966-613-532-5.

2. Семеренко, В. П. Теория и практика CRC кодов: новые результаты на основе автоматных моделей / В. П. Семеренко // Східно-Європейський журнал передових технологій. – 2015. – Т. 4, № 9 (76). – С. 38–48. doi: 10.15587/1729-4061.2015.47860.

3. Koopman, P. Effcient high hamming distance CRCs for embedded networks / J. Ray and P. Koopman // The International Conference on Dependable Systems and Networks (DSN2006), 2006, Philadelphia PA, June 25-28. – P. 3–12.

4. Baicheva, T. S. Determination of the best CRC codes with up to 10-bit redundancy / T. S. Baicheva // IEEE Transactions on Communications. – 2008. – Vol. 56, Issue 8. – P. 1214–1220. doi: 10.1109/tcomm.2008.070033.

5. McLoughlin, A. The complexity of computing the covering radius of a code / A. McLoughlin // IEEE Trans. Inf. Theory. – 1984. – Vol. 30. – P. 800–804.

6. Семеренко В. П. Теорія циклічних кодів на основі автоматних моделей : монографія / В. П. Семеренко. – Вінниця : ВНТУ, 2015. – 444 с. –ISBN 978-966-641-624-0.

7. Семеренко, В. П. Оценка корректирующей способности циклических кодов на основе автоматных моделей / В. П. Семеренко // Східно-Європейський журнал передових технологій. – 2015. – Т. 2, № 9 (74). – С. 16–24. doi: 10.15587/1729-4061.2015.39947.