УДК 621.391.037.372

Савченко Ю. Г., *д.т.н.*

ОЦІНКА РЕАЛЬНОЇ СПРОМОЖНОСТІ ДО ВИЯВЛЕННЯ ПОМИЛОК ГРУПОВИМИ КОДАМИ

Савченко Ю. Г. Оцінка реальної спроможності до виявлення помилок груповими кодами. Проведена оцінка реальної спроможності групових кодів до виявлення помилок високої кратності. Показано, що така оцінка практично завжди є вищою, ніж при розрахунках на основі мінімальної кодової відстані. Розглянуті приклади аналізу конкретних кодів.

Ключові слова: ГРУПОВИЙ КОД, КОДОВА ВІДСТАНЬ, ВИЯВЛЯЮЧИ СПРОМОЖНІСТЬ, ВИСОКОКРАТНІ ПОМИЛКИ

Савченко Ю. Г. Оценка реальной способности к выявлению ошибок групповыми кодами. Проведена оценка реальной способности групповых кодов к обнаружению ошибок высокой кратности. Показано, что такая оценка практически всегда является более высокой по сравнению с расчетами на основе минимального кодового расстояния. Приведены примеры анализа конкретных кодов.

Ключевые слова: ГРУППОВОЙ КОД, КОДОВОЕ РАССТОЯНИЕ, ОБНАРУЖИВАЯ СПОСОБНОСТЬ, ВЫСОКОКРАТНЫЕ ОШИБКИ

Savchenko Yu. H. Estimation of the real possibility to detect of errors by group codes. The research is dedicated to real ability of group codes to detect multiple errors. It is shown that traditional linear code distance based evaluations can only be used for rough computation of information exchange reliability. Analysis of specific codes is presented.

Key words: group code, code distance, **detect** ability, multiple errors

Сьогодні застосування кодів із корекцією помилок в телекомунікаційних системах стало вже не винятком, а нормою. Це пов'язано, передусім, із непередбачуваним рівнем завад навіть в стаціонарних каналах інформаційного обміну та, з іншого боку, відносно невеликими витратами на реалізацію процедур кодування та декодуваня. Тому переважна більшість протоколів і стандартів міжкомп'ютерного зв'язку передбачає завадозахищене кодування.

В той же час оцінка ефекту від застосування кодів, як правило, провадиться на основі відомих співвідношень, які зв'язують кратність t помилок, що виявляються або виправляються, з мінімальною кодовою відстанню [1]:

$$d_{\min} = t_{\text{вияв}} + 1,$$

$$d_{\min} = 2t_{\text{випр}} + 1.$$

Але ці прості співвідношення, фактично, визначають лише мінімальний гарантований ефект від застосування кодів (принаймні, це справедливо до виявлення помилок). Насправді ж, як буде показано, будь-який код має суттєво більшу виявляючу спроможність, враховуючи помилки високої кратності.

Можна використати також оцінку на основі співвідношення між кількістю заборонених та дозволених кодових слів [2]:

$$\delta = \frac{2^n - 2^k}{2^n - 1}, \tag{1}$$

де n – повна довжина кодової комбінації, k – кількість інформаційних символів, δ – доля (частина) помилок, які виявляються кодом.

Співвідношення (1) хоч і дає точну оцінку спроможності кода до виявлення помилок, але не розкриває їх розподілу по кратності.

З практичної точки зору, користувача послуг зв'язку цікавить, насамперед, реальна достовірність інформаційного обміну, тобто ймовірність того, що отримане повідомлення не містить помилок.

Традиційно завадозахищені коди ділять на два класи: коди з виправленням помилок і коди з виявленням помилок (є, правда, деякі винятки, коли код виправляє деяку сукупність помилок та одночасно і додатково виявляє іншу сукупність, але це для подальшого аналізу не є суттєвим). В даній роботі обмежимося лише аналізом спроможності будь-якого коду саме до виявлення помилок. З практичної точки зору ця задача є досить актуальною, оскільки коди з виправленням помилок мають значну інформаційну надлишковість, що помітно зменшує реальну пропускну спроможність каналу. До того ж процедури декодування при виправленні помилок (у порівнянні з їх виявленням) вносять значну затримку, що не завжди є припустимим при деяких застосуваннях. При реальних рівнях завад та наявності зворотного каналу для передачі інформації про виявлені помилки та організації повторної передачі спотвореного блоку застосування саме кодів із виявленням помилок можна вважати найбільш доцільним.

Застосування кодів із корекцією помилок завжди пов'язано з введенням інформаційної надлишковості, тобто подовженням вихідних груп (пакетів, блоків, кадрів). В умовах незалежності помилок в окремих символах бітового потоку зі збільшенням довжини груп n сукупна ймовірність виникнення помилок більш високої кратності t також зростає відповідно до біноміального закону

$$P(t) = \sum_{i=1}^{t} C_n{}^i q^i (1-q)^{n-i} .$$

Будь-яка система передачі даних може знаходитись лише в одному з трьох станів:

A – відсутність помилок при передачі;

B – невиявлення помилок;

C – виявлення помилок.

Тоді ймовірність того, що передані дані дійсно є безпомилковими визначається як

$$P(A) = 1 - P(B) - P(C).$$

Із усієї сукупності можливих помилок кратності t, кількість яких визначається комбінаторним числом $C^t{}_n$, кожен конкретний код виявляє лише деяку частину з них, а саме N_t помилок (при $C^t{}_n = N_t$ код виявляє *всі* помилки відповідної кратності). Тому попередній вираз для оцінки ймовірності відсутності помилок в переданому повідомленні можна записати у вигляді

$$P(A) = \sum_{t=1}^{n} (C^t{}_n - N_t) q^t (1-q)^{n-t} . \qquad (2)$$

Таким чином, виходячи з виразу (2), для оцінки *реальної* спроможності конкретного коду необхідно обчислити ряд значень N_t, $t = 1, 2, ...n$. Очевидно, ці значення можуть бути отримані лише на основі властивостей породжуючої або перевірочної матриці групового коду (для циклічних кодів – породжуючого поліному). Так, наприклад, найпростішому коду із однією перевіркою на парність відповідає перевірочна матриця

$$P = \|11...1\|,$$

де кожний стовпчик – синдром однократної помилки у відповідному розряді бітової групи з n символів. Використовуючи цю очевидну властивість перевірочної матриці, можна стверджувати, що код буде виявляти всі однократні помилки та всі помилки непарної кратності, тобто

$$(C^t{}_n - N_t) = 0 \text{ для всіх } t = 1, 3, 5, ...,$$

оскільки сума по модулю 2 будь-якої непарної кількості стовпчиків матриці буде дорівнювати 1. Аналогічно, помилки парної кратності код не виявляє, оскільки суми по модулю 2 парної кількості стовпчиків будуть дорівнювати 0.

Це був найпростіший приклад, а оцінка спроможності до виявлення помилок коду з однією перевіркою на парність добре відома.

У загальному випадку для більш потужних кодів така оцінка може виявитися досить трудомістською. Зокрема, маючи як вихідну перевірочну матрицю коду, необхідно перебором сполучень стовпчиків виявити такі з них, порозрядна сума по модулю 2 котрих не є нульовою, - це ті вектори кратних помилок, що виявляються кодом. Підрахунок кількості таких сполучень визначить необхідні для оцінки значення N_t.

Розглянемо для ілюстрації такого підходу перевірочну матрицю класичного (7,4)-коду Хеммінга

$$\mathrm{P}=\begin{Vmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

В табл.1 наведені результати обчислень для усіх можливих кратних помилок

Табл. 1

t	1	2	3	4	5	6	7
$C^t{}_7$	7	21	35	35	21	7	1
N_t	7	21	28	28	21	7	0
$C^t{}_7 - N_t$	0	0	7	7	0	0	1

Із таблиці видно, що цей код *виявляє* всі помилки кратності 1, 2, 5 та 6, *не виявляє* деяку кількість помилок кратності 3 і 4 та 7-кратну помилку. Безпосередній розрахунок показує, що код виявляє 88,2% усіх можливих помилок.. Отримане значення співпадає з розрахунком на основі співвідношення (1)

$$\delta = \frac{2^n - 2^k}{2^n - 1} = \frac{128 - 16}{127} = 0{,}882.$$

Аналогічно можна розрахувати спроможність цього коду до виявлення пачок помилок довжини b (табл.2).

В таблиці символами M та N, відповідно. позначено комбінаторне число теоретично можливих помилок в пачці та таких, що реально виявляються.

Табл. 2

b	2	3	4	5	6	7
M	6	5	4	3	2	1
N	6	4	2	3	2	0
$M-N$	0	1	2	0	0	1

Ще один приклад спроможності до виявлення помилок популярного (8,4)-коду з перевірочною матрицею

$$P=\begin{Vmatrix}1&0&1&1&0&0&0&0\\1&1&0&0&1&1&0&0\\1&1&1&0&1&0&1&0\\1&1&1&1&1&1&1&1\end{Vmatrix}.$$

Як і слід очікувати, спроможність до виявлення помилок для цього коду зростає у порівнянні з кодом (2), що видно з табл. 3.

Табл. 3

t	1	2	3	4	5	6	7	8
C^t_8	8	28	56	70	56	28	8	1
N_t	8	28	56	56	56	28	8	0
$C^t_8-N_t$	0	0	0	14	0	0	0	1

Доля помилок, які цей код виявляє, складає вже 94%.

Можна було б навести багато прикладів інших кодів, які використовуються в телекомунікаційних системах. Зокрема, аналогічний аналіз проведено для ASC11 (11,7)-коду, міжнародного телеграфного (15,11)- коду МТК-2 та багатьох циклічних кодів класу CRC.

Задаючи ймовірність виникнення помилки в одному біті цифрового потоку q (числове значення визначається фізичними властивостями реального каналу передачі) та закон розподілу ймовірностей виникнення помилок в групах символів, можна розрахувати реальну достовірність отриманого блоку. Так, при застосуванні (7,4)-коду та $q=10^{-8}$ ймовірність невиявлення помилки складає $7\cdot10^{-24}$ при незалежних помилках в окремих символах бітового потоку.

Таким чином, запропонований підхід дозволяє *точно* визначити спроможність заданого коду до виявлення конкретних помилок. Саме ця особливість підходу створює передумови для реалізації телекомунікаційних систем із гарантованою достовірністю інформаційного обміну, яка за фізичним змістом є, по суті, ймовірністю відсутності помилок в повідомленні.

Література

1. Берлекемп Э. Алгебраическая теория кодирования / Э. Берлекемп. – М.: Мир, 1971. – 477 с.

2. Локазюк В. М. Надійність, контроль, діагностика і модернізація ПК / В. М. Локазюк, Ю. Г. Савченко. –К.: Видавничий центр «Академія», 2004. –373 с.