УДК 004.052.42

МАРКОВСКИЙ А.П.,
УВАРОВ Н.В.,
УВАРОВА Н.В.

ГАРАНТИРОВАННОЕ ОБНАРУЖЕНИЕ ЧЕТЫРЕХКРАТНЫХ ОШИБОК С ИСПОЛЬЗОВАНИЕМ МИНИМАЛЬНОГО ЧИСЛА КОНТРОЛЬНЫХ РАЗРЯДОВ

В статье описан метод гарантированного обнаружения четырехкратных ошибок передачи данных в симметричном двоичном канале. Особенность разработанного метода заключается в использовании теоретически минимального числа контрольных разрядов. Метод основывается на использовании весовых коэффициентов, сформированных на основе специального нелинейного функционального преобразования. Показано, что использование таких коэффициентов для вычисления контрольных разрядов обеспечивает гарантированное обнаружение четырехкратных ошибок при минимально возможной длине контрольного кода. В статье проанализированы свойства специальных нелинейных преобразований. Определены условия, которым должны соответствовать как само преобразование, так и ее дифференциал.

The article is dedicated to the method of the guaranteed detection of quadruple errors during transmission via symmetrical binary channel. The main feature of the proposed method is the theoretically minimal length of the error detection code. The method is based on the use of the weight coefficients formed on the basis of a special non-linear functional transformation. It is shown that use of such coefficients in control code calculation provides guaranteed detection of quadruple errors with the minimal length of the error detection code. Also in the article the properties of such non-linear transformation has been analyzed. The conditions to be met for both the transformation itself and its differential are determined.

Введение

В условиях углубления процессов информационной интеграции и расширения использования распределенных компьютерных систем возрастает роль средств передачи данных. Важная роль в прогресса средств передачи и обмена информации принадлежит технологиям контроля возникающих ошибок, частью которых являются способы обнаружения битовых искажений различной кратности.

В современных условиях объективно существуют факторы, педалирующие рост кратности возникающих ошибок. В частности, динамичный рост скорости передачи данных имеет следствием рост числа ошибок, вызванных межсигнальной интерференцией [1]. Расширяющееся использование беспроводных средств передачи данных существенно повысило интенсивность внешних электромагнитных помех, что приводит к увеличению вероятности возникновения ошибок. Отмечается [1] и тенденция роста длины контролируемых блоков, данных, которые составляют в настоящее время тысячи байт. Ясно, что с ростом блины блока возрастает вероятность появления многократных битовых искажений.

В отличие от систем связи или цифрового телевидения, при передаче данных между компонентами компьютерных систем необходимо обеспечить качественно более высокий уровень надежности. Соответственно, ошибка передачи данных в компьютерных системах имеет значительно большую цену в сравнении с системами связи.

Таким образом, проблема повышения эффективности обнаружения ошибок большей кратности, возникающих при передаче данных является актуальной в реалиях современного этапа развития компьютерных технологий.

Анализ кодов, используемых для обнаружения ошибок

Для обнаружения ошибок передачи данных в компьютерных системах преимущественно используются:

– циклические избыточные коды (CRC – Cyclic Redundancy Codes);

– контрольные суммы (CS – Check Sum).

И контрольные суммы, и CRC относятся к средствам блокового контроля.

При использовании этих кодов, к блоку из m битов $B=\{b_1,b_2,\ldots,b_m\}$, $\forall l=1,\ldots,m$: $b_l\in\{0,1\}$ добавляется k контрольних разрядов

$\{b_{m+1},b_{m+2},\ldots,b_{m+k}\}$. Совокупность всех возможных блоков вместе с зависящими от них контрольными разрядами образует множество Ω разрешенных кодов. Получение приемником неразрешенного кода означает ошибку. Обнаруживающие возможности кодов характеризуются минимальной величиной кодового расстояния d_{min} между разрешенными кодами, которое определяется следующей формулой [1]:

$$d_{min} = \min_{X,Y\in\Omega} \sum_{j=1}^{m+k} (x_j \oplus y_j).$$

Для обнаружения ошибок кратностью h необходимо и достаточно, чтобы выполнялось условие [1]:

$$d_{min} > h. \qquad (1)$$

Выражение (1) устанавливает теоретические границы числа необходимых контрольных разрядов.

Для гарантированного выявления 4-кратных ошибок, которые возникают при передаче m-битового блока, минимальное количество k контрольных разрядов должно быть таким, что минимальное хемингово расстояние между $(m+k)$-битовыми кодами равнялась пяти. Таким образом, можно показать, что: $k = \lfloor 2 \cdot log_2 m \rfloor$.

Наиболее распространенным в вычислительной технике способом контроля правильности передачи данных является CRC.

Сущность контроля с использованием циклических избыточных кодов состоит в том, что контролируемый блок B представляется в виде полинома $P(B)$ степени $m+k$:

$$P(B)=b_1 \cdot x^k + b_2 \cdot x^{k+1} + \ldots + b_{m-1} \cdot x^{k+m-1} + b_m \cdot x^{k+m} \qquad (2)$$

Контрольный код $R(B)$ вычисляется как остаток от деления полинома $P(B)$, определяемого (2), на образующий полином $Q(X)$ степени k циклического избыточного кода.

По основному критерию эффективности – надежности обнаружения ошибок, циклические коды превосходят контрольные суммы. Ошибки при передаче блока данных не обнаруживается, если полином $E(X)$, соответствующий вектору ошибки, делится на образующий полином CRC $Q(X)$ без остатка. Показано [2], что все полиномы $E(X)$, соответствующие указанным ниже ошибкам, не делятся на специальным образом выбранный базовый полином $Q(X)$, а, следовательно, они обнаруживаются гарантированно:

1. Все искажения битов $b_1,b_2,\ldots,b_m$ нечетной кратности, если базовый полином $Q(X)$ может быть представлен в виде произведения полиномов: $Q(X) = (x+1) \cdot S(X)$;

2. Все двукратные искажения битов контролируемого блока, если базовый полином $Q(X) = q_0 + q_1 \cdot x + q_2 \cdot x^2 + \ldots + q_k \cdot x^k$ содержит не менее трех ненулевых компонент: $\sum_{i=0}^{k} q_i \geq 3$;

3. Группа ошибок, локализованных в рамках k разрядов.

Для остальных ошибок показано [2], что остаток $R(B)$ представляет собой результат хеширования блока B данных в пространство 2^k всех возможных контрольных кодов. Соответственно, вероятность P_{CRC} того, что эти ошибки не будут обнаружены с использованием CRC с образующим полиномом $Q(X)$ степени k, определяется как: $P_{CRC} = \frac{1}{2^k}$.

Наряду с высокой надежностью, CRC обладает рядом недостатков, наиболее важным из которых является принципиально последовательный характер вычисления контрольного кода, что обуславливает существование ограничений на скорость выполнения операций, связанных с контролем ошибок [3]. Этот недостаток особенно актуален в современных условиях быстрого роста скоростей передачи.

Как и все способы блокового контроля, CRC обладает низкой скоростью реакции на возникновение ошибки передачи данных, что ограничивает его использование в компьютерных системах реального времени [3].

Важным достоинством контроля передачи блоков информации в компьютерных сетях с использованием CS, по сравнению с другими методами блокового контроля, является простота реализации и высокая скорость [4].

Традиционно, контрольная сумма S блока данных, представляющего собой n k-разрядных символов $D_1,D_2,\ldots D_n$, формируется в виде поразрядной суммы по модулю 2 всех символов блока: $S = D_1 \oplus D_2 \oplus \ldots \oplus D_n$.

Вполне очевидно, что при использовании традиционной контрольной суммы, обнаруживаются все ошибки нечетной кратности.

Основным достоинством контрольных сумм является отсутствие ограничений на скорость реализации вычислительных операций, связан-

ных с контролем ошибок. Это обусловлено тем, что структура вычислений контрольных сумм допускает возможность широкого распараллеливания при аппаратной реализации.

Как следует из проведенного анализа, CRC и CS имеют ограниченный класс гарантированно обнаруживаемых классов ошибок. Для расширения этого класса необходимо существенно изменить организацию этих средств контроля. Вместе с тем, использование кодов Хемминга позволяет гибко расширять класс гарантированно обнаруживаемых ошибок [2].

Тем не менее, q-мерными кодами Хемминга принципиально не может быть обнаружены ошибки кратности 2^q, которые образуют куб в q-мерном пространстве. Это означает, что двумерными контрольными кодами Хемминга не может быть обнаружены 4 ошибки, которые образуют прямоугольник на плоскости. При применении трехмерных кодов Хемминга (q=3), наименьшая кратность ошибки, для которой не обеспечивается гарантированное обнаружение, равна 2^3=8. Так, не обнаруживаются ошибки, точки локализации которых в трехмерном пространстве образуют куб. То есть, при использовании трехмерных контрольных кодов Хемминга, гарантированно могут быть обнаружены ошибки кратностей 2,4 и 6.

Поэтому для гарантированного обнаружения четырех ошибок необходимо наличие $k_3 = 3 \cdot (\sqrt[3]{m})^2$ контрольных разрядов. Основным недостатком контрольных сумм Хемминга является быстрый рост контрольных разрядов, необходимых для гарантированного обнаружения многократных ошибок, при увеличении их кратности. Так, для гарантированного обнаружения 4-кратных ошибок при передаче 1024 бит, использование кода Хемминга требует 305 контрольных разрядов, что резко снижает эффективность контроля ошибок.

В таблице 1 приведены значения числа контрольных разрядов, которые обеспечивают гарантированное обнаружение 4-кратных ошибок. Вполне очевидно, что приведенные в таблице 1 значения существенно превышают теоретический минимум числа контрольных разрядов, устанавливаемого формулой (1).

Это свидетельствует о невысокой эффективности кодирования контрольной информации кодами Хемминга для случая 4-кратных ошибок.

Таким образом, использование кодов Хемминга не решает в достаточной мере эффективно актуальную в современных условиях проблему расширения класса гарантированно обнаруживаемых ошибок

Табл. 1. Количество контрольных разрядов кода Хемминга для гарантированного обнаружения 4-кратных ошибок

Длина кода	Число контрольных разрядов	Длина кода	Число контрольных разрядов
64	48	1024	305
128	76	2048	485
256	121	4096	768
512	193	8192	1223

Метод гарантированного обнаружения 4-кратных ошибок с использованием взвешенных контрольных сумм

Пусть контролируемый блок B состоит из m бит: $B=\{b_1,b_2,\ldots,b_m\}$, $b_l \in \{0,1\}$, $l=1,\ldots,m$.

Код W_j весового коэффициента j-го бита b_j контролируемого блока предлагается формировать в виде 3-х компонент. Первой компонентой является n-разрядный ($n=\log_2 m$) двоичный код $X_j=\{x_{1j},\ldots,x_{nj}\}$, $\forall i \in 1,..,n$: $x_{ij} \in \{0,1\}$ номера j такой, что $j=x_{1j}\cdot 2^{n-1}+x_{2j}\cdot 2^{n-2}+\ldots+x_{(n-1)j}\cdot 2+x_{nj}$,. Вторая компонента – это n-разрядный результат булевого преобразования $F(X_j)$ над кодом X_j. Третья компонента – единичный бит. Таким образом, W_j имеет вид: $W_j=<X_j, F(X_j), 1>$.

Функциональное преобразование $F(X)$ выбирается таким образом, чтобы для любой пары битов блока сумма по модулю 2 их весовых коэффициентов не повторялась. Формально это означает, что для любых 4-х различных номеров битов $q,l,g,r \in \{1,\ldots.m\}$ всегда выполняется неравенство:

$$W_q \oplus W_l \neq W_g \oplus W_r . \qquad (3)$$

Описанный выбор функционального преобразования $F(X)$ иллюстрируется следующим примером. Если блок содержит 15 бит (m=15), то одним из преобразований $F(X)$, при котором выполняется (3), является преобразование, представленное в таблице 2.

Если выбрать, например, четыре номера: q=5, l=7, g=12 и r=14, то их весовые коэффициенты равны соответственно:

$W_q = <5,11,1>$, $W_l = <7,10,1>$, $W_g = <12,9,1>$ и $W_r = <14,4,1>$.

Легко убедиться, что для выбранных номеров неравенство (3) выполняется:

$$W_q \oplus W_l = <2,1,0> \neq W_g \oplus W_r = <2,13,0>.$$

Табл. 2. Пример преобразования F(X), для которого выполняется (3)

X	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$F(X)$	0	1	2	3	7	11	12	10	13	14	5	8	9	6	4	15

При использовании взвешенных контрольных сумм, контрольный код C блока данных на приемнике и передатчике вычисляется в виде суммы по модулю два произведений битов на соответствующие им весовые коэффициенты:

$$C = \bigoplus_{j=1}^{m} b_j \cdot W_j . \qquad (4)$$

Таким образом, число k_c разрядов контрольного кода C совпадает с разрядностью весовых коэффициентов и определяется формулой:

$$k_c = 2 \cdot n + 1 = 2 \cdot log_2 m + 1 . \qquad (5)$$

Контрольный код C состоит из 3-х компонент: $C=<C_X, C_F, C_p>$, где C_X – сумма по модулю 2 номеров единичных битов передаваемого блока, C_F – сумма по модулю 2 результатов функционального преобразования $F(X)$ номеров единичных битов, а C_p – сумма по модулю 2 битов блока, или бит четности.

В процессе получения блока, приемник формирует свой контрольный код C_R и, по завершению передачи, вычисляет сумму по модулю 2 сформированного и принятого от приемника контрольного кода C_S: $\Delta = C_R \oplus C_S = <\Delta_X, \Delta_F, \Delta_p>$. Если разность Δ контрольных кодов приемника и передатчика равна нулю, то считается, что блок передан без ошибок.

Если в процессе передачи ошибочно передано нечетное количество битов, то значения бита четности контрольных кодов приемника и передатчика отличны и, соответственно, $\Delta_p=1$, то есть $\Delta \neq 0$. Это значит, что битовые ошибки нечетной кратности гарантированно обнаруживаются.

При ошибочной передаче 2-х бит, имеющих номера q и l, код Δ разности контрольных кодов приемника и передатчика равен сумме соответствующих весовых коэффициентов: $\Delta = W_q \oplus W_l = <X_q \oplus X_l, F(X_q) \oplus F(X_l), 0>$. Поскольку первая компонента Δ являет собой сумму по модулю 2 двух различных кодов: $\Delta_X = X_q \oplus X_l$, то $\Delta_X \neq 0$. Это означает, что при двукратной ошибке первая компонента Δ не равна нулю, следовательно, такая ошибка гарантированно обнаруживается.

При ошибочной передаче 4-х битов, номера которых равны q, l, g и r, разность контрольных сумм приемника и передатчика $\Delta = W_q \oplus W_l \oplus W_g \oplus W_r$. Поскольку, согласно (3), $W_q \oplus W_l \neq W_g \oplus W_r$, то $\Delta \neq 0$, что означает гарантированное обнаружение любой 4-кратной ошибки.

Выводы

В результате проведенного исследования разработан метод гарантированного обнаружения четырехкратных ошибок передачи данных в каналах, теоретической моделью которых является двоичный симметричный канал.

Главное отличие предложенного метода от существующих – использование теоретически минимального числа контрольных разрядов. При вычислении контрольных разрядов используются весовые коэффициенты, формируемые с помощью специального нелинейного функционального преобразования. Исследованы свойства таких преобразований.

Вычисление контрольных разрядов легко может быть распараллелено, что обеспечивает возможность контроля ошибок в темпе передачи данных. Предложенный метод может быть эффективно использован для контроля ошибок передачи данных в высокоскоростных каналах между компонентами компьютерных систем, работающих в режиме реального времени.

Список литературы

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: Изд. Дом "Вильямс".– 2004.– 1104 С.
2. Klove T., Korzhik V. Error Detecting Codes: General Theory and Their Application in Feedback CommunicationSystems. Norwell, MA: Kluwer, 1995. – 433 p.
3. Nikolaos G. Bardis, Athanasios Drigas and Oleksandr P. Markovskyi, Performance Increase of Error Control Operation on Data Transmission", 3rd International Conference on New Technologies, Mobility and Security, IEEE, *IEEE COMMUNICATIONS SOCIETY*, Egypt – Cairo, 20-23 December 2009.
4. Самофалов К.Г., Марковский А.П., Мулки Яссин Ахмед Ал Бадайнех. Обнаружение и исправление ошибок передачи данных с использованием взвешенных контрольных сумм // Проблеми інформатизації та управління. Збірник наукових праць: Випуск 3(14). – К.,НАУ. – 2008. – С.121– 128.