УДК 004.452.42

Самофалов К.Г., чл.-кор. НАН Украины
Марковский А.П., канд. техн. наук
Мулки Яссин Ахмед Ал Бадайнех

ОБНАРУЖЕНИЕ И ИСПРАВЛЕНИЕ ОШИБОК ПЕРЕДАЧИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ВЗВЕШЕННЫХ КОНТРОЛЬНЫХ СУММ

Национальный технический университет Украины "КПИ"

Рассмотрена проблема повышения эффективности обнаружения и исправления ошибок передачи данных с использованием взвешенных контрольных сумм. Разработаны процедуры коррекции однократных и многократных ошибок с повторной пересылкой части контролируемого блока данных. Для модели симметричного двоичного канала получены теоретические оценки вероятностей ошибок коррекции, которые позволяют обоснованно подойти к выбору технологии исправления ошибок. Проведен теоретический анализ эффективности предложенного подхода, на основе которого получены оценки уменьшения объема передаваемых для исправления ошибок данных по сравнению с известными технологиями

Введение

Передача цифровых данных является одним из наименее надежных среди компонент систем обработки информации. Трудности обеспечения высокой достоверности передачи данных обусловлены сложной природой физических процессов в длинных линиях, отражением сигналов и их взаимным влиянием (межсигнальная интерференция), влиянием внешних помех. Поэтому, с конца 40-х годов интенсивно развивается технология обнаружения и исправления ошибок передачи данных.

В последнее десятилетие отмечается качественный прогресс с области технологии передачи данных: многократно возросли объемы и скорость передачи информации. Вместе с тем, проблема обеспечения эффективного контроля ошибок приобрела большую важность. Уменьшение временных интервалов между сигналами, расширение применения спектрального уплотнения имеют следствием рост числа ошибок интерференционной природы [1]. Увеличение интенсивностей электромагнитных полей, обусловленных расширением применения беспроводных линий передачи данных вызывает рост числа ошибок, вызванных внешними помехами. Для новых технологий передачи данных характерны специфические типы возникающих искажений информации. Все это диктует необходимость развития средств контроля и исправления ошибок, адекватного прогрессу технологии передачи данных.

Таким образом, проблема повышения эффективности обнаружения и исправления ошибок с учетом особенностей современных технологий передачи данных является важной и актуальной для развития компьютерных систем и сетей.

Анализ методов обнаружения и исправления ошибок передачи данных

Для контроля и исправления ошибок, возникающих при передаче данных используются два основных подхода: обнаружение и исправлением ошибок без повторной передачи с использованием корректирующих кодов, а также обнаружение ошибок с повторной передачей контролируемого блока с случае обнаружения ошибок. Основными недостатками использования корректирующих кодов является необходимость в большом числе контрольных разрядов и неэффективность исправления многократных ошибок. На практике использование корректирующих кодов оправдано при относительно большой вероятности ошибок (в частности, при передаче данных в эфирных каналах) и в случае невозможности повторения передачи [4]. В проводных линиях передачи, для которых характерной является ма-

лая интенсивность ошибок, для исправления ошибок используется повторная передача.

Для обнаружения ошибок чаще всего используются циклические избыточные коды (*CRC- Cyclic Redundancy Check*) и различные разновидности контрольной суммы (*CS-Check Sum*). Одной из наиболее эффективных из них является взвешенная контрольная сумма (*WCS-Weighed Check Sum*) [2, 3].

Пусть контролируется правильность передачи блока B данных, состоящего из m битов: $B=\{b_1,b_2,...,b_m\}$, $b_i\in\{0,1\}$, $i=1,\ldots,m$. Взвешенная контрольная сумма WCS на приемнике и передатчике формируется, как сумма по модулю 2 m $(k+1)$-разрядных ее компонент $W_1,W_2,...,W_m$:

$$CS=W_1\oplus W_2\oplus\ldots\oplus W_m. \qquad (1)$$

Значение каждой j–той компонента контрольной суммы – W_j, $j\in\{1,\ldots,m\}$ определяется значением одноименного биту b_j контролируемого блока и j-тым весовым кодом U_j. При этом, компонента W_j формируется как конкатенация j-того бита b_j блока B и логического произведения этого бита на каждый из разрядов кода U_j:

$$\forall j\in\{1,...,m\}: W_j=b_j\,\|\,b_j\cdot U_j. \qquad (2)$$

Контрольная сумма, определяемая в соответствии с (1) и (2) вычисляется отдельно на передатчике и приемнике. Блоки данных на передатчике и приемнике, равно как и составляющие их биты, обозначим как $B_S=\{b_{1S},b_{2S},...,b_{mS}\}$ и $B_R=\{b_{1R},b_{2R},...,b_{mR}\}$ соответственно. После передачи блока данных, контрольная сумма CS_S передатчика передается на приемник, где сравнивается с контрольной суммой CS_R, вычисленной на приемнике.

Решение о ошибочной передаче блока принимается, если отличен от нуля $(k+1)$-разрядный код Δ разницы контрольных сумм передатчика CS_S и приемника CS_R : $\Delta=\{\delta_1,\delta_2,...,\delta_{k+1}\}=CS_S\oplus CS_R$, причем компоненты кода Δ разности определяются в виде:

$$\delta_1=\bigoplus_{j=1}^{m}(b_{jS}\oplus b_{jR})$$
$$Z=\{\delta_2,\delta_3,...,\delta_{k+1}\}=\bigoplus_{j=1}^{m}(b_{jS}\oplus b_{jR})\cdot U_j \qquad (3)$$

При возникновении в процессе передачи блока нечетного числа битовых искажений, δ_1 кода Δ разницы контрольных сумм передатчика и приемника в силу (3) представляет собой сумму по модулю 2 нечетного количества единичных компонент, которые соответствуют несовпадающим одноименным битам блоков данных на приемнике и передатчике. Соответственно, $\delta_1=1$. Это означает, что нечетное число ошибок, возникших в процессе передачи блока данных, гарантированно обнаруживается при использовании взвешенной контрольной суммы.

Множество весовых кодов $\Omega_m=\{U_1,U_2,...,U_m\}$ представляет собой множество частично-ортогональных кодов, таких, что сумма по модулю 2 любого их подмножества ϑ, которое включает в себя не более h таких кодов не равно нулю:

$$\forall\,\vartheta=\{V_1,V_2,...,V_{q_\vartheta}\}\subset\Omega_m,$$
$$q_\vartheta\le h: V_1\oplus V_2\oplus\ldots\oplus V_{q_\vartheta}\neq 0 \qquad (4)$$

Фактически, это означает, что любое подмножество $\vartheta\subset\Omega_m$, $|\vartheta|\le h$ представляет собой ортогональный базис в h-мерном пространстве.

При возникновении в процессе передачи m-битового блока данных ошибок четной кратности $d\le h$, номера d искаженных битов образуют множество Θ : $|\Theta|\le h$. Бит δ_1 кода Δ разности контрольных сумм передатчика и приемника принимает нулевое значение в силу того, что $d\ mod\ 2 = 0$, а значение k-битовой компоненты Z кода Δ определяется следующим выражением:

$$Z=\bigoplus_{i\in\Theta}(b_{iS}\oplus b_{iR})\cdot U_i=\bigoplus_{i\in\Theta}U_i \qquad (5)$$

В силу свойства (4) сумма по модулю 2 не более, чем h частично-ортогональных кодов не может быть равна нулю, а поскольку $|\Theta|\le h$, то $Z\neq 0$, а соответственно и разность контрольных сумм приемника и передатчика $\Delta\neq 0$. Это

означает, что любое искажение не более, чем h битов будет гарантированно обнаружено при использовании взвешенной контрольной суммы.

В простейшем случае, при h=2 множество Ω_m представляет собой $\log_2 m$-разрядные коды порядковых чисел от 1 до m. При h=4 процедура формирования весовых кодов имеет более сложный характер, а их разрядность составляет $2.4 \cdot \log_2 m$ [3]. Таким образом, использование *WCS* при h>2 обеспечивает более широкий класс гарантированно обнаруживаемых ошибок по сравнению с *CRC*.

Обычно, исправление ошибок, обнаруженных с использованием *WCS* или *CRC* производится путем повторной передачи всего контролируемого блока данных. Однако повторная передача всего блока заметно снижает эффективность исправления ошибок, поскольку требует значительного времени и не исключает возможности возникновения ошибок при вторичной передаче.

Целью исследований является повышение эффективности коррекции выявленных с использованием взвешенной контрольной суммы ошибок передачи данных за счет уменьшения объема повторно передаваемой информации.

Исправление однократной ошибки с использованием WCS

При возникновении однократной ошибки, то есть искажения в процессе передачи e-того бита b_e блока, код Δ разности *WCS* приемника и передатчика позволяет однозначно определить его позицию, поскольку $\Delta = U_e$. При этом δ_1=1, а $Z = W_e$. При h=2 W_e представляет собой порядковый номер искаженного бита блока, то есть $W_e = e$, и коррекция осуществляется путем инвертирования W_e-го бита без повторной передачи.

При h>2 код W_e также однозначно соотносится с позицией e искаженного бита, однако $W_e \neq e$ и для получения номера e по коду W_e необходимо выполнять дополнительные преобразования, либо обращаться к таблице $T(W_e, e)$ соответствия кодов W_e и e. В сущности, аналогичная ситуация имеет место и при использовании *CRC*: контрольный код Δ_{CRC} представляет собой остаток от деления полинома

$$b_1 \cdot x^{m+k-1} + b_2 \cdot x^{m+k-2} + \ldots + b_m \cdot x^k + c_k \cdot x^{k-1} + \ldots + c_2 \cdot x + c_1$$

на образующий полином $P(X)$ *CRC* степени k. В случае искажения одного, e-того бита Δ_{CRC} представляет собой остаток от деления полинома x^{m+k-e} на $P(X)$. При $2^k < m$ имеет место однозначное соответствие Δ_{CRC} позиции e искаженного бита, однако, как и в случае *WCS* при h>2 , для определения кода e по Δ_{CRC} необходимо выполнить дополнительные преобразования, либо использовать таблицу.

Таким образом, при использовании *WCS* коррекция однократной ошибки всегда возможна без повторной передачи, причем при h=2 процедура коррекции выполняется без дополнительных преобразований.

Основной проблемой коррекции однократных ошибок является неразличимость последних с ошибками большей нечетной кратности, и, в частности, с трехкратными.

При использовании *WCS* с h>3 коды Δ при возникновении однократной и трехкратной ошибок не могут совпасть, поэтому ошибка кратности не менее 5-ти может быть ошибочно классифицирована как однократная.

При использовании *CRC* и *WCS* (h=2) по коду Δ невозможно различить однократную ошибку и ошибку большей нечетной кратности, включая трехкратную. Соответственно, коррекция одиночной ошибки без повторной передачи, в этих случаях, будет иметь следствием риск того, что коррекция ошибок будет выполнена неверно.

Вероятность указанного риска зависит от характера доминирующего в канале типа ошибок.

В частности, для двоичного симметричного канала оценка вероятности риска может выполнена следующим образом. Предположим, что произошла ошибка нечетной кратности. Для выбора эффективной стратегии ее исправления необходимо оценить вероятность P_1 того, что эта

ошибка – одиночная. Если линия передачи данных соответствует модели двоичного симметричного канала и p- вероятность ошибочной передачи бита, при возникновении ошибки нечетной кратности, апостериорная вероятность P_j того, что она имеет кратность j определяется по форме Бейеса в виде:

$$P_j = \frac{P_j'}{\sum_{t=0}^{m/2-1} P_{2\cdot t+1}'}, \qquad (6)$$

где P_j'- априорная вероятность возникновения ошибки кратности j. При имеющих место на практике малых значения p, априорные вероятности возникновения ошибок для двоичного симметричного канала могут быть определены с использованием модели Пуассона в виде:

$$P_j' = \frac{e^{-m\cdot p}\cdot (m\cdot p)^j}{j!}. \qquad (7)$$

С учетом (7), находящаяся в знаменателе (6) вероятность того, что при передаче блока число ошибок будет нечетным может быть преобразована к виду:

$$\sum_{t=0}^{m/2-1} P_{2\cdot t+1}' = e^{-m\cdot p}\cdot \sum_{t=0}^{m/2-1} \frac{(m\cdot p)^{2\cdot t+1}}{(2\cdot t+1)!} \approx$$
$$e^{-m\cdot p}\cdot \sum_{t=0}^{\infty} \frac{(m\cdot p)^{2\cdot t+1}}{(2\cdot t+1)!} = \qquad (8)$$
$$= e^{-m\cdot p}\cdot sh(m\cdot n) = \frac{e^{-m\cdot p}\cdot (e^{m\cdot p} - e^{-m\cdot p})}{2}.$$

Подстановка (8) в (6) позволяет выразить апостериорную вероятность P_j ошибки кратности j при возникновении ошибки нечетной кратности в виде:

$$P_j = \frac{e^{-m\cdot p}\cdot (m\cdot n)^j \cdot 2}{j!\cdot e^{-m\cdot p}\cdot (e^{m\cdot p} - e^{-m\cdot p})} = \frac{(m\cdot n)^j}{j!\cdot sh(m\cdot p)}. \qquad (9)$$

Вероятность $P_{>1}$ риска неверной классификации ошибки нечетной кратности, большей единицы, как однократной, при использовании *CRC* и *WCS*(h=2), определяется суммой вероятностей ошибок, кратностью больше единицы:

$$P_{>1} = 1 - P_1 = 1 - \frac{m\cdot p}{sh(m\cdot p)}. \qquad (10)$$

Аналогично, $P_{>3}$ риска неверной классификации ошибки нечетной кратности, большей трех, как однократной ошибки, имеющего место при использовании *WCS*(h=4) определяется суммой вероятностей ошибок, кратностью больше трех:

$$P_{>3} = 1 - P_1 - P_3 =$$
$$= 1 - \frac{m\cdot p}{sh(m\cdot p)} - \frac{(m\cdot p)^3}{6\cdot sh(m\cdot p)}. \qquad (11)$$

В табл. 1 приведены результаты вычисления $P_{>1}$ и $P_{>3}$ по формулам (10) и (11) для значений от $m\cdot p$=0.5 (соответствует ситуации передачи без ошибок, в среднем, половины блоков) до $m\cdot p$=0.001 (с ошибками передается, в среднем, один блок из тысячи).

Таблица 1. Вероятности $P_{>1}$ и $P_{>3}$ риска неверной классификации однократной ошибки для *WCS*(h=2) и *WCS*(h=4)

$m\cdot p$	$P_{>1}$	$P_{>3}$
0.5	0.0405	0.0005
0.45	0.0330	0.0003
0.4	0.0262	0.0002
0.35	0.0201	0.0001
0.3	0.0148	$6.7\cdot10^{-5}$
0.25	0.0103	$3.2\cdot10^{-5}$
0.2	0.0066	$1.3\cdot10^{-5}$
0.15	0.0037	$4.2\cdot10^{-6}$
0.1	0.0017	$8.3\cdot10^{-7}$
0.05	0.0004	$5.2\cdot10^{-8}$
0.01	$1.7\cdot10^{-5}$	$8.3\cdot10^{-11}$
0.001	$1.7\cdot10^{-7}$	$8.3\cdot10^{-15}$

Вероятность риска неправильной коррекции одиночной ошибки при использовании *WCS* или *CRC* может быть сведена к нулю путем повторной передачи одного бита, номер e которого определяется по коду Δ.

Если повторно переданный один бит r_e не совпадает с кодом b_e, то есть: $r_e \neq b_e$, то имело место искажение одного бита с номером e. Ошибка исправляется инвертированием соответствующего бита. Если $r_e = b_e$, то произошла ошибка, нечетной кратности большей единицы и коррекция искаженных битов производится следующим образом.

Организуется повторная передача битов блока до тех пор, пока не будут обнаружены два неверно переданные биты, с номерами e и q. Если $W_x=W_e\oplus W_q\oplus Z$. Повторно передается бит с номером x. Если полученный при этом бит r_x не совпадает с кодом b_x, то есть: $r_x\neq b_x$, то инвертируются биты с номерами e, q и x. В противном случает ($r_x=b_x$) продолжается повторная передача всего блока.

При использовании предложенной процедуры коррекции ошибок, среднее число R_o повторно переданных битов при возникновении ошибки нечетной кратности определяется формулой:

$$R_o = P_1 + \frac{P_3 \cdot m}{2} + m \cdot (1 - P_1 - P_3) = \\ = m - (m-1) \cdot P_1 - \frac{m \cdot P_3}{2} \quad . \quad (12)$$

Например, в случае передачи блока 1 Кбит по проводной линии, для которой $p=10^{-4}$ [4] среднее число повторно передаваемых бит для исправления ошибки нечетной кратности составляет всего 2.

Организация исправления ошибок четной кратности

При возникновении двукратной ошибки, то есть искажения в процессе передачи e-того и q-того битов контролируемого блока, код Δ разности WCS приемника и передатчика содержит $\delta_1=0$ и $Z=W_e\oplus W_q\neq 0$. При $h>3$ код Z позволяет однозначно определить номера искаженных битов и коррекция может быть выполнена без повторной передачи. Технологически, для получение номеров e и q по коду Z необходимо выполнять дополнительные преобразования, либо обращаться к таблице $T(Z, e, q)$.

Если для обнаружения ошибок используется $WCS(h=2)$, то коррекция двукратной ошибки выполняется путем повторной передачи части блока до тех пор, пока не будет обнаружен первый искаженный бит $r_e\neq b_e$. Затем вычисляется позиция q второго искаженного бита: $q=W_q=Z\oplus W_e$. После этого формируется запрос на повторную передачу q-го бита. Если переданный принятый в результате бит r_q не равен ранее принятому одноименному биту b_q: $r_q\neq b_q$, то позиции искаженных битов локализованы и коррекция на этом закончена. В противном случае, если $r_q=b_q$, кратность ошибок больше 2-х и повторно передается весь блок.

Вполне очевидно, что при использовании $WCS(h=2)$ для коррекции двукратной ошибки требуется повторная передача, в среднем, $m/3$ битов блока.

Проведя выкладки, аналогичные выполненные в предыдущем разделе можно показать, что при возникновении ошибки четной кратности, вероятность P_l того, что кратность ошибки равна l, определяется следующей формулой:

$$P_l = \frac{(m \cdot p)^l}{l! \cdot (ch(m \cdot p) - 1)} . \quad (13)$$

Анализ эффективности

Предложенный подход позволяет уменьшить число повторно передаваемых битов при исправлении ошибок, кратность которых не превышает 3-х.

При традиционном способе исправления ошибок, заключающемся в повторной передаче блока при возникновении ошибок среднее число M_0 передаваемых битов (без учета контрольных битов, число которых не изменяется) определяется формулой:

$$M_0 = \sum_{u=1}^{m} e^{-m \cdot p} \cdot u \cdot m \cdot (1 - e^{-m \cdot p})^{u-1} \approx \\ \approx e^{-m \cdot p} \cdot m \cdot \sum_{u=1}^{\infty} u \cdot (1 - e^{-m \cdot p})^{u-1} = m \cdot e^{m \cdot p} \quad . (14)$$

Предлагаемый способ исправления ошибок позволяет, за счет использования информации, содержащейся в разности взвешенных контрольных сумм приемника и передатчика, снизить количество повторно передаваемых бит. Так, при использовании $WCS(h=2)$ среднее число M_2 передаваемых бит определяется в виде:

$$M_2 = m \cdot e^{-m \cdot p} \cdot (1 + m \cdot p + \frac{2 \cdot (m \cdot p)^2}{3} + \\ + \frac{2 \cdot (m \cdot p)^3}{9} + 2 \cdot \sum_{i=4}^{m} \frac{(m \cdot p)^i}{i!}) \quad . \quad (15)$$

Эффективность предлагаемого способа коррекции ошибок передачи данных при использовании *WCS*(h=2) можно оценить коэффициентом β уменьшения среднего объема данных, пересылаемых по каналу для передачи одного блока:

$$\beta = \frac{M_0}{M_2}. \qquad (16)$$

При заданном среднем числа ошибочно переданных блоков - $m \cdot p$ значение β не зависит от длины m контролируемого блока. Значения коэффициента β уменьшения объема передаваемой информации для значений $m \cdot p$ от 0.001 (из тысячи передаваемых блоков, только один передан с ошибкой) до 0.5 (каждый второй блок передается с ошибкой), вычисленные по формулам (14-16) приведены в таблице 2.

Таблица 2. Значения коэффициента β уменьшения объема передаваемых данных при использовании предлагаемого способа исправления ошибок

$m \cdot p$	$\beta = M_0/M_2$
0.01	1.01
0.05	1.06
0.1	1.11
0.2	1.22
0.25	1.27
0.3	1.34
0.4	1.46
0.5	1.61

Например, при передаче блока объемом 1 Кбайт по линии, соответствующей модели двоичного симметричного канала, с вероятностью ошибочной передачи бита – $5 \cdot 10^{-5}$ [4], значение $m \cdot p = 0.41$ (41% блоков передаются с ошибкой). При традиционном способе исправления ошибок для успешной передачи одного блока требуется передать (без учета 14-разрядного контрольного кода *WCS*) в среднем, 1543 бита, а при предлагаемом – только 1046, то есть в β=1.475 раз меньше.

Выводы

В результате проведенных исследований предложен способ применения *WCS* для обнаружения и исправления ошибок передачи данных. Способ позволяет уменьшить объем повторно передаваемой информации при обнаружении ошибки за счет использования информации, содержащейся в коде разностей *WCS* приемника и передатчика.

Применение предложенного способа позволяет на 2-3 порядка уменьшить объем данных повторно передаваемых при обнаружении ошибки по сравнению с *CRC*, что обеспечивает увеличение на 5-15% скорости передачи.

Полученные результаты могут быть использованы в линиях передачи данных компьютерных систем и сетей.

Список литературы

1. *Скляр Б.* Цифровая связь. Теоретические основы и практическое применение. М.: Изд. дом "Вильямс", 2004. – 1104 с.

2. *Марковский А.П., Мулки Ахмед Яссин Ал Бадайнех, Корниец Е.В.* Обнаружение многократных ошибок передачи данных с использованием контрольной суммы // Труды 8-й международной научно-технической конференции "Современные информационные и электронные технологии", 21-25 мая 2007, г. Одесса. – С. 193.

3. *Klove T., Korzhik V.* Error Detecting Codes: General Theory and Their Application in Feedback Communication Systems. Norwell, MA: Kluwer, 1995. – 433 p.