

*МАРКОВСКИЙ А.П.,
МУЛКИ АХМЕД ЯССИН АЛ БАДАЙНЕХ,
ПУЯ СОЛЕЙМАНИ НЕЖАДИАН*

ОБ ОДНОМ ПОДХОДЕ К ПОВЫШЕНИЮ ЭФФЕКТИВНОСТИ ОБНАРУЖЕНИЯ И ИСПРАВЛЕНИЯ ОШИБОК ПЕРЕДАЧИ ДАННЫХ

Статья посвящена решению проблемы повышения эффективности обнаружения и исправления ошибок передачи данных за счет расширения класса гарантированно обнаруживаемых ошибок и класса ошибок, которые могут быть исправлены без повторной передачи. Для решения проблемы предложено использовать двумерную взвешенную контрольную сумму. Разработан алгоритм обнаружения и коррекции ошибок. Показано, что предложенная технология обеспечивает большую эффективность обнаружения и исправления ошибок по сравнению с CRC и традиционной двумерной контрольной суммой.

Paper is dedicated to solving the problem of increasing the efficiency of data transmission error detecting and correcting by expanding the class of guaranteed detecting errors and class of ones which can be corrected without repeat transmission. For solving the problem the utilization of two-dimension weighed check sum has been proposed. Algorithm for error detection and correction has been worked out. It has been shown that proposed techniques ensure superior data transformation error detecting and correcting effectiveness in compare to CRC and traditional two-dimension checksum.

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 спектрально 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 средств контроля ошибок передачи данных

В современных шинах (USB, PCI) компьютерных систем и большинстве сетевых линий информация передается блоками, которые могут разделяться на символы [2]. Пусть передаваемый блок B состоит из m бит: $B = \{b_1, b_2, \dots, b_m\}$, $b_l \in \{0, 1\}$, $l = 1, \dots, m$ и разделяется на $t = m/n$ n -разрядных символов. Если обозначить j -тый символ передаваемого блока в виде $X_j = \{x_1, x_2, \dots, x_n\} = \{b_{(j-1)n+1}, b_{(j-1)n+2}, \dots, b_{j \cdot n}\}$, $j = 1, 2, \dots, t$, $\forall i \in \{1, \dots, n\}$: $x_i \in \{0, 1\}$, то блок B может быть представлен в виде последовательности $t = m/n$ символов: $B = \{X_1, X_2, \dots, X_t\}$, $X_j \in \{0, 1, \dots, 2^n - 1\}$.

Для обеспечения надежности передачи цифровых данных используются две базовые технологии - прямое исправление ошибок (FEC-forward error correction) с применением корректирующих кодов и обнаружение ошибок с автоматическое повторной передачей при их обнаружении (ARQ-Automatic Repeat Request) [1]. При использовании технологии FEC контрольный код передается после каждого из символов блока, а технология ARQ предусматривает передачу контрольного кода только после всего блока.

В рамках технологии FEC для исправления ошибок используется широкая гамма корректирующих кодов, наиболее известными из которых являются коды Голея, Рида-Маллера, Боуза-Чоудхури-Хоквингема (БЧХ), турбо-коды [1]. Недостатками корректирующих кодов является относительно большой объем контрольной информации, сложность процедур кодирования и декодирования. Число контрольных разрядов быстро растет с увеличением кратности исправляемых ошибок, поэтому в рамках технологии FEC обычно используют перемешивание битов символов блока с тем, чтобы многократная ошибка возникающая, обычно, внутри символа распределялась по многим символам. Это существенно усложняет процедуры кодирования и декодирования. Поэтому, практическое использование корректирующих кодов оправдано, если интенсивность возникающих ошибок достаточно высока, а также, в случае отсутствия возможности для повторной передачи блока. Соответственно, основными сферами практического применения корректирующих кодов являются системы беспроводной связи, для которых характерным является отно-

сительно высокая интенсивность ошибок, вызванных внешними помехами, а также системы хранения информации на дисках.

Главное преимущество ARQ перед FEC заключается в том, что обнаружение ошибок производится гораздо быстрее и проще по сравнению с коррекцией ошибок. При использовании технологии ARQ объем контрольной информации, передаваемой по линии существенно меньше. В проводных, кабельных и опто-волоконных линиях передачи цифровых данных интенсивность возникновения ошибок относительно мала, поэтому использование корректирующих кодов оказывается неэффективным. Для обеспечения безошибочной передачи в таких линиях преимущественно используется технология ARQ [3].

Для обнаружения ошибок передачи данных в рамках технологии ARQ используются циклические коды (CRC- Cyclic Redundancy Check), и контрольные суммы (CS-Check Sum). Использование CRC позволяет обнаруживать все ошибки нечетной кратности, все двойные ошибки при передаче блока данных, а также "пачку" ошибок, локализованных в участке длиной не превышающего степень образующего полинома CRC [4]. Традиционная CS, вычисляемая как побитовая сумма по модулю 2 всех символов блока, позволяет гарантированно обнаруживать только все ошибки нечетной кратности. Однако, модификации CS обеспечивают существенно большую достоверность обнаружения ошибок. Так, двумерная CS (TDCS-Two Dimension Check Sum) позволяет гарантированно обнаруживать ошибки нечетной кратности и двукратные ошибки. Некоторые варианты взвешенной CS позволяют обнаруживать, кроме ошибок нечетной кратности, все ошибки кратностью 2 и 4 [5].

Существенным отличием перечисленных кодов для обнаружения ошибок передачи информации от корректирующих кодов является отсутствие функциональной зависимости между длиной m контролируемого блока и разрядностью k контрольного кода.

Как уже отмечалось выше, в современных условиях существуют эффективные факторы, действие которых имеет следствием увеличение числа возникающих ошибок в линиях передачи данных. Это требует адекватного повышения эффективности контроля правильности передачи, а также технологии исправления возникающих ошибок. В частности, рост интенсивности внешних электро-магнитных полей требует увеличения кратности гарантированно обнаруживаемых ошибок, а также повышения достоверности выявления "пачек" ошибок.

Кратность ошибок, гарантированно обнаруживаемых CRC и известными модификациями CS не превышает 4-х, поэтому необходимо создание средств с большей кратностью гарантированно обнаруживаемых ошибок. Другой важной проблемой, которая не решается с использованием существующих средств обнаружения ошибок является обеспечение эффективности исправления выявленных ошибок. CRC и известные моди-

фикации CS не позволяют локализовать часть блока, содержащей неверно переданные биты. Поэтому для исправления выявленных ошибок выполняется повторная передача всего блока. Это сопряжено во-первых, с заметными затратами времени, а во-вторых, не исключает возможности возникновения ошибок при повторной передаче. Учитывая тенденцию роста объемов передаваемых блоков, значимость указанного недостатка также увеличивается.

Целью работы является увеличение кратности гарантированно обнаруживаемых ошибок и повышение эффективности исправления выявленных ошибок.

Модификация TDCS на основе использования взвешенных контрольных сумм

Повышение кратности обнаруживаемых ошибок передачи данных необходимо связано с увеличением объема передаваемой контрольной информации.

В качестве основы для разработки эффективных средств обнаружения ошибок большой кратности выбрана TDCS. Это обусловлено тем, что стандарты большинства линий передачи вычислительных систем, в частности COM-портов и шин PCI, а также линий компьютерных сетей предусматривают пересылку бита четности на каждый передаваемый байт [2].

При контроле с использованием TDCS передаваемых блок данных можно представить в виде матрицы D , имеющей n столбцов m строк. Вместе с блоком передается m битов четности $p_{r1}^S, p_{r2}^S, \dots, p_{rm}^S$ строк и n битов четности $p_{c1}^S, p_{c2}^S, \dots, p_{cn}^S$ столбцов матрицы D^S передатчика. При этом, бит четности i -той, $i \in \{1, \dots, m\}$, строки $\{d_{i1}^S, d_{i2}^S, \dots, d_{in}^S\}$ матрицы D^S вычисляется в виде: $p_{ri}^S = d_{i1}^S \oplus d_{i2}^S \oplus \dots \oplus d_{in}^S$, а бит четности j -того, $j \in \{1, \dots, n\}$, столбца $\{d_{1j}^S, d_{2j}^S, \dots, d_{mj}^S\}$ матрицы D^S вычисляется в виде: $p_{cj}^S = d_{1j}^S \oplus d_{2j}^S \oplus \dots \oplus d_{mj}^S$. На приемнике биты четности строк $p_{r1}^R, p_{r2}^R, \dots, p_{rm}^R$ и биты четности столбцов $p_{c1}^R, p_{c2}^R, \dots, p_{cn}^R$ вычисляются по матрице D^R принятого блока и сравниваются с принятыми от передатчика. Решение о наличии ошибок передачи производится путем анализа вектора $\Delta_R = \{r_1, r_2, \dots, r_m\}$ несовпадения битов четности строк на приемнике и передатчике, а также вектора $\Delta_C = \{c_1, c_2, \dots, c_m\}$ несовпадения битов четности столбцов матриц D^R и D^C . Это означает, что каждый i -тый бит Δ_R формируется как: $r_i = p_{ri}^S \oplus p_{ri}^R$ и каждый j -тый бит Δ_C вычисляется в виде: $c_j = p_{cj}^S \oplus p_{cj}^R$.

При использовании такой TDCS можно гарантированно обнаруживать все ошибки нечетной кратности и двойные ошибки. 4-кратные ошибки,

попарно локализованные в одноименных разрядах двух строк матрицы D не обнаруживаются. Кроме обнаружения ошибок указанной выше кратности, TDCS позволяет локализовать битовую позицию однократной ошибки. Это позволяет, в принципе, использовать TDCS для исправления наиболее часто встречающихся в проводных линиях однократных ошибок. При этом существует риск не различить однократную ошибку от пятикратной. Действительно, если две пары ошибок локализованы в одинаковых битах двух строк, а пятая ошибка локализована в j -том бите i -той строки, то по два ошибочных бита q_1 и q_2 локализованы в строке и их строковые биты четности равны нулю.

Основным недостатком описанной традиционной TDCS является малая кратность гарантированно обнаруживаемых ошибок.

Для повышения достоверности обнаружения ошибок передачи данных и повышения эффективности их исправления с использованием двумерной контрольной суммы предлагается ее модификация, отличающаяся наличием контроля с использованием взвешенных сумм битов столбцов матрицы D . Предлагаемая модификация предполагает вычисление для каждого j -го столбца матрицы D^S h -разрядного кода S_j^S взвешенной контрольной суммы $S_j^S = d_{1j}^S \cdot W_1 \oplus d_{2j}^S \cdot W_2 \oplus \dots \oplus d_{mj}^S \cdot W_m$, где W_1, W_2, \dots, W_m – h -разрядные коды весовых коэффициентов строк матрицы D . В простейшем случае $h = \log_2 m$, а W_i – двоичный код номера i строки. Аналогичным образом вычисляется значение h -разрядного кода S_j^R взвешенной суммы для j -го столбца матрицы D^R приемника.

Вычисленные на передатчике взвешенные контрольные суммы $S_1^S, S_2^S, \dots, S_n^S$ передаются на приемник, где сравниваются с взвешенными контрольными кодами $S_1^R, S_2^R, \dots, S_n^R$, вычисленными на приемнике. Результатом сравнения является набор Δ_V h -разрядных кодов $\{V_1, V_2, \dots, V_n\}$, каждый j -тый из которых вычисляется как поразрядная сумма по модулю 2 соответствующих взвешенных сумм передатчика и приемника $V_j = S_j^S \oplus S_j^R$. Передача считается выполненной без ошибок, если все компоненты множеств $\Delta_R = \{r_1, r_2, \dots, r_m\}$, $\Delta_C = \{c_1, c_2, \dots, c_m\}$ и $\Delta_V = \{V_1, V_2, \dots, V_n\}$ равны нулю. Суммарное число k передаваемых контрольных бит в предлагаемой модификации TDCS определяется формулой:

$$k = m + n(1 + \log_2 m) \quad (1)$$

Введение взвешенных контрольных сумм столбцов матрицы D позволяет гарантированно обнаруживать двойную ошибку в столбце. Действительно, если ошибка происходит в q -том и g -том, $q, g \in \{1, \dots, m\}$, битах j -того столбца матрицы D , то $V_j = W_q \oplus W_g \neq 0$, в силу того, что не равна нулю сумма по модулю 2 двух различных кодов.

Анализ возможностей модифицированной TDCS по обнаружению многократных ошибок

Предложенная модификация TDCS позволяет гарантированно обнаруживать все ошибки нечетной кратности, поскольку при возникновении таких ошибок не менее, чем один из битов $\{r_1, r_2, \dots, r_m\}$ и не менее чем один из битов $\{c_1, c_2, \dots, c_m\}$ равен единице.

Для того, чтобы ошибка четной кратности не была обнаружена при использовании предлагаемой TDCS, то есть, для того, чтобы ее проверочные коды $\Delta_R = \{r_1, r_2, \dots, r_m\}$, $\Delta_C = \{c_1, c_2, \dots, c_m\}$ и $\Delta_V = \{V_1, V_2, \dots, V_n\}$ не были равны нулю необходимо выполнение 3-х условий:

1. Для того, чтобы $\Delta_R = \{r_1, r_2, \dots, r_m\}$ не содержал единичных компонент, необходимо, чтобы число ошибочных бит в каждой из искаженных строк матрицы D^R было четным.
2. Для того, чтобы $\Delta_C = \{c_1, c_2, \dots, c_m\}$ не содержал единичных компонент, необходимо, чтобы число ошибочных бит в каждом из искаженных столбцов матрицы D^R было четным.
3. Для того, чтобы все компоненты $\Delta_V = \{V_1, V_2, \dots, V_n\}$ были равны нулю, необходимо, чтобы число ошибочных бит в каждом из искаженных столбцов матрицы D^R было больше 2-х.

Из первых двух условий следует, что кратность ошибок должна нацело делиться на 4, а из третьего следует, что она должна быть больше 4-х. Минимальная кратность ошибок, удовлетворяющих приведенным выше условиям равна 8-ми. Действительно, если 8 ошибочных бит локализованы попарно в разрядах x_1 и x_2 4-х строках с номерами y_1, y_2, y_3, y_4 , то все битовые компоненты $\Delta_R = \{r_1, r_2, \dots, r_m\}$ и $\Delta_C = \{c_1, c_2, \dots, c_m\}$ равны нулю и существует ненулевая вероятность того, что сумма 4-х весовых коэффициентов равна нулю: $V_{x_1} = W_{y_1} \oplus W_{y_2} \oplus W_{y_3} \oplus W_{y_4} = 0$. Количество E_8 вариантов таких локализаций определяется формулой:

$$E_8 = \binom{m}{4} \cdot \binom{n}{2} \quad (2)$$

Поскольку сумма по модулю 2 4-х h -разрядных двоичных кодов $V_{x_1} = W_{y_1} \oplus W_{y_2} \oplus W_{y_3} \oplus W_{y_4}$ может с равной вероятностью принимать $2^h = m$ значений, то вероятность p_{E8} того, что эта сумма равна нулю составляет: $p_{E8} = 1/m$. Из этого и выражения (2) следует, что вероятность P_{N8} того, что возникшая 8-кратная ошибка не будет обнаружена определяется в виде:

$$P_{N8} = \frac{E_8 \cdot p_{E8}}{\binom{m \cdot n}{8}} \approx \frac{m^4 \cdot n^2 \cdot 8!}{m^9 \cdot n^8 \cdot 4! \cdot 2} = \frac{480}{(m \cdot n)^5 \cdot n} \quad (3)$$

Например, при контроле правильности передачи блока длиной 1 Кбайт ($n=8, m=1024$), вероятность P_{N8} того, что возникшая 8-кратная ошибка не будет обнаружена составляет $3 \cdot 10^{-18}$.

Рассуждая аналогичным образом, несложно показать, что для 12-кратной ошибки, вероятность P_{N12} того, что она не будет обнаружена определяется формулой:

$$P_{N12} = \frac{\binom{m}{6} \cdot \binom{n}{2} \cdot m^{-1}}{\binom{m \cdot n}{12}} \approx \frac{m^6 \cdot n^2 \cdot 12!}{m^{13} \cdot n^{12} \cdot 6! \cdot 2} = \frac{332640}{(m \cdot n)^7 \cdot n^3} \quad (4)$$

Так, для $n=8, m=1024$, в соответствии с (3), вероятность $P_{N8} = 2.6 \cdot 10^{-25}$. Для больших кратностей 16, 20, ... ошибок, вероятности того, что они не будут обнаружены, быстро убывают.

Таким образом, по сравнению с CRC и обычной TDCS, использование предложенной модификации TDCS позволяет существенно расширить класс гарантированно обнаруживаемых ошибок за счет безусловного выявления ошибок кратности 4 и 6.

С практической точки зрения важной является оценка возможности гарантированного обнаружения "пачки" ошибок, вызванной внешними помехами и наводками. Анализ сформулированных выше трех условий при выполнении которых ошибки не обнаруживаются, позволяет сделать вывод о том, что "пачка" битовых ошибок гарантированно обнаруживается, если ее длина не превышает $4 \cdot n - 1$. Следовательно, по возможности обнаружения "пачек" ошибок предложенная модификация TDCS практически соответствует CRC-32 и существенно превышает обычную TDCS, которая позволяет гарантированно обнаруживать "пачки" ошибок, длина которых не превышает $4 \cdot n - 1$.

Организация исправления ошибок передачи данных в TDCS на основе взвешенных сумм

Использование модифицированной TDCS позволяет существенно расширить класс ошибок, которые могут быть исправлены без повторной передачи всего блока.

При организации исправления возникших ошибок возможными являются три режима:

- исправление без передачи дополнительной информации, то есть на основе наличных проверочных кодов $\Delta_R = \{r_1, r_2, \dots, r_m\}$, $\Delta_C = \{c_1, c_2, \dots, c_m\}$ и $\Delta_V = \{V_1, V_2, \dots, V_n\}$;
- формирование требований на передачу отдельных фрагментов блока и исправление выявленных ошибок на их основе.
- формирование требования на повторную передачу всего блока.

Для исправления возникших ошибок передачи блока предлагается следующий алгоритм обработки на приемнике проверочных кодов $\Delta_R = \{r_1, r_2, \dots, r_m\}$, $\Delta_C = \{c_1, c_2, \dots, c_n\}$ и $\Delta_V = \{V_1, V_2, \dots, V_n\}$. Алгоритм реализует коррекцию тех ошибок, которые могут быть исправлены с использованием векторов Δ_R , Δ_C и Δ_V , а также составление списка Ξ скорректированных строк матрицы D^R . Предлагаемый алгоритм включает две базовых процедуры:

А. Процедуру коррекции ошибок по проверочным кодам Δ_R , Δ_C и Δ_V .

В. Процедуру коррекции ошибок за счет повторной передачи.

Процедура **А** коррекции ошибок по проверочным кодам Δ_R , Δ_C и Δ_V , в свою очередь, состоит из двух циклов:

А.1. Цикла коррекции ошибок нечетной кратности в строках матрицы D^R .

А.2. Цикла коррекции одиночных ошибок в столбцах матрицы D^R .

Реализация первого из перечисленных циклов – **А.1** состоит в последовательном выполнении следующей последовательности действий:

1. Построить множество Ω номеров строк, у которых проверочный бит четности равен единице: $r_i = 1: i \in \Omega$. $\Xi = \emptyset$. Установить счетчик s скорректированных битов в нуль: $s = 0$.

2. Сформировать множество $\Theta = \{\Theta_1, \Theta_2, \dots, \Theta_t\}$ всех возможных подмножеств множества Ω . Для каждого u -го из подмножеств Θ , $u \in \{1, \dots, t\}$, сформировать:

- бит c'_u , равный единице, если число элементов Θ_u – нечетное и равный нулю, если количество элементов Θ_u – четное.
- h -разрядный код $V'_u = \bigoplus_{g \in \Theta_u} W_g$.

3. Установить в нуль флаг f корректирования матрицы D^R - $f := 0$.

4. Установить индекс столбца $j := 1$.

5. Если $c_j = 0$ и $V_j = 0$ перейти на пп.9.

6. Установить индекс $u := 1$.

7. Если $\Theta_u \neq \emptyset$, $c_j = c'_u$ и $V_j = V'_u$ то:

Перебрать все $t-1$ возможные пары $\langle \Theta_u, \Theta_l \rangle$, $l \in \{1, \dots, t\}$ подмножеств множества Θ . Если среди подмножеств Θ нет пары $\langle \Theta_u, \Theta_l \rangle$ для которой выполняются условия: $c'_u = c'_l$ и $V'_u = V'_l$ то выполнить коррекцию:

7.0. Установить $f := 0$.

7.1. Инвертировать j -тый разряд строк матрицы D^R , номера которых принадлежат множеству Θ_u : $\forall g \in \Theta_u: d_{gj}^R = 1 \oplus d_{gj}^R$.

7.2. Установить в нуль c_j и V_j .

7.3. Включить в список Ξ номера всех строк, принадлежащих Θ_u : $\Xi = \Xi \cup \Theta_u$.

7.4. Исключить из множества Θ все подмножества, которые имеют общие номера строк с подмножеством \mathcal{G}_u : $\forall l=1, \dots, t$: если $\mathcal{G}_l \cap \mathcal{G}_u \neq \emptyset \Rightarrow \Theta = \Theta - \mathcal{G}_l$. Скорректировать значение t .

7.5. Установить в нуль проверочные биты четности строк, номера которых принадлежат \mathcal{G}_u : $\forall g \in \mathcal{G}_u: r_g = 0$.

7.6. Увеличить значение счетчика s на число, соответствующее количеству элементов множества \mathcal{G}_u ; перейти на пп.9.

8. Если $u < t$: $u := u + 1$, возврат на пп.7.

9. Если $j < n$: $j := j + 1$, возврат на пп.5.

10. Если матрица D^R корректировалась, то есть $f=1$, возврат на пп.3., иначе конец цикла **A.1**.

Реализация цикла **A.2** коррекции ошибок единичной кратности в столбцах матрицы D^R состоит в выполнении следующей последовательности действий:

11. Установить флаг ϕ коррекции в цикле **A.2**. в нуль: $\phi := 0$.

12. Установить счетчик j цикла столбцов в единицу: $j = 1$.

13. Если $c_j = 0$ и $V_j = 0$ перейти на пп.15.

14. Если $c_j = 1$ и $V_j \neq 0$, то :

14.1. Инвертировать j -тый разряд строки матрицы D^R с номером V_j :
 $d_{V_j}^R = 1 \oplus d_{V_j}^R$

14.2. Установить в нуль c_j, V_j : $c_j = 0, V_j = 0$

14.3. Включить в список Ξ номер V_j скорректированной строки: $\Xi = \Xi \cup V_j$.

14.4. Установить $\phi := 1$.

14.5. Инвертировать бит четности скорректированной строки $r_{V_j} = 1 \oplus r_{V_j}$.

14.6. Инкрементировать значение счетчика s числа исправленных битов.

15. Если $j < n$: $j := j + 1$, возврат на пп.12.

16. Если $\phi = 1$, возврат на пп.1, иначе конец цикла **A.2**.

После завершения выполнения процедуры **A**, реализуется процедура **B** повторной передачи отдельных строк или всего блока, которая осуществляется в следующем порядке:

1. Если все компоненты Δ_R, Δ_C и Δ_V равны нулю, то, в зависимости от значения s – количества скорректированных бит матрицы D^R выполнить запрос на повторную передачу строк, номера которых составляют множество Ξ .

2. Если существуют не равные нулю компоненты Δ_R, Δ_C и Δ_V , выполнить запрос на повторную передачу всего блока.

3. По получении копий строк, составляющих множество Ξ , сравнить их с одноименными скорректированными строками матрицы D^R : если они совпадают, то исправления выполнены правильно, если не совпадает – выполнить запрос на повторную передачу блока.

Предложенный алгоритм иллюстрируется следующими примерами. Пусть, передается блок данных объемом 16 байт, причем вместе с каждым байтом передается его бит четности. Соответственно матрица D содержит 16 строк ($m=16$) по 8 бит в каждой ($n=8$). Рассмотрим работу описанного выше алгоритма для нескольких вариантов локализации ошибок.

Пример 1. Пусть позиции двух искаженных бит в матрице D задаются двумя тремя парами координат: $\langle 3,3 \rangle, \langle 7,3 \rangle$, причем первая цифра соответствует номеру строки, а вторая – номеру столбца. Фактически искажен третий бит двух строк: 3-й и 7-й матрицы D^R . В этом случае на приемнике сформируются следующие не равные нулю проверочные коды: $r_3=1, r_7=1, V_3=W_3 \oplus W_7=0011 \oplus 0111=0100$.

В процессе выполнения цикла **A** строится множество $\Omega=\{3,7\}$ и его подмножества $\mathcal{Q}_1=\{3\}, \mathcal{Q}_2=\{7\}, \mathcal{Q}_3=\{3,7\}$, то есть $t=6$. Вычисляются $c_1'=1, V_1'=0011, c_2'=1, V_2'=0111, c_3'=0, V_3'=0100$. Очевидно, что операции пп.7 коррекции ошибок будут выполняться только при $j=3, u=3, c_j=c_u'=0, V_j=V_u'$. В п.7.1. инвертируются третьи биты строк 3 и 7 матрицы D^R . В пп.7.2 устанавливается в нуль V_3 . В пп.7.3 в множество Ξ дописываются номера скорректированных строк: $\Xi=\{3,7\}$. В пп.7.5 обнуляются биты r_3 и r_7 .

Поскольку матрица D^R корректировалась и $f=1$, то цикл **A1** выполняется еще раз. Поскольку все проверочные коды равны нулю, то цикл реализуется без коррекции D^R . Последующий цикл **A2** также выполняется без коррекции, поскольку $\forall j \in \{1, \dots, 16\}: c_j=0$ и $V_j=0$.

В рамках процедуры **B**, в зависимости от значения s , можно осуществить запрос на повторную передачу 3-й и 7-й строк матрицы D и по их получении проверить правильность выполненной выше коррекции ошибок в означенных строках.

Пример 2. Пусть позиции трех искаженных бит в матрице D задаются двумя тремя парами координат: $\langle 3,3 \rangle, \langle 5,3 \rangle, \langle 5,7 \rangle$. Фактически искажен третий бит двух строк: 3-й и 5-й матрицы D^R , а также 7-й бит 5-й строки. В этом случае на приемнике сформируются следующие, не равные нулю проверочные коды: $r_3=1, V_3=W_3 \oplus W_5=0011 \oplus 0101=0110$, а также $c_7=1, V_7=W_5=0101$.

В процессе выполнения цикла **A.1** строится множество $\Omega=\mathcal{Q}_1=\{3\}, t=1, c_1'=1, V_1'=W_3=0011$. Однако, коррекции матрицы D^R , в этом цикле не производится, поскольку не существует пары $\langle j, u \rangle, j \in \{1, \dots, 16\}, u=1: c_j=c_j', V_j=V_1'$. При реализации цикла **A.2**. при $j=7$ выполняется условие пп.14: $c_7=1, V_7=W_5=0101 \neq 0$. Соответственно, осуществляется коррекция 7-го бита строки матрицы D^R с номером $V_7=5$ (пп.14.1). В пп.14.2 c_7, V_7 устанавливаются в нуль. В список Ξ дописывается номер скорректированной строки: $\Xi=\{5\}$. Флаг ϕ устанавливается в единицу. Инвертируется проверочный бит r_5 четности 5-той строки: $r_5=1$.

Поскольку $\varphi=1$, по пп.16 осуществляется возврат на повторное выполнение цикла **A1**. При переходе к выполнению этого цикла установлены в единицу проверочные биты четности 3-й и 5-той строк: $r_3=1$ и $r_5=1$. Проверочные коды всех столбцов, кроме 3-го равны нулю. Для 3-го столбца $c_3=0$, $V_3 = 0110$. Эта ситуация полностью аналогична описанной выше в рамках примера 1. Соответственно корректировка двух искаженных битов производится по описанному в примере 1 способу. После выполнения цикла **A1** множество $\Xi=\{3,5\}$. Эти строки могут быть переданы повторно для проверки правильности выполненной коррекции.

Анализ предложенного алгоритма коррекции ошибок показывает, что он позволяет гарантированно корректировать без повторной передачи искажения до 4-х бит передаваемого блока. Единственным вариантом локализации 4-кратных ошибок, при котором не обеспечивается их исправление без повторной передачи является ситуация, когда их матричные координаты могут быть представлены в виде: $\langle b_1, e_1 \rangle, \langle b_1, e_2 \rangle, \langle b_2, e_1 \rangle, \langle b_2, e_2 \rangle$, $b_1, b_2 \in \{1, \dots, m\}$, $e_1, e_2 \in \{1, \dots, n\}$. Поскольку количество вариантов такой локализации равно $m \cdot (m-1) \cdot n \cdot (n-1)/4$, то вероятность P_{NC4} появления такой локализации четверки искаженных бит определяется формулой:

$$P_{CN4} = \frac{m \cdot (m-1) \cdot n \cdot (n-1) \cdot 3!}{m \cdot n \cdot (m \cdot n - 1) \cdot (m \cdot n - 2) \cdot (m \cdot n - 3)} \approx \frac{6}{m^2 \cdot n^2} \quad (5)$$

Анализ выражения (5) свидетельствует о том, что практически все 4-кратные ошибки также могут быть исправлены без повторной передачи.

Анализ достоверности коррекции ошибок

Исправление ошибок на основе анализа на приемнике проверочных кодов $\Delta_R = \{r_1, r_2, \dots, r_m\}$, $\Delta_C = \{c_1, c_2, \dots, c_m\}$ и $\Delta_V = \{V_1, V_2, \dots, V_n\}$, как и при использовании всех корректирующих кодов, сопряжено с риском неверной классификации ошибок. Так, однократная ошибка передачи бита, локализованного в строке x и столбце y матрицы D , при которой $r_y=1$, $c_x=1$ и $W_x=y$ по значениям проверочных кодов не отличима от 7-кратной ошибки, при которой 4 искаженных бита локализованы в одном столбце z , причем одна из этих четырех ошибок находится в строке y , а остальные три – в строках y_1, y_2, y_3 , а другие три искаженных бита находятся в столбце x и строках y_1, y_2, y_3 . Очевидно, что при такой локализации 7-ми искажаемых битов $c_z=0$, $r_{y1}=0$, $r_{y2}=0$, $r_{y3}=0$, $c_x=1$, $r_y = 1$, а $W_z = W_{y1} \oplus W_{y2} \oplus W_{y3} \oplus W_x$. Существует вероятность (численно равная $1/m$) того, что $W_z = 0$, тогда $W_z = W_{y1} \oplus W_{y2} \oplus W_{y3} = y$ и тогда такая 7-кратная ошибка по значению проверочных кодов не отличима от однократной. Проведя рассуждения, аналогичные использованным при выводе формулы (1) можно показать, что вероятность P_{S7} того, что возникшая 7-кратная ошибка будет иметь

проверочные коды совпадающие с однократной определяется выражением:

$$P_{s7} = \frac{105}{(m \cdot n)^4 \cdot n} \quad (6)$$

Так, для $n=8$, $m=1024$, в соответствии с (6), вероятность $P_{s7} = 2.9 \cdot 10^{-15}$. Необходимо отметить, что рассмотренная конфигурация 7-кратной ошибки не может возникнуть в "пачке", вызванной внешней помехой. С учетом того, вероятность возникновения 7-кратной ошибки, в проводных линиях передачи данных достаточно мала.

Аналогично можно показать, при коррекции двукратной ошибки существует риск, связанный с тем, что такая ошибка, исходя из проверочных кодов не отличима от 6-кратной. Вероятность P_{s6} такого риска определяется вероятностью специфической локализации в матрице D 6-ти искаженных бит, которая определяется формулой:

$$P_{s6} = \frac{60}{(m \cdot n)^3 \cdot n} \quad (7)$$

Соответственно, при коррекции 3-кратной ошибки существует риск не отличить ее от 5-кратной. Сопоставление формул (6) и (7) показывает, что вероятность риска неверной коррекции на основе проверочных кодов растет с увеличением количества s скорректированных битов матрицы D^R .

Устранение риска неверной коррекции достигается повторной передачей скорректированных строк матрицы D . Целесообразность повторной передачи определяется значением s , зависимостью вероятностей появления ошибок от их кратности для конкретного типа канала, требованиям к достоверности передачи данных.

Для оценки эффективности предложенного алгоритма коррекции ошибок проводилось статистическое моделирование его работы. Для матрицы D , состоящей из 32-х строк и 8-ми столбцов сгенерировано 40000 ошибок кратностей от 1 до 8. Полученные результаты приведены в таблице 1. Анализ результатов экспериментальных исследований показывает, что предложенный алгоритм коррекции без повторной передачи исправленных строк позволяет надежно корректировать ошибки, кратность которых не превышает 3-х.

Табл. 1. Результативность алгоритма коррекции ошибок без повторной передачи исправленных строк

Результат работы алгоритма коррекции	Кратность ошибок							
	1	2	3	4	5	6	7	8
Исправлены без повторной передачи (%)	100	100	100	99	94	81	58	38
Исправлено за счет повторной передачи (%)	0	0	0	1	5.5	16	35	50
Неверно исправлены без повторной передачи (%)	0	0	0	0	0.5	2	7	12

Выводы

В результате проведенных исследований, направленных на повышение эффективности обнаружения и коррекции ошибок передачи данных, предложена модификация TDCS. За счет использования взвешенных контрольных сумм, применение предложенной TDCS позволяет расширить класс гарантированно обнаруживаемых ошибок по сравнению с традиционной TDCS и CS. Разработан алгоритм коррекции ошибок для предложенной TDCS без повторной передачи блока данных. Выполненный теоретический анализ и экспериментальные исследования показали его высокую эффективность для исправления ошибок, кратность которых не превышает 3-х.

Предложенная модификация TDCS позволяет повысить эффективность обнаружения и коррекции ошибок в линиях с передачей бита четности.

Список использованной литературы:

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: Издательский дом "Вильямс", 2004.- 1104 с.
2. Петров С.В. Шины PCI, PCI Express. Архитектура, дизайн, принципы функционирования.-СПб.:БХВ-Петербург.-2006.-416 с.
3. Дансмор Б., Скандьер Т. Справочник по телекоммуникационным технологиям. М.: Издательский дом "Вильямс", 2004.-640 с.
4. Klove T., Korzhik V. Error Detecting Codes: General Theory and Their Application in Feedback Communication Systems. Norwell, MA: Kluwer, 1995. – 433 p.
5. Али Тауфик Окла Аль-Хавальди, Мулки Ахмед Яссин Ал Бадайнех, Антоненко А.А. Об одном подходе к повышению надежности ошибок передачи данных методом контрольных сумм // Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка, – Київ: ВЕК+ – 2006 – № 45. с.27-35.