УДК 004.056

Марковський А.П., к.т.н.,
Клименко И.А., к.т.н.,
Иванов А.Н.

СПОСОБ ЭФФЕКТИВНОЙ КОРРЕКЦИИ "ПАЧКИ" ОШИБОК В КАНАЛАХ С КОДОВО-ИМПУЛЬСНОЙ МОДУЛЯЦИЕЙ

Национальный технический университет Украины "КПИ"

Предложен простой и эффективный способ коррекции одиночной "пачки" ошибок, возникающих при передаче данных в каналах с кодово-импульсной модуляцией. В основе способа лежит использование двухкомпонентной взвешенной контрольной суммы, при вычислении которой применяются логические операции. Доказано, что предложенный способ позволяет решать задачу коррекции одиночной "пачки" ошибок эффективнее кодов Рида-Соломона за счет существенного уменьшения вычислительной сложности при использовании меньшего числа контрольных разрядов.

Введение

Развитие технологий передачи цифровых данных претерпело качественное изменение. На порядки выросли скорости передачи данных, прогресс цифровой обработки сигналов дал возможность реализовать сложные технологии спектрального уплотнения данных для наиболее полного использования возможностей каналов, стабильный характер приобретает тенденция расширения использования беспроводных линий. Достигнутый прогресс в области интегральной технологии позволил существенно усложнить и интеллектуализировать протоколы обмена данными на всех уровнях их компьютерной обработки.

Динамическое развитие средств передачи и обмена данными существенно влияет на характер ошибок, которые возникают в процессе передачи цифровой информации.

Каналы передачи данных с импульсно-кодовой модуляцией широко применяются на практике [2]. Характерной особенностью таких линий является отсутствие модулирования цифровых данных высокочастотным синусоидальным сигналом - соответственно отсутствует и физическое разделение данных на символы - группы битов, которые модулируются одним несущим сигналом. Схожая ситуация имеет место и в дисковых накопителях, в частности, при использовании *RLL*-модуляции. В последние годы доминирующим типом ошибок в таких линиях и дисковых накопителях стали "пачки" битовых искажений.

Поэтому, важность задачи исправления "пачек" ошибок в современных условиях возрастает.

С другой стороны, меняется акценты критериев эффективности средств контроля и исправления возникающих при передаче ошибок. В частности, в условиях динамического роста объемов и скорости передачи снижается значимость такого традиционного критерия эффективности средств обнаружения и коррекции ошибок, как количество используемых контрольных разрядов. И напротив, в условиях увеличения пропускной способности линий весьма важно обеспечить контроль и обработку ошибок в темпе передачи данных - это объективно имеет следствием рост значимости таких критериев эффективности, как вычислительная и временная сложность процедур, связанных с обнаружением и коррекцией ошибок [2].

Таким образом, задача повышения эффективности обнаружения и исправления ошибок, возникающих при передаче данных между компонентами компьютерных систем, а также при их хранении на оптических и магнитных дисках является актуальной.

Анализ существующих средств коррекции "пачки ошибок"

При высокой скорости передачи под влиянием внешней помехи возникают ошибки, которые носят характер "пачки".

При этом понятие "пачки ошибок" (*burst errors*) определяется [2] следующим образом. Считается, что время воздействия внешней помехи ограничено величиной t_n. За максимальное время t_n воздействия помехи на линию по ней передается m бит. Под воздействием внешней помехи передаваемые биты с высокой вероятностью (до 0.5) подвергаются искажению. Исходя из этого, "пачкой" называется группа из m бит, передаваемых во время действия внешней помехи, каждый из которых с высокой вероятностью подвержен искажению [2]. В работе описанная модель "пачки" конкретизирована в следующем виде: под "пачкой ошибок" понимается группа из m бит, первый из которых искажен. Поэтому, в современных условиях, для широкого класса линий передачи данных доминирующим типом ошибок становится "пачка".

Анализ публикаций, посвященных возникновению "пачек ошибок" указывают на то, что всегда можно указать максимальное число m бит в "пачке". Большинство моделей [2] исходят из того, что возникновение "пачек ошибок" при передаче блока данных носит биномиальный характер. Это означает, что при передаче блока наиболее вероятным является отсутствие "пачек ошибок". Вероятность того, что передаче блока возникнет одна "пачка ошибок" на порядки меньше. Вероятность возникновения двух "пачек ошибок" на порядки меньше вероятности возникновения одной "пачки". Поэтому на практике ограничивают кратность возникающих при передаче блока "пачек ошибок" одной [2].

Применительно к задаче коррекции "пачки" ошибок наиболее эффективным корректирующим кодом является код Рида-Соломона [2]. Этот код предполагает разделение контролируемого блока данных на n символов разрядностью m бит. С использованием $2 \cdot h$ контрольных символов код Рида-Соломона позволяет корректировать любые битовые искажения в h символах.

Код Рида-Соломона представляет собой недвоичный циклический код существенным недостатком которого является большая вычислительная и временная сложность, которая практически исключает возможность контроля ошибок в темпе передачи с использованием программных средств. Обычно, кодирование и декодирование с использованием кодов Рида-Соломона реализуется специальными аппаратными средствами.

Для коррекции такой ошибки с использованием кодов Рида-Соломона надо гарантировать возможность коррекции 2-х (h=2) m-битовых символов, поскольку "пачка ошибок" может возникнуть начиная с любого бита.

Для вычисления контрольного кода Рида-Соломона надо выполнить деление полинома информационного сообщения на генерирующий полином поля. Операция деления полинома степени n на полином $g(X)$ степени 4 требует выполнения $n \cdot 4$ операций умножения символов и такого же количества операций символьного сложения. Таким образом, время T_c кодирования при использовании кода Рида-Соломона определяется формулой:

$$T_c = 4 \cdot n \cdot (t_m + t_{XOR}), \tag{1}$$

где t_m - время умножения на поле Галуа двух m-разрядных символов, t_{XOR} - время сложения на поле Галуа двух символов. Учитывая, что $t_m \approx 2 \cdot m \cdot t_{XOR}$, выражение (1) можно упростить:

$$T_c \approx 8 \cdot n \cdot m \cdot t_{XOR}. \tag{2}$$

Для обнаружения ошибок в коде Рида-Соломона, рассчитанного на исправление 2-х символов вычисляется 4 синдрома: $S_1=R(\alpha), S_2=R(\alpha^2), \ldots, S_{2 \cdot h}=R(\alpha^{2 \cdot h})$. Если значения всех синдромов равно нулю, то полагается, что сообщение передано без ошибок. При вычислении каждого из синдромов производится n операций умножения и такое же количество операций сложения. Поэтому, время T_O обнаружения ошибок оценивается следующим выражением:

$$\begin{aligned} T_o &= 4 \cdot n \cdot (t_m + t_{XOR}) \approx \\ &\approx 8 \cdot n \cdot m \cdot t_{XOR} \end{aligned}. \tag{3}$$

Для коррекции выявленных ошибок в коде Рида-Соломона решается 2 системы их двух уравнений и вычисляется n 3-х компонентных полиномов.

Таким образом, общее время T_s коррекции искаженных символов определяется следующим образом:

$$T_s \approx 4 \cdot n \cdot m \cdot t_{XOR} \quad . \qquad (4)$$

Если полагать, что "пачка" возникающих ошибок локализирована в рамках одного символа, то можно говорить, что код Рида-Соломона использует для коррекции "пачки" теоретически минимальное количество контрольных разрядов.

Возникновение "пачки" ошибок в каналах с кодово-импульсной модуляцией имеет специфические особенности, учет которых позволяет существенно повысить эффективность исправления "пачки" ошибок по сравнению с кодами Рида-Соломона.

Для коррекции такой ошибки с использованием кодов Рида-Соломона число контрольных символов при этом равно 4-м ($4 \cdot m$ битов), при том, что теоретический минимум - $2 \cdot m$. То есть число контрольных разрядов в коде Рида-Соломона более чем вдвое превышает теоретический минимум.

Следовательно, для рассматриваемого канала коррекцию пачки ошибок можно реализовать с использованием меньшего, чем в коде Рида-Соломона числа контрольных разрядов.

Очевидным является тот факт, что в рамках частного метода для отдельного, часто встречающегося на практике типа ошибок можно обеспечить эффективность больше, чем достигается применением универсального метода, в данном случае кода Рида-Соломона.

Фактически при коррекции "пачки" ошибок надо решить две задачи - локализовать начало "пачки" в блоке и локализовать позиции искаженных битов в самой "пачке".

В коде Рида-Соломона решение этих двух задач сильно связано - это и обуславливает высокую вычислительную сложность исправления "пачки", а также зависимость числа символов от его размера.

Это позволяет добиться кроме уменьшения вычислительной сложности уменьшения требуемого числа контрольных разрядов.

Способ коррекции "пачки" ошибок на основе двухкомпонентной логической взвешенной контрольной суммы

Для того, чтобы скорректировать "пачку" надо знать решить две задачи:

- локализовать положение искаженных битов в "пачке";

- локализовать начало "пачки".

В кодах Рида-Соломона эти две задачи сильно связаны: отсюда получается большая вычислительная сложность и жесткая связь числа символов с их разрядностью [2].

Для того, чтобы кардинально уменьшить вычислительную и временную сложность решения этих задач представляется целесообразным разделить решение этих разные задачи, используя при этом наиболее подходящие методы.

Так получить информацию о положении искаженных битов в "пачке" проще всего используя обычную контрольную сумму, но так, чтобы исключить наложение битовых искажений - для этого разрядность контрольной суммы должна быть примерно вдвое больше максимальной длины "пачки".

Для нахождения начала "пачки" проще всего использовать взвешенную контрольную сумму, которая позволяет "привязать" к искаженным битам их позицию в виде весового коэффициента.

Идея, положенная в основу предлагаемого метода коррекции "пачек ошибок", состоит в том, чтобы добиться уменьшения вычислительной сложности путем раздельного решения двух задач:

1) получение информации о том, какие именно биты "пачки" искажены, то есть получение m-битового вектора битовых ошибок $E=\{e_1,e_2,...,e_m\}$, $\forall l=1,...,m$: $e_l \in \{0,1\}$;

2) локализация первого бита "пачки" битовых ошибок.

Для обнаружения и исправления пачки ошибок передачи цифровых данных в низкочастотном канале с использованием взвешенной контрольной суммы на основе логических операций контролируемый N-битовый блок $B=\{b_1,b_2,...,b_N\}$, $\forall l\in\{1,...,N\}$: $b_l\in\{0,1\}$ данных предлагается разделить на n $2\cdot(m-1)$-битовых символов, где m- максимальная длина “пачки ошибок”. Каждому из N битов контролируемого блока ставятся в соответствие весовые коэффициенты W_1, W_2, ..., W_N, разрядностью $q=\lfloor\log_2 N\rfloor$. Набор весовых коэффициентов W_1, W_2, ..., W_N представляет собой последовательность элементов α, α^2,...α^N поля Галуа, образованного неразложимым полиномом $P(x)$ степени q или, другими словами, последовательность кодов, образуемых при сдвиге q-разрядного сдвигового регистра с линейной функцией обратной связи, соответствующей неразложимому полиному $P(x)$ на поле Галуа.

Контрольный код взвешенной контрольной суммы предлагается вычислять в виде 2-х компонент C_1 и C_2, первая из которых, разрядностью $2\cdot(m-1)$ представляет собой логическую сумму символов $X_1, X_2, \ldots, X_n$ блока:

$$C_1 = X_1 \oplus X_2 \oplus \ldots \oplus X_n. \tag{5}$$

Вторая компонента C_2 формируется как логическая сумма произведений битов блока на их весовой коэффициент:

$$C_2 = b_1\cdot W_1 \oplus b_2\cdot W_2 \oplus \ldots \oplus b_N\cdot W_N. \tag{6}$$

Например, если контролируется 15-разрядный блок B = 1001 1010 1100 101 и m=4, то блок разделяется на три 6-битовых символа (недостающие биты принимаются равными нулю) : X_1 = 100110, X_2 = 101100 и X_3=101000. Соответственно, первая компонента взвешенной контрольной суммы вычисляется как: $C_1 = X_1 \oplus X_2 \oplus X_3$= 100110 $\oplus$ 101100 $\oplus$ 101000 = 100010. Вторая компонента C_2 взвешенной контрольной суммы вычисляется в виде: $S_2 = b_1\cdot W_1 \oplus b_2\cdot W_2 \oplus \ldots \oplus b_{15}\cdot W_{15} = 1 \oplus 15 \oplus 14 \oplus 10 \oplus 11 \oplus 6 \oplus 2 \oplus 8 = 13_{10} = 1101_2$. Таким образом, двухкомпонентный контрольный код на передатчике равен: <100010,1101>.

На приемнике контрольный код взвешенной контрольной суммы вычисляется аналогичным образом. Например, пусть в процессе передачи вследствие воздействия внешней помехи возникает “пачка ошибок”, которая начинается с 5-го бита состоит в искажении 5-го, 6-го и 8-го битов. Соответственно, принятый на принимающей стороне блок B' имеет следующий вид: B' = 1001 0111 1100 101. На приемнике: X_1' = 100101, X_2' = 111100 и X_3'=101000. Соответственно $C_1' = X_1' \oplus X_2' \oplus X_3'$= 100101 $\oplus$ 111100 $\oplus$ 101000 = 110001. $C_2' = b_1'\cdot W_1 \oplus b_2'\cdot W_2 \oplus \ldots \oplus b_{15}'\cdot W_{15} = 1 \oplus 15 \oplus 13 \oplus 10 \oplus 5 \oplus 11 \oplus 6 \oplus 2 \oplus 8 = 11_{10} = 1011_2$.

На приемнике вычисляется состоящая из 2-х компонент разность Δ полученных и вычисленных контрольных кодов $\Delta = \langle\Delta_1,\Delta_2\rangle$, причем : $\Delta_1=C_1\oplus C_1'$ $\Delta_2=C_2\oplus C_2'$. Если обе компоненты разности контрольных кодов приемника и передатчика равны нулю: Δ_1=0 и Δ_2=0, то считается, что блок передан без ошибок. В противном случае, анализируется код Δ_1: если в нем не содержится (m-2)-х следующих подряд нулей (с учетом того, что последний бит Δ_1 считается предшествующим первому), то считается, что при передаче возникло более одной “пачки” ошибок, то есть имеет место ошибка, которая не может быть исправлена. В этом случае приемная сторона посылает приемнику специальный сигнал запроса на повторную передачу блока. Если $\Delta_1\neq0$ и содержит m-2 следующих подряд нулей, то такая ситуация классифицируется как одиночная “пачка” ошибок, которая может быть исправлена путем выполнения изложенных ниже операций над кодами Δ_1 и Δ_2.

В рамках приведенного выше примера Δ_1=100010$\oplus$110001 = 010011 и Δ_2= 1101 $\oplus$ 1011 = 0110 то есть имеет место ситуация возникновения одиночной “пачки” ошибок.

Процедура коррекция включает два этапа: получение вектора E битовых искажений в “пачке” и локализации битовой позиции начала “пачки”.

Для получения кода вектора E битовых искажений по коду Δ_1 определяется положение первого в "пачке" искаженного бита, который соответствует единице, перед которой находится в точности m-2 следующих подряд нулей (с учетом того, что последний бит Δ_1 считается предшествующим первому). Технологически, при программной реализации, указанная операция выполняется с использованием простых операций сдвига и маскированного сравнения. При аппаратной реализации операция просто реализуется на сдвиговом регистре и логической схеме ИЛИ на m-2 входа, которая детектирует наличие последних m-2 нулей. Так для приведенного выше примера Δ_1= 010011. После 4-х операций циклического сдвига влево Δ_1 = 110100 последние m-2=2 бита равны нулю. Это свидетельствует о том, что компонентами вектора E являются первые m=4 бит кода Δ_1, то есть $E=\{1,1,0,1\}$.

Определение весового коэффициента $W_j=\{w_{j,1},w_{j,2},...,w_{j,q}\}$, $j\in\{1,...,N\}$ первого искаженного бита "пачки" на основе известных вектора $E=\{e_1,e_2,...,e_m\}$ битовых искажений и кода $\Delta_2=\{\delta_1, \delta_2,...,\delta_q\}$ разностей вторых компонент контрольных кодов может быть сведено к решению системы линейных уравнений. Заданный неразложимый полином $P(x)$ поля Галуа определяет линейную функцию $\lambda_1(w_{j,1},w_{j,2},...,w_{j,q})$ получения q-той битовой компоненты $w_{j+1,q}$ весового коэффициента W_{j+1} следующего бита блока : $w_{j+1,q}=\lambda_1(w_{j1},w_{j2},...,w_{jq})$, в то время. как первые его q-1 компоненты представляют собой сдвинутые компоненты коэффициента W_j : $w_{j+1,1}=w_{j,2}$, $w_{j+1,2}=w_{j,3},...$, $w_{j+1,q-1}=w_{j,q}$ так, что $W_{j+1}=\{w_{j2},w_{j3},...,w_{jq}, \lambda_1(w_{j1},w_{j2},...,w_{jq})\}$. Например, при $P(x)=x^4+x+1$ линейная функция $\lambda_1=w_{j,1}\oplus w_{j,4}$. Если $W_{11}=\{1,1,0,0\}$ то $W_{12}=\{1,0,0, 1\oplus 0\}$. Аналогично, $W_{j+2}=\{w_{j,3},w_{j,4},...,w_{j,q}, \lambda_1(w_{j,1},w_{j,2},...,w_{j,q}),\lambda_1(w_{j,2},...,w_{j,q},\lambda_1(w_{j,1},w_{j,2},...,w_{j,q}))\}$. Если обозначить линейную функцию $\lambda_1(w_{j,2},...,w_{j,q},\lambda_1(w_{j,1},w_{j,2},...,w_{j,q}))$ как функцию $\lambda_2(w_{j,1},w_{j,2},...,w_{j,q})$, то $W_{j+2}=\{w_{j,3},w_{j,4},...,w_{j,q}, \lambda_1(w_{j,1},w_{j,2},...,w_{j,q}),\lambda_2(w_{j,1},w_{j,2},...,w_{j,q})\}$. Так в рамках рассматриваемого примера $\lambda_2=w_{j,2}\oplus\lambda_1(w_{j,1},...,w_{j,4}))=w_{j,2}\oplus w_{j,1}\oplus w_{j,4}$, соответственно весовой коэффициент W_{13} можно представить в виде: $W_{13}=\{w_{11,3},w_{11,4},w_{11,1}\oplus w_{11,4},w_{11,2}\oplus w_{11,1}\oplus w_{11,4}\}=\{0,0,1\oplus 0,1\oplus 1\oplus 0\}=\{0,0,1,0\}$. Выполнив аналогичные рассуждения и введя соответствующие обозначения, можно прийти к выводу о том, что битовые компоненты весового коэффициента W_{j+m} последнего бита "пачки" могут быть представлены в виде линейных функций от компонент весового коэффициента W_j первого бита "пачки":

$$\begin{aligned}
&W_{j+m}=\{w_{j,m},\ldots,w_{j,q},\lambda_1(w_{j,1},\ldots,w_{j,q}), \ldots, \lambda_{m-1}(w_{j,1},\ldots,w_{j,q})\} \text{ при } m<q;\\
&W_{j+m}=\{w_{j,q},\lambda_1(w_{j,1},\ldots,w_{j,q}), \ldots,\lambda_{q-1}(w_{j,1},\ldots,w_{j,q})\} \text{ при } m=q;\\
&W_{j+m}=\{\lambda_{m-q+1}(w_{j,1},\ldots,w_{j,q}), \ldots, \lambda_m(w_{j,1},\ldots,w_{j,q})\} \text{ при } m>q.
\end{aligned} \quad (7)$$

Каждая l-тая, где l=2,...,m-1, линейная функция λ_l рекурсивно выражается через ранее определенные функции: λ_1 $\lambda_2,...,\lambda_{l-1}$ следующим образом:

$$\lambda_l(w_{j,1},...,w_{j,q})=\lambda_1(w_{j,i},...,w_{j,q..}, \lambda_1(w_{j,1},...,w_{j,q}),..., \lambda_{l-1}(w_{j,1},...,w_{j,q})\ ). \quad (8)$$

Если предположить, что "пачка" ошибок начинается с искаженного бита с порядковым номером j, то первая компонента δ_1 разности Δ_2 представляет собой логическую сумму первых компонент весовых коэффициентов тех бит, которые подверглись искажению, то есть тем битам "пачки" которым соответствуют единичные компоненты вектора E битовых искажений:

$$e_1\cdot w_{j,1}\oplus e_2\cdot w_{j+1,1}\oplus e_3\cdot w_{j+2,1}\oplus ... \oplus e_m\cdot w_{j+m,1}=\delta_1 \quad (9)$$

Выражение (9) может быть обобщено и для всех остальных компонент $\delta_2,...,\delta_q$ разности Δ_2, i-тая, i=1,...,q компонента δ_i разности Δ_2 представляет собой логическую сумму i-тых компонент весовых коэффициентов тех бит, которые подверглись искажению. Следовательно, можно составить q уравнений вида:

$$e_1 \cdot w_{j,1} \oplus e_2 \cdot w_{j+1,1} \oplus e_3 \cdot w_{j+2,1} \oplus \cdots$$
$$\oplus e_m \cdot w_{j+m,1} = \delta_1$$
$$e_1 \cdot w_{j,2} \oplus e_2 \cdot w_{j+1,2} \oplus e_3 \cdot w_{j+2,2} \oplus \cdots$$
$$\oplus e_m \cdot w_{j+m,2} = \delta_2 \qquad (10)$$
$$\cdots$$
$$e_1 \cdot w_{j,q} \oplus e_2 \cdot w_{j+1,q} \oplus e_3 \cdot w_{j+2,q} \oplus \cdots$$
$$\oplus e_m \cdot w_{j+m,q} = \delta_q$$

С учетом того, что каждая из компонент весовых коэффициентов W_{j+1}, W_{j+2}, ..., W_{j+m} может быть выражена как линейная функция от компонент весового коэффициента W_j, все уравнения системы (10) могут быть линейно выражены через компоненты $w_{j,1}$, $w_{j,2}$,..., $w_{j,q}$ коэффициента W_j. Следовательно, система (6) может быть представлена в виде системы q линейных уравнений от q бинарных неизвестных $w_{j,1}$, $w_{j,2}$,..., $w_{j,q}$ для ситуации когда $m \geq q$ следующим образом:

$$e_1 \cdot w_{j,1} \oplus e_2 \cdot w_{j,2} \oplus \cdots$$
$$\oplus e_m \cdot \lambda_{m-q+1}(w_{j,1},..,w_{j,q}) = \delta_1$$
$$e_1 \cdot w_{j,2} \oplus e_2 \cdot w_{j,3} \oplus \cdots$$
$$\oplus e_m \cdot \lambda_{m-q+2}(w_{j,1},..,w_{j,q}) = \delta_2 \qquad (11)$$
$$\cdots$$
$$e_1 \cdot w_{j,q} \oplus e_2 \cdot \lambda_1(w_{j,1},..,w_{j,q}) \oplus \cdots$$
$$\oplus e_m \cdot \lambda_m(w_{j,1},..,w_{j,q}) = \delta_q$$

Если $m<q$, то система (11) может быть представлена как система линейных уравнений в следующем виде:

$$e_1 \cdot w_{j,1} \oplus e_2 \cdot w_{j,2} \oplus \cdots \oplus e_m \cdot w_{j,m} = \delta_1$$
$$e_1 \cdot w_{j,2} \oplus e_2 \cdot w_{j,3} \oplus \cdots \oplus e_m \cdot w_{j,m+1} = \delta_2$$
$$\cdots$$
$$e_1 \cdot w_{j,q-m+1} \oplus e_2 \cdot w_{j,q-m+2} \oplus \cdots$$
$$\oplus e_m \cdot w_{j,q} = \delta_{q-m+1}$$
$$e_1 \cdot w_{j,q-m+2} \oplus e_2 \cdot w_{j,q-m+3} \oplus \cdots \qquad (12)$$
$$\oplus e_m \cdot \lambda_1(w_{j,1},..,w_{j,q}) = \delta_{q-m+2}$$
$$\cdots$$
$$e_1 \cdot w_{j,q} \oplus e_2 \cdot \lambda_1(w_{j,1},..,w_{j,q}) \oplus \cdots$$
$$\oplus e_m \cdot \lambda_{m-1}(w_{j,1},..,w_{j,q}) = \delta_q$$

Системы (11) и (12) могут быть достаточно просто решены известными способами и, как результат, получены значения двоичных компонент $w_{j,1}$, $w_{j,2}$,..., $w_{j,q}$ весового коэффициента W_j с которого начинается "пачка" ошибок, то есть таким образом локализовано ее начало в блоке.

Приведенные теоретические выкладки могут быть иллюстрированы в рамках рассмотренного выше примера следующим образом. Ранее определены компоненты вектора E битовых искажений внутри "пачки" следующим образом: $E=\{1,1,0,1\}$, то есть $e_1=1$, $e_2=1$, $e_3=0$ и $e_4=1$. Разность $\Delta_2=\{0,1,1,0\}$, то есть $\delta_1=0$, $\delta_2=1$, $\delta_3=1$ и $\delta_4=0$. Значение $q=m=4$. δ_1- логическая сумма первых разрядов взвешенных контрольных сумм для искаженных разрядов: $e_1 \cdot w_{j,1} \oplus e_2 \cdot w_{j+1,1} \oplus e_3 \cdot w_{j+1,1} \oplus e_4 \cdot w_{j+1,1} = \delta_1$ Учитывая, что $e_1=1$, $e_2=1$, $e_3=0$ и $e_4=1$, а $\delta_1=0$ первое уравнение системы получается в виде: $w_{j,1} \oplus w_{j+1,1} \oplus w_{j+3,1} = 0$. Аналогичным образом составляются остальные 3 уравнения:

$$w_{j,1} \oplus w_{j+1,1} \oplus w_{j+3,1} = 0$$
$$w_{j,2} \oplus w_{j+1,2} \oplus w_{j+3,2} = 1$$
$$w_{j,3} \oplus w_{j+1,3} \oplus w_{j+3,3} = 1$$
$$w_{j,4} \oplus w_{j+1,4} \oplus w_{j+3,4} = 0$$

Код W_{j+1} - это сдвинутый влево код W_j, а младший разряд - логическая сумма 1-го и 4-го разрядов. Это означает, что разряды W_{j+1} можно представить как линейную комбинацию разрядов W_j : $w_{j+1,1}= w_{j+1,2}=w_{j,3}$, $w_{j+1,3}=w_{j,4}$ $w_{j+1,4} = w_{j,1} \oplus w_{j,4}$. Аналогично для W_{j+2}: $w_{j+2,1}=w_{j+1,2}=w_{j,3}$; $w_{j+2,2}=w_{j,4}$; $w_{j+2,3}= w_{j,1} \oplus w_{j,4}$; $w_{j+2,4} = w_{j+1,1} \oplus w_{j+1,4}= w_{j,2} \oplus w_{j,1} \oplus w_{j,4}$. Для W_{j+3} линейное представление его компонент через битовые компоненты W_j имеет вид: $w_{j+3,1}=w_{j,4}$; $w_{j+3,2}= w_{j,1} \oplus w_{j,4}$; $w_{j+3,3}= w_{j,2} \oplus w_{j,1} \oplus w_{j,4}$; $w_{j+3,4} = w_{j+2,1} \oplus w_{j+2,4}= w_{j,3} \oplus w_{j,2} \oplus w_{j,1} \oplus w_{j,4}$. После подстановки и элементарных преобразований получается система (7) линейных уравнений, неизвестными которой являются двоичные разряды W_j - весового коэффициента бита, с которого начинается "пачка":

$$w_{j,1} \oplus w_{j,2} \oplus w_{j,4} = 0$$
$$w_{j,2} \oplus w_{j,3} \oplus w_{j,1} \oplus w_{j,4} = 1$$
$$w_{j,3} \oplus w_{j,1} \oplus w_{j,2} = 1$$
$$w_{j,4} \oplus w_{j,2} \oplus w_{j,3} = 0$$

Решениями системы являются значения: $w_{j,1}=1$, $w_{j,2}=1$, $w_{j,3}=1$ и $w_{j,4}=0$. Таким образом $W_j = 1110_2=14$ и это означает, что "пачка" ошибок начинается с 5-го бита (j=5). Используя значения компонент вектора $E=\{1,1,0,1\}$, исправление "пачки" осуществляется инвертированием 5-го, 6-го и 8-го битов блока.

Технологически, наиболее быстрым вариантом решения системы линейных уравнений с двоичными коэффициентами (11) или (12) является использование табличной памяти. Для фиксированного образующего полинома $P(x)$ несложно в кодированном виде представить вычисление каждой из компонент W_j как линейных функций от компонент Δ_2 для каждого из 2^m-1 возможных значений m-разрядного вектора E. Так, например, для $E=\{1,1,0,1\}$ значения компонент W_j в рамках рассматриваемого примера могут быть представлены в виде:

$$w_{j,1} = \delta_2 \oplus \delta_4$$
$$w_{j,2} = \delta_1 \oplus \delta_3$$
$$w_{j,3} = \delta_1 \oplus \delta_2$$
$$w_{j,4} = \delta_2 \oplus \delta_3$$

Общий объем требуемой табличной памяти не превышает $2^m \cdot q$ бит. Учитывая, что на практике значение m не превышает 8-ми [2], а q редко превышает 16, объем табличной памяти вполне укладывается в рамки десятков килобайт, что далеко не выходит за рамки технических возможностей как при программной, так и при аппаратной реализации.

Количество k контрольных разрядов определяется суммой числа разрядов первой компоненты C_1 контрольного кода - $2\cdot(m$-1) и второй компоненты C_2 контрольного кода - $\log_2 N$:

$$k = 2\cdot(m-1)+\log_2 N. \qquad (13)$$

При использовании кодов Рида-Соломона для решения аналогичной задачи - исправления в низкочастотном канале одной "пачки" ошибок длиной не более m бит необходимо обеспечивать возможность коррекции 2-х m-битовых символов, что требует $4\cdot m$ контрольных разрядов. Поскольку в кодах Рида-Соломона количество контрольных разрядов рассчитывается исходя из предельного значения числа n символов равного 2^m, то для более наглядного сравнения с предлагаемым методом можно считать, что $N=n\cdot m=m\cdot 2^m$, соответственно, $\log_2 N=m + \log_2 m$, так, что формула (13) преобразуется к виду:

$$k' = 3\cdot m + \log_2 m - 2\,. \qquad (14)$$

Сравнение оценки (10) с количеством контрольных разрядов ($4\cdot m$), используемым для решения аналогичной задачи кодом Рида-Соломона, свидетельствует о том, что в предлагаемом методе, за счет его специализации, то есть в ущерб универсальности, достигается уменьшение количества контрольных разрядов по сравнению с кодом Рида-Соломона.

На практике, если $N < m\cdot 2^m$ уменьшение числа контрольных разрядов, достигаемое при использовании разработанного метода в сравнении с кодами Рида-Соломона еще значительнее. Например, если решается задача коррекции одиночной "пачки" ошибок, длина которой не превышает 8 (m=8) в блоке длиной в 64 байта то при использовании кодов Рида-Соломона требуется $4\cdot m$=32 контрольных бита, в то время, как предложенный способ решает ту же задачу с использованием всего $2\cdot(m$-1)+$\log_2 64\cdot m$ = 23 контрольных разрядов.

В кодах Рида-Соломона процедуры получения информации о позиции искаженного символа и о искаженных в нем битах интегрированы в единый механизм. Это обстоятельство связывает длину блока с разрядностью символа и, тем самым, накладывает ограничения на максимальную длину блока - она не должна превышать 2^m символов или $2^m\cdot m$ битов. Так, при m=4 длина контролируемого блока ограничивается всего 8-мю байтами. Предложенный метод благодаря разнесению механизмов получения информации о позиции "пачки" и искаженных е ней битах не накладывает никаких ограничений на длину контролируемого блока.

В кодах Рида-Соломона отсутствует механизм контроля появления ошибок, кратность которых превышает возможность их исправления. Поэтому, при возникновении двух "пачек" ошибочно "корректируются" неискаженные символы. В предлагаемом методе есть механизм контроля неисправимых ошибок.

Основным преимуществом предложенного метода коррекции одиночной "пачки" ошибок в сравнении с кодами Рида-Соломона является существенно меньшая вычислительная сложность операций, выполняемых в ходе контроля и исправления ошибок.

Время вычисления T'_c контрольного кода в разработанной модификации контрольного кода определяется временем вычисления ее 2-й компоненты:

$$T'_c = n \cdot m \cdot t_{XOR} \, . \qquad (15)$$

Сравнение полученного выражения (15) с (2) и (3) позволяет сделать вывод о том, что время вычисления контрольного кода и декодирования уменьшается по сравнению с кодами Рида-Соломона в $T_c/T'_c = 8$ раз.

Общее время коррекции T'_s "пачки" ошибок в разработанном способе определяется временем решения системы из $q=\log_2(n \cdot m)$ линейных булевых уравнений оценивается как:

$$T'_s \approx 8 \cdot q^2 \cdot t_{XOR} \, . \qquad (16)$$

Сравнение (16) с выражением (4) также позволяет сделать вывод о том, что операция коррекции "пачки" ошибок также выполняется в разработанном методе значительно быстрее по сравнению с кодами Рида-Соломона. Например, при n=256, m=8, q=11 выигрыш во времени коррекции составляет T_s/T'_s= 8192/968=8.5.

Таким образом, предложенный способ коррекции одиночной пачки ошибок в низкочастотном канале имеет существенно меньшую вычислительную сложность по сравнению с кодом Рида-Соломона, а также использует меньшее число контрольных разрядов.

Выводы

Предложен подход к повышению эффективности коррекции "пачки" ошибок за счет раздельного выполнения двух задач - локализации положение искаженных битов в "пачке" и локализации начала "пачки" в контролируемом блоке. Для уменьшения вычислительной и временной сложности в рамках этого подхода для решения первой задачи предлагается использовать обычную контрольную сумму, а для решения второй - взвешенную контрольную сумму. Для реализации такого подхода разработан способ коррекции одиночной "пачки ошибок", длина которой не превышает m битов с использованием двухкомпонентной взвешенной логической контрольной суммы, отличающийся использованием $2 \cdot (m-1)$-разрядной контрольной суммы для локализации позиций искаженных битов в "пачке" и логической взвешенной контрольной суммы с отдельным весовым коэффициентом для каждого бита - для локализации первого бита "пачки". Доказано, что применение метода позволяет на порядок уменьшить вычислительную сложность операций, связанных с коррекцией "пачки" ошибок по сравнению с кодами Рида-Соломона, при том, что в нем используется меньшее число контрольных разрядов. Кроме того, предложенный метод, позволяет распознавать одиночные "пачки", которые могут быть исправлены от двойных, которые не могут быть скорректированы.

Список литературы

1. Марковский А.П., Саидреза Махмали, Турченко Ю.А., Сакун В.Н. Использование взвешенных контрольных сумм для исправления "пачек" ошибок передачи данных // Вісник національного технічного університету України "КПІ". Інформатика, управління та обчислювальна техніка. – 2009. - № 51. – С. 90 – 96.

2. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. – М.: Издательский дом "Вильямс". – 2004. – 1104 с.