

Восстановление искажённых пакетов данных без их повторной передачи по линии связи

При передаче пакетов данных по линии связи возможны их искажения в результате действия помех. Обычно для обнаружения ошибок в каждый передаваемый пакет вводится контрольная сумма. На приёмной стороне также вычисляется контрольная сумма пакета и сверяется с принятой. При несовпадении контрольных сумм пакет считается ошибочным. Ошибки невысокой кратности могут быть устранены, если контрольное суммирование выполнялось с использованием корректирующих кодов, таких как циклический избыточный код CRC. Если число ошибок в пакете велико (в пределе, каждый бит пакета ошибочен), то, на первый взгляд, для их исправления необходима весьма значительная избыточность. Например, если бы каждый пакет передавался в трёх или более экземплярах, то при благоприятном распределении ошибок между ними можно было бы по “большинству голосов” реконструировать правильный пакет. В данной статье рассмотрены более экономичные методы [1, 2] восстановления искажённых пакетов.

1. Введение

В традиционной системе передачи, показанной на рис. 1, а, данные, упакованные в информационные пакеты одинаковой длины, передаются по линии связи из устройства А в устройство В. В результате действия помех на линию связи некоторые пакеты искажаются. В примере, приведенном на рисунке, при передаче пакета с номером i в нём возникли искажения четырёх битов (искажения помечены “кляксами”).

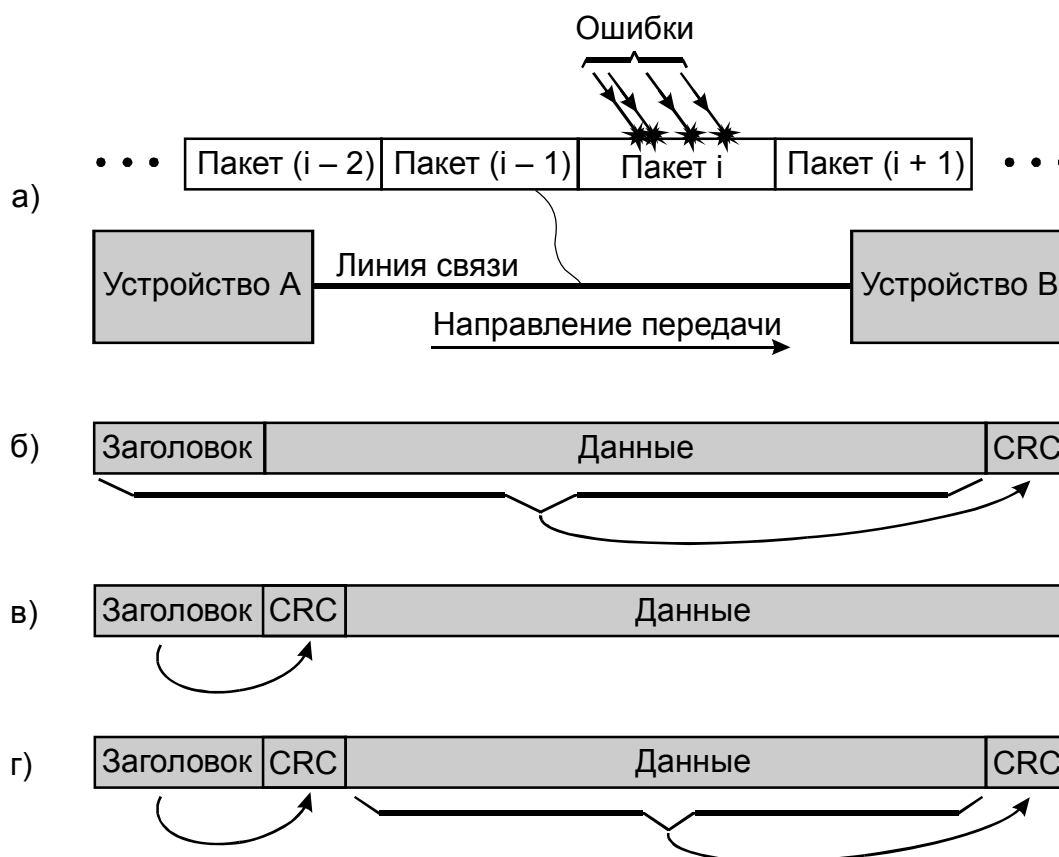


Рис. 1. Передача пакетов данных: а — схема; б, в, г — структуры пакетов

Для обнаружения ошибок и, если это возможно, их исправления без повторной передачи, в пакет обычно вводятся один или несколько циклических избыточных кодов CRC (рис. 1, б — г). Пакет в общем случае содержит заголовок и поле данных.

В структуре пакета, показанного на рис. 1, б, для вычисления кода CRC и его последующей проверки используются биты заголовка и биты поля данных как единый мас-

сив. Чем больше длина контролируемого массива битов, тем меньше корректирующие возможности кода CRC фиксированной длины.

В структуре, показанной на рис. 1, *в*, код CRC относится только к заголовку, так как в данном случае от ошибок защищается лишь наиболее значимая часть пакета. Поле данных в данном случае не контролируется. Такая структура пакета предусмотрена в технологии ATM [3].

В структуре пакета, представленной на рис. 1, *г*, использованы два кода CRC, которые отдельно контролируют поле заголовка и поле данных. Корректирующие способности первого и второго кодов могут быть разными. Например, с помощью первого кода CRC можно исправлять двойные ошибки в заголовке, в то время как с помощью второго кода исправляются только одиночные ошибки в поле данных.

Если устройство-приёмник В не способно исправить имеющиеся в пакете ошибки, когда их число превышает корректирующие возможности кода CRC, то оно в зависимости от принятого протокола обмена данными либо запрашивает от устройства А повторную передачу искажённого пакета, либо просто отбрасывает его.

Помеха, действующая на линию связи, может создавать групповые ошибки, т. е. исказить целую группу битов передаваемого пакета. В этом случае использование кодов CRC в традиционных структурах рассмотренных пакетов (рис. 1, *б* — *г*) оказывается неэффективным, так как с увеличением корректирующих возможностей этих кодов возрастает их длина. При этом если, например, искажен весь пакет, включая коды CRC, то восстановить его в рамках традиционных структур невозможно. Единственный способ восстановления — получение безошибочной или содержащей допустимое число ошибок корректируемой копии пакета от устройства А по запросу устройства В. Но этот способ связан с большими потерями времени.

Рассмотренные далее методы восстановления пакетов позволяют противостоять групповым ошибкам без повторной передачи пакетов. Длина каждой отдельной “зоны поражения” ошибками потока данных может достигать длины передаваемого по линии связи пакета. В эту зону может попадать либо один пакет, либо части двух смежных пакетов.

2. Метод “вертикального” контрольного суммирования группы пакетов

Далее рассмотрен способ восстановления искажённого пакета с произвольным числом имеющихся в нём ошибок [1]. В худшем случае пакет может содержать до ста процентов ошибочных битов. Исправляются также смежные пакеты, если общая “зона поражения” не превышает длины одного пакета. Способ основан на двумерном представлении группы передаваемых пакетов (рис. 2).

Прежде чем начать передачу группы из N пакетов данных $D1 — DN$, устройство-передатчик А (рис. 1, *а*) размещает эти пакеты в памяти друг под другом, как показано в верхней половине рис. 2. Пакеты имеют фиксированную длину (L бит) и содержат заголовки длиной $M - 1$ бит и поля данных длиной $L - M + 1$ бит. Совокупность полей данных пакетов $D1 — DN$ рассматривается как матрица битов, содержащая N строк и $(L - M + 1)$ столбцов.

После просмотра и обработки группы пакетов данных устройство А формирует и записывает в память группу пакетов контрольных кодов CRC, как показано в нижней половине рис. 2. В этой группе также присутствует матрица битов, содержащая J строк и $L - M + 1$ столбцов, где J — разрядность применяемых кодов CRC. Коды CRC ($Y1Y2...YJ$), как и порождающие их коды $X1X2...XN$ размещаются в одноимённых столбцах соответствующих матриц.

Предположим сначала, что заголовки пакетов не защищаются от ошибок.

После формирования пакетов данных и пакетов контрольных кодов в соответствии с рис. 2 устройство А выдаёт их в линию связи в следующем порядке: $D1 — DN$, $CRC1 — CRCJ$. Передача каждого пакета начинается с бита 1 и завершается битом L . Передача пакетов данных может совмещаться во времени с формированием пакетов кон-

трольных кодов. Тогда к моменту выдачи в линию связи последнего бита L последнего пакета данных DN пакеты контрольных кодов уже сформированы, так что в последующих тактах в линию передаётся бит 1 заголовка пакета $CRC1$, затем бит 2 этого заголовка и т. д.

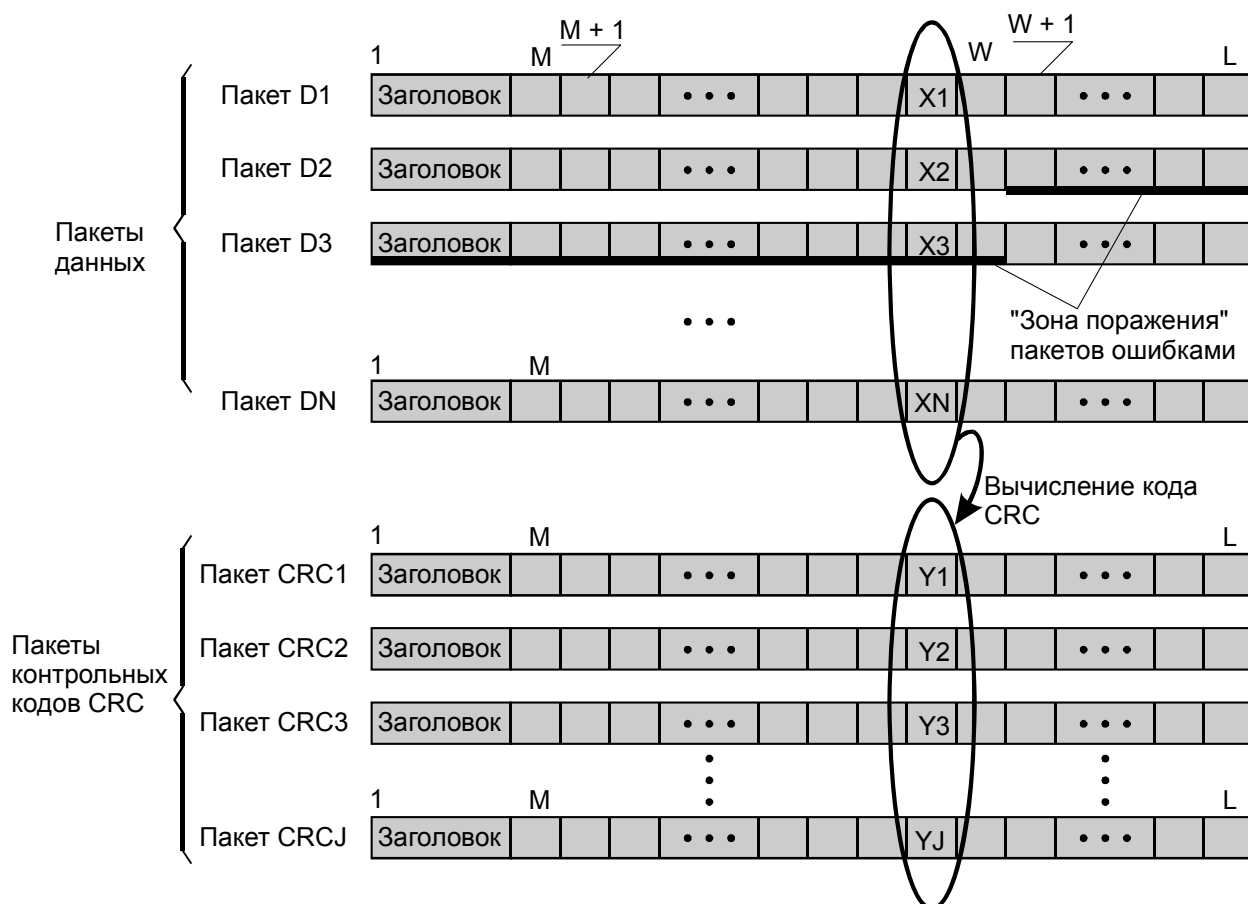


Рис. 2. Пакеты данных и контрольных кодов, подготовленные устройством-передатчиком (А) к выдаче в линию связи

Основная идея предложенного в [1] способа исправления ошибок состоит в том, что корректируемые коды данных и сопровождающие их корректирующие коды CRC располагаются в памяти не по горизонтали (как на рис. 1), а по вертикали. Это имеет принципиальное значение.

Действительно, даже при минимальных корректирующих возможностях кода CRC, когда исправляются только одиночные ошибки в столбцах, поле данных некоторого пакета из группы $D1 — DN$ на приёмной стороне может состоять только из ошибочных битов (остальные пакеты в группе, как предполагаем, в этом случае правильны). Тем не менее, все ошибки искажённого пакета будут скорректированы благодаря тому, что в каждом столбце реконструированной устройством В пары матриц (рис. 3) присутствует не более одной ошибки, а она исправляется кодом CRC.

Это же относится и к ситуациям, при которых групповая ошибка искажает два смежных пакета. На рис. 2 утолщёнными горизонтальными линиями показана одна из возможных допустимых “зон поражения” смежных пакетов $D2$ и $D3$. Части этой зоны, принадлежащие смежным пакетам, не должны перекрываться по вертикали, иначе в столбцах матрицы появятся двойные (некорректируемые) ошибки. Вслед за выдачей битов $W + 1, W + 2, \dots, L - 1, L$ пакета $D2$ без какой-либо паузы в линию связи выдаются биты $1, 2, \dots, W - 1, W$ следующего пакета $D3$. Эти биты принадлежат общей “зоне поражения”, но, благодаря тому, что они размещены в разных столбцах матрицы, все ошибки исправляются. Как будет показано, заголовки пакетов данных можно защищать так же, как и сами данные.

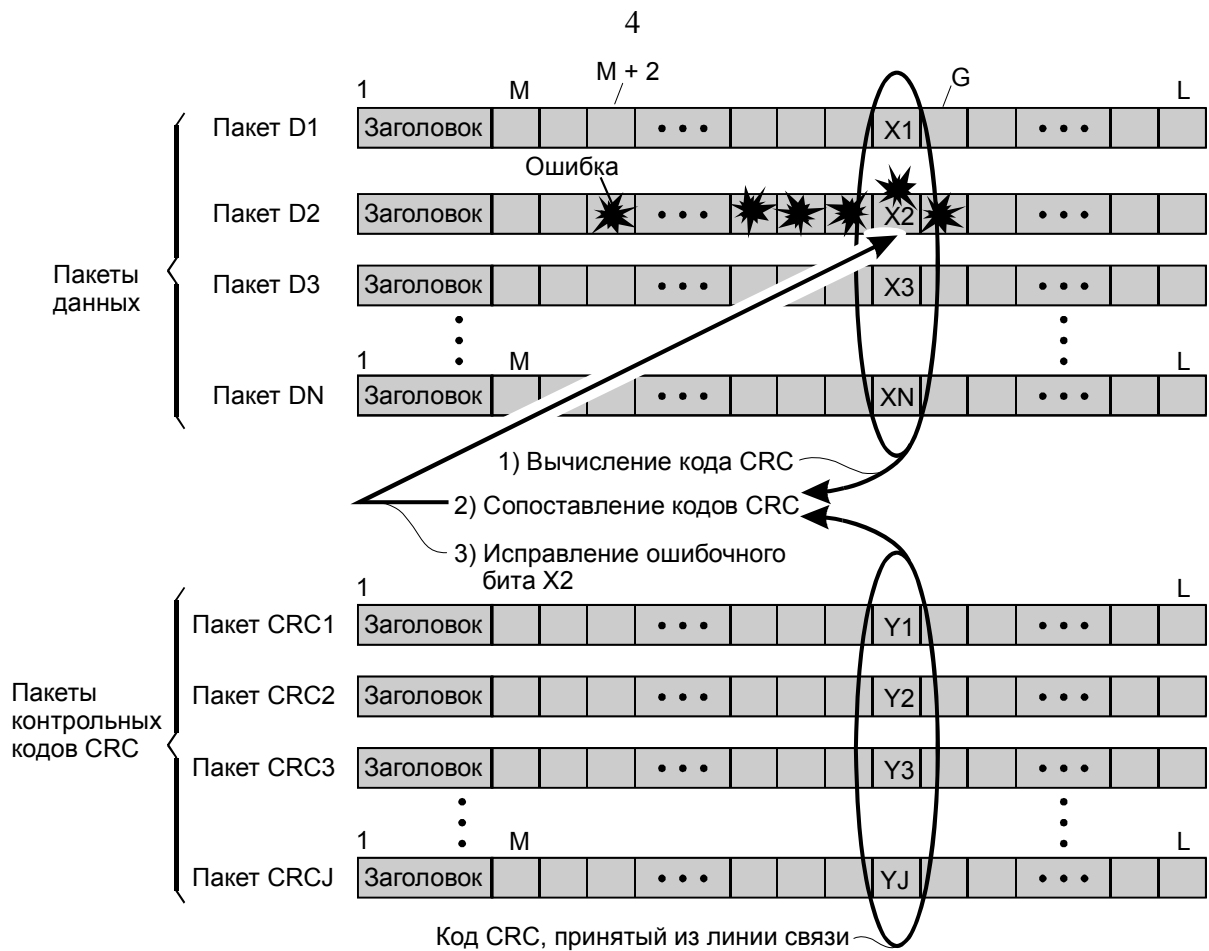


Рис. 3. Пакеты данных и контрольных кодов, принятые в память устройства-приёмника (В) из линии связи

Устройство-приёмник размещает полученные из линии связи пакеты в памяти так же, как они первоначально размещались в памяти устройства-передатчика. Как показано на рис. 3, пакет данных D2 содержит ряд ошибочных битов, помеченных “кляксами”, но устройство-приёмник пока “не знает” об этом.

Далее выполняются следующие действия.

1. Для каждого столбца битов данных, начиная от столбца M и заканчивая столбцом L, вычисляется контрольная сумма (код CRC). Эти суммы на рисунке не показаны.
2. Вычисленные коды CRC сопоставляются с теми, которые поступили из устройства A и размещены в пакетах CRC1 — CRCJ. В тех столбцах, где ошибок нет, вновь вычисленные и принятые из линии связи коды CRC совпадают.
3. В тех столбцах, где вычисленные и хранимые в памяти коды CRC не совпадают, осуществляется коррекция одиночных ошибок. В данном примере показано исправление очередного ошибочного бита X2 в пакете D2.

Таким образом, групповая ошибка, вызванная помехой в линии связи, распадается на отдельные биты, каждый из которых исправляется независимо от соседних. Иными словами, все биты данных некоторого пакета могут быть ошибочными и тем не менее пакет будет восстановлен!

Ранее предполагалось, что заголовки пакетов данных не защищаются от ошибок. Как осуществить такую защиту? Возможны разные способы, например, такие.

- Каждый заголовок защищается индивидуальным кодом CRC, как было показано на рис. 1, в, причём для повышения степени защиты этот код может включаться в состав поля данных соответствующего пакета. В остальном, описанный способ коррекции остаётся тем же (рис. 2, 3).
- Заголовки пакетов контрольных кодов CRC (рис. 2) исключаются. На их месте располагаются столбцы из кодов Y1Y2...YJ, соответствующих кодам X1X2...XN. По-

следние в данном случае представляют собой “вертикальные срезы” массива заголовков пакетов данных $D1 — DN$. Особенность такого решения состоит в том, что устройство В должно “знать”, что в каждом цикле приёма полной группы пакетов первая её часть (пакеты данных) содержит заголовки, а вторая лишена их.

Данный способ коррекции групповых ошибок, разумеется, имеет ограничения. Он не обеспечивает исправления ошибок, кратность которых превышает корректирующие возможности применяемого кода CRC. Имеются в виду ситуации, при которых ошибки размещены в одном и том же столбце $X1X2...XN — Y1Y2...YJ$. Такое размещение ошибок, однако, не характерно для реальных условий работы телекоммуникационных систем. Обычно помеха искажает часть одного или части двух или нескольких смежных пакетов, последовательно передаваемых по линии связи. Иными словами, групповая ошибка, вероятнее всего, распределена “по горизонтали”, а не “по вертикали” (в принятых здесь “геометрических” терминах).

Вносимая в процесс передачи данных кодовая избыточность относительно невелика и полностью соответствует той, которая имеется при обычном “горизонтальном” расположении кодов CRC (рис. 1). Так, если $N = 32$, то $J = 8$, т. е., по “глобальной” оценке, в дополнение к каждому четырём байтам данных в передаваемый поток вводится один корректирующий байт. При этом исправляется одна ошибка в каждом столбце, так что допустимая кратность групповой ошибки равна длине пакета L .

3. Метод контрольного суммирования пакетов в “скользящем окне”

С помощью метода [2] решается та же задача, что и рассмотренная ранее — устраняются искажения пакетов, вызванные действием помех на линию связи. Как и ранее, предполагаем, что искажаться может любая часть пакета или даже весь пакет. В отличие от предыдущего метода [1], в поток данных вводится более заметная избыточность — длина каждого исходного пакета удваивается. Но при этом для проверки данных и выполнения коррекции требуется меньшая вычислительная мощность и снижается объём буферной памяти. Поэтому уменьшается задержка потока данных, связанная с анализом и исправлением ошибок, а это существенно при передаче информации в реальном времени.

Устройство-передатчик А (рис. 1, *a*) формирует очередь из исходных пакетов P_i (рис. 4, *a*). Каждый такой пакет, как и ранее, содержит заголовок и поле данных. На рисунке эти детали не отражены, как несущественные для описания идеи [2]. Исходные пакеты обрабатываются устройством А в порядке возрастания их номеров.

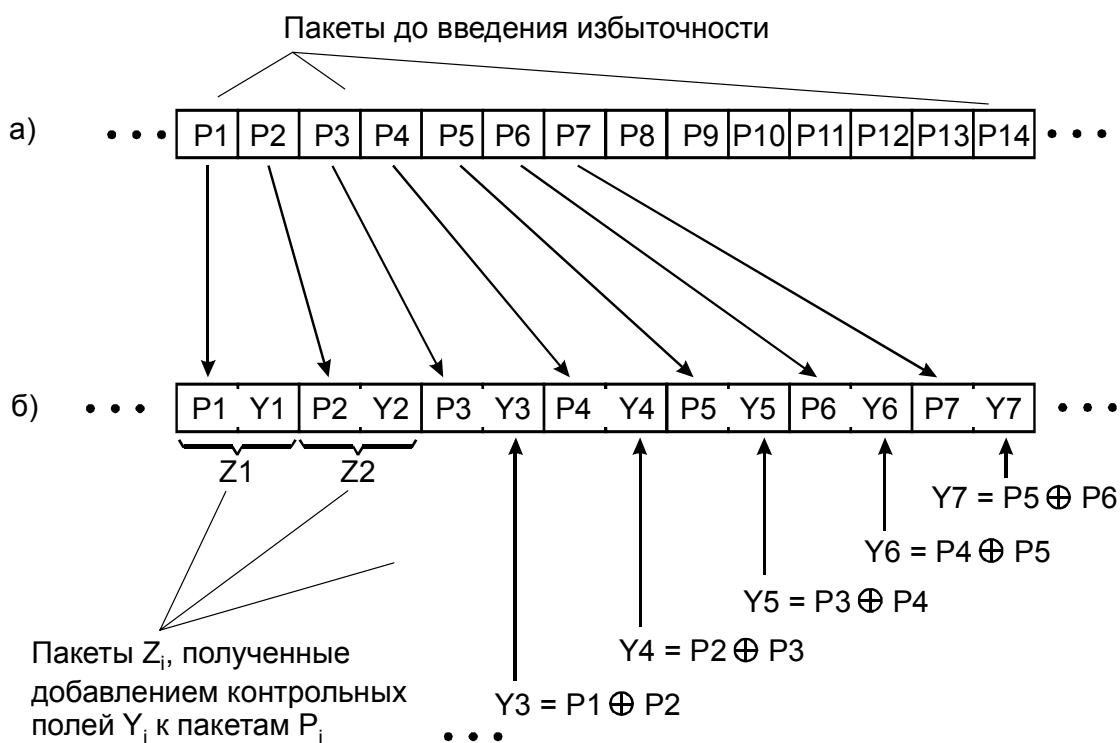


Рис. 4. Подготовка пакетов к передаче по линии связи:
a — исходная последовательность P_i ; *b* — эта же последовательность после введения в неё избыточности в виде контрольных полей Y_i

Прежде чем передать пакет P_i в линию связи, устройство А дополняет его контрольным полем Y_i (рис. 4, б); в результате формируются пакеты Z_i . Длина контрольного поля равна длине исходного пакета. Сформированные пакеты Z_i передаются в линию связи в порядке возрастания их номеров i .

Контрольные поля Y_i формируются суммированием по модулю два двух предыдущих исходных пакетов: P_{i-2} и P_{i-1} . В свою очередь, каждый исходный пакет участвует в формировании двух контрольных полей. Такое “переплетение” данных, как показано далее, создаёт предпосылки для их коррекции в случае возникновения ошибок в связи с действием помех на линию связи.

При использовании данного метода допустимая длина “зоны поражения” потока данных одной группой ошибок не должна превышать длины пакета Z_i . Это означает, что возможны четыре кодовые ситуации, при которых ошибки могут исказить:

- пакет Z_i , в том числе, входящий в него пакет P_i и контрольное поле Y_i (первая кодовая ситуация);
- часть пакета Z_i , а именно, входящий в него пакет P_i ; (вторая кодовая ситуация);
- часть пакета Z_i , а именно, контрольное поле Y_i ; (третья кодовая ситуация);
- два смежных пакета Z_i и Z_{i+1} , а именно, контрольное поле Y_i и пакет P_{i+1} (четвёртая кодовая ситуация).

Эти кодовые ситуации могут в произвольном порядке чередоваться во времени. Для правильной работы системы коррекции ошибок необходимо, чтобы между моментами появления этих кодовых ситуаций существовали интервалы времени, в которых, в зависимости от предыстории, наблюдается поступление одного, двух или трёх неискажённых пакетов Z_i . Иными словами, отдельные и, возможно, весьма протяжённые “зоны поражения” пакетов групповыми ошибками не должны сливаться между собой — между этими зонами должны оставаться области, в которых размещены неискажённые пакеты.

Рассмотрим перечисленные кодовые ситуации подробнее.

Первая кодовая ситуация (рис. 5) — ошибки присутствуют в обеих половинах пакета Z_i

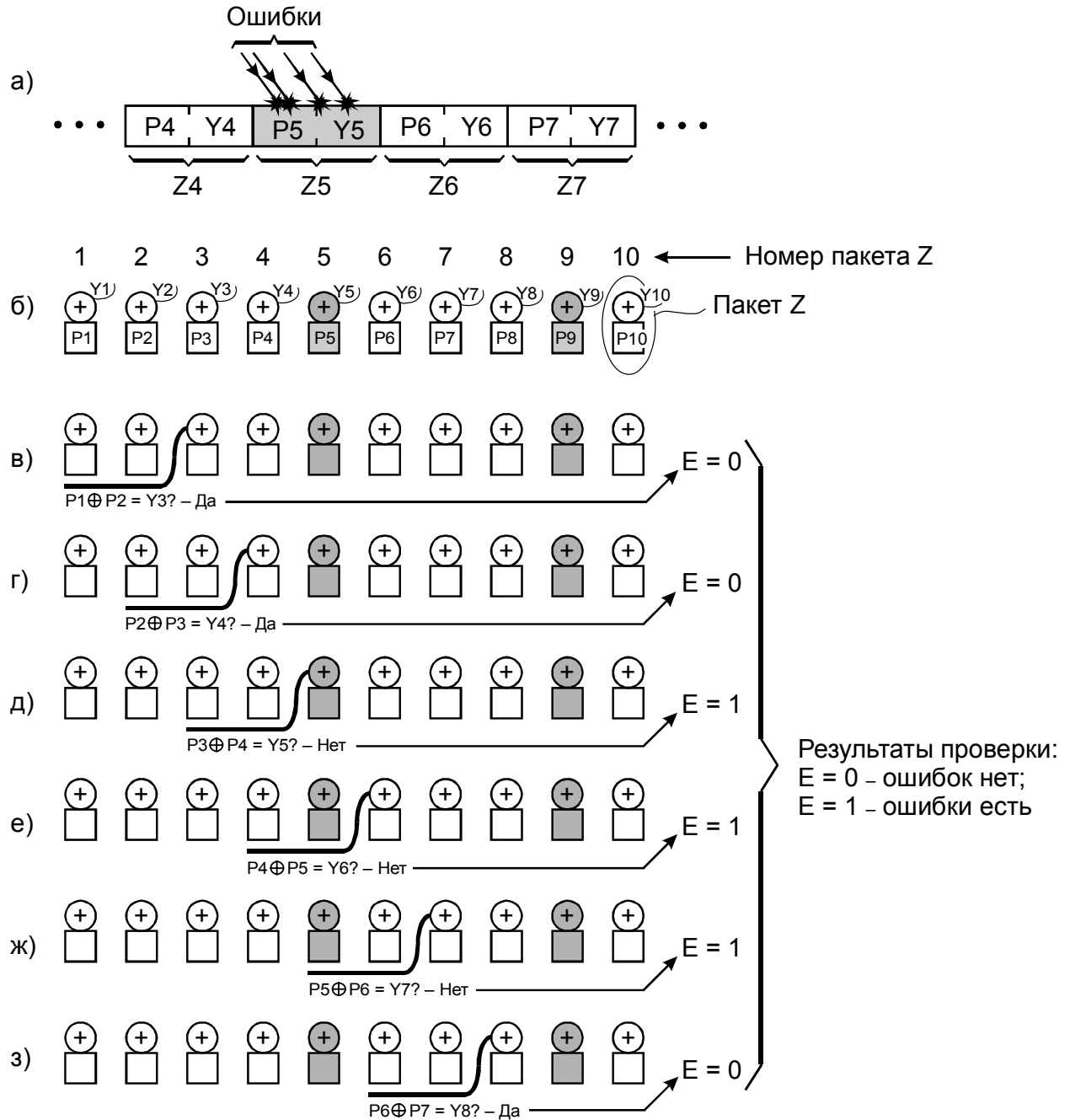


Рис. 5. Приём, размещение (а, б) и анализ (в — з) группы пакетов Z_i устройством-приёмником (первая кодовая ситуация)

События развиваются следующим образом.

Устройство В принимает пакеты Z_i и размещает их в буферной памяти (рис. 5, а). Для более наглядного отображения процесса обработки входного потока пакетов разместим их в двухъярусной строке, состоящей из квадратов и кружков (рис. 5, б). В нижнем ярусе строки показаны пакеты P_i , в верхнем — соответствующие контрольные поля Y_i . Искорёженные пакеты (P_5 , P_9) и искорёженные контрольные поля (Y_5, Y_9) выделены на рисунке серым фоном.

Строка из пакетов Z_i непрерывно пополняется новыми данными и анализируется. Для определённости предположим, что анализ начинается с сопоставления составных частей пакетов Z_1 — Z_3 (рис. 5, в). По определению, $Y_3 = P_1 \oplus P_2$. Проверка этого условия устройством В даёт положительный результат, поэтому переменной E (признаку ошибки) присваивается нулевое значение ($E = 0$).

Далее (рис. 5, *з*) анализируются составные части пакетов $Z2 — Z4$. Так как ошибок по-прежнему нет, то условие $Y4 = P2 \oplus P3$ выполняется, $E = 0$.

После этого в поле зрения устройства В попадает очередная триада пакетов: $Z3 — Z5$ (рис. 5, *д*). Поле зрения, как видим, постоянно перемещается вправо и, по сути, представляет собой “скользящее окно”, выделенное на рисунке утолщённой изогнутой линией. Так как контрольное поле $Y5$ кадра $Z5$ содержит, как минимум, один ошибочный бит (устройство В пока не “знает” об этом), проверка условия $Y5 = P3 \oplus P4$ даёт отрицательный результат, признак ошибки E устанавливается в состояние лог. 1.

В данной ситуации устройство В может “рассуждать” следующим образом. Пакет $P3$ безошибочен, иначе на предыдущем шаге проверка условия $Y4 = P2 \oplus P3$ оказалась бы неудачной. Одновременное искажение пакета $P4$ и контрольного поля $Y5$ исключено, так как в силу ограниченности корректирующих возможностей метода рассматриваются только те ситуации, при которых ошибкам подвергаются лишь достаточно удалённые друг от друга пакеты Z_i (например, $Z5$ и $Z9$). Поэтому под подозрение в наличии искажений попадает либо пакет $P4$, либо контрольное поле $Y5$. Уточнить истинное положение вещей позволяют последующие шаги перемещения “скользящего окна”.

При проверке последующих условий: $Y6 = P4 \oplus P5$ (рис. 5, *е*) и $Y7 = P5 \oplus P6$ (рис. 5, *ж*) также обнаруживается их невыполнение (признак ошибки E в третий раз подряд устанавливается в состояние лог. 1). И, наконец, проверка условия $Y8 = P6 \oplus P7$ (рис. 5, *з*) даёт положительный результат.

Таким образом, ошибочный пакет $Z5$ при просмотре области, в которой он находится, оставил уникальный “отпечаток” в виде последовательности признаков $E: \dots 01110\dots$, которая позволяет однозначно определить его местоположение. Точнее, как следует из рис. 5, *з*, при обнаружении признака $E = 0$ после комбинации $\dots 0111$ можно утверждать, что ошибочный пакет ($Z5$) расположен в непосредственной близости от левой границы “скользящего окна”. Теперь устройство В может приступить к коррекции пакета $P5$ и локализации следующего ошибочного пакета $Z9$.

Для исправления ошибочного пакета $P5$ можно использовать одно из условий: $Y6 = P4 \oplus P5$ или $Y7 = P5 \oplus P6$, в которых присутствует переменная $P5$. Из первого условия следует, что $P5 = P4 \oplus Y6$. В правой части этого равенства содержатся безошибочные коды одинаковой разрядности, равной длине пакета $P5$. В результате поразрядного суммирования этих кодов по модулю два (используется операция Иключающее ИЛИ), формируется исправленный пакет $P5$.

Между текущим и последующим ошибочными пакетами ($Z5$ и $Z9$) должны располагаться, как минимум, три безошибочных пакета. В противном случае, единичные признаки ошибок (E), сформированные при просмотре близлежащих искажённых пакетов, будут “склеены” между собой, и распознавание ошибочной ситуации окажется неоднозначным.

Вторая кодовая ситуация (рис. 6) — ошибки присутствуют в первой половине пакета Z_i

Если в избыточном пакете $Z5$ ошибочен только исходный пакет $P5$, то, как показано на рис. 6, после его просмотра сквозь “скользящее окно” он оставляет уникальный “отпечаток” из признаков E : ...0110... . Это по аналогии с предыдущей кодовой ситуацией (рис. 5) однозначно определяет местоположение ошибочного пакета. Для его исправления, как и в предыдущей кодовой ситуации, можно применить одну из двух формул: $P5 = P4 \oplus Y6$ или $P5 = P6 \oplus Y7$.

Между текущим и последующим ошибочными пакетами ($Z5$ и $Z8$) должны располагаться, как минимум, два безошибочных пакета ($Z6$ и $Z7$).

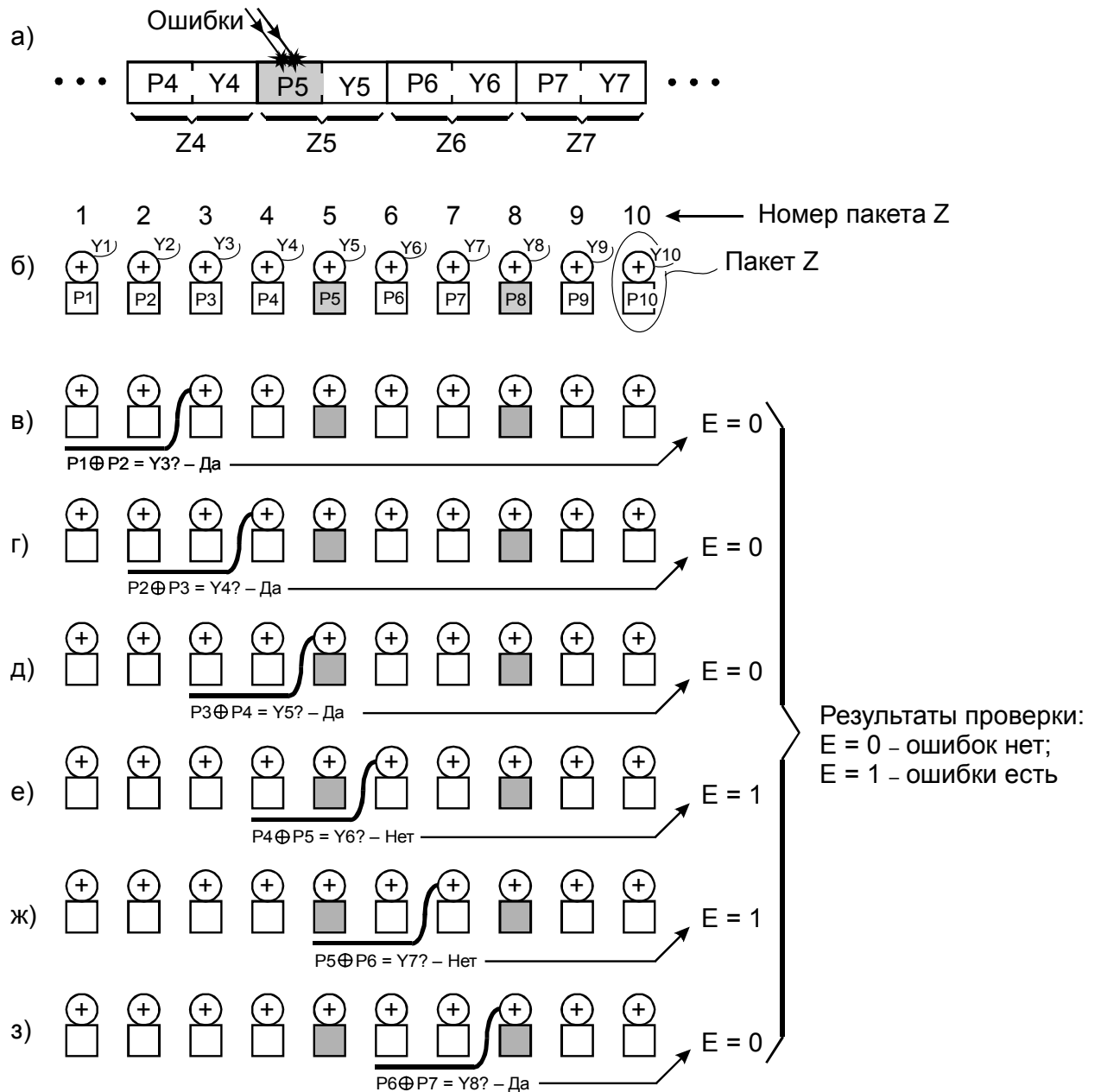


Рис. 6. Приём, размещение (а, б) и анализ (в — з) группы пакетов Z_i устройством-приёмником (вторая кодовая ситуация)

Третья кодовая ситуация (рис. 7) — ошибки присутствуют во второй половине пакета Z_i

Если в избыточном пакете $Z5$ ошибочно только контрольное поле $Y5$ (рис. 7), то в результате перемещения “скользящего окна” через этот пакет формируется признак $E = 1$, обрамлённый с обеих сторон логическими нулями. Этот признак в данном случае игнорируется, так как исходные пакеты переданы правильно, а в этом, собственно, и состоит задача. Ошибочные контрольные поля $Y5, Y7, Y9$ не корректируются, так как они далее не используются устройством В.

Между текущим и последующим ошибочными пакетами ($Z5$ и $Z7$) должен располагаться, как минимум, один безошибочный пакет ($Z6$).

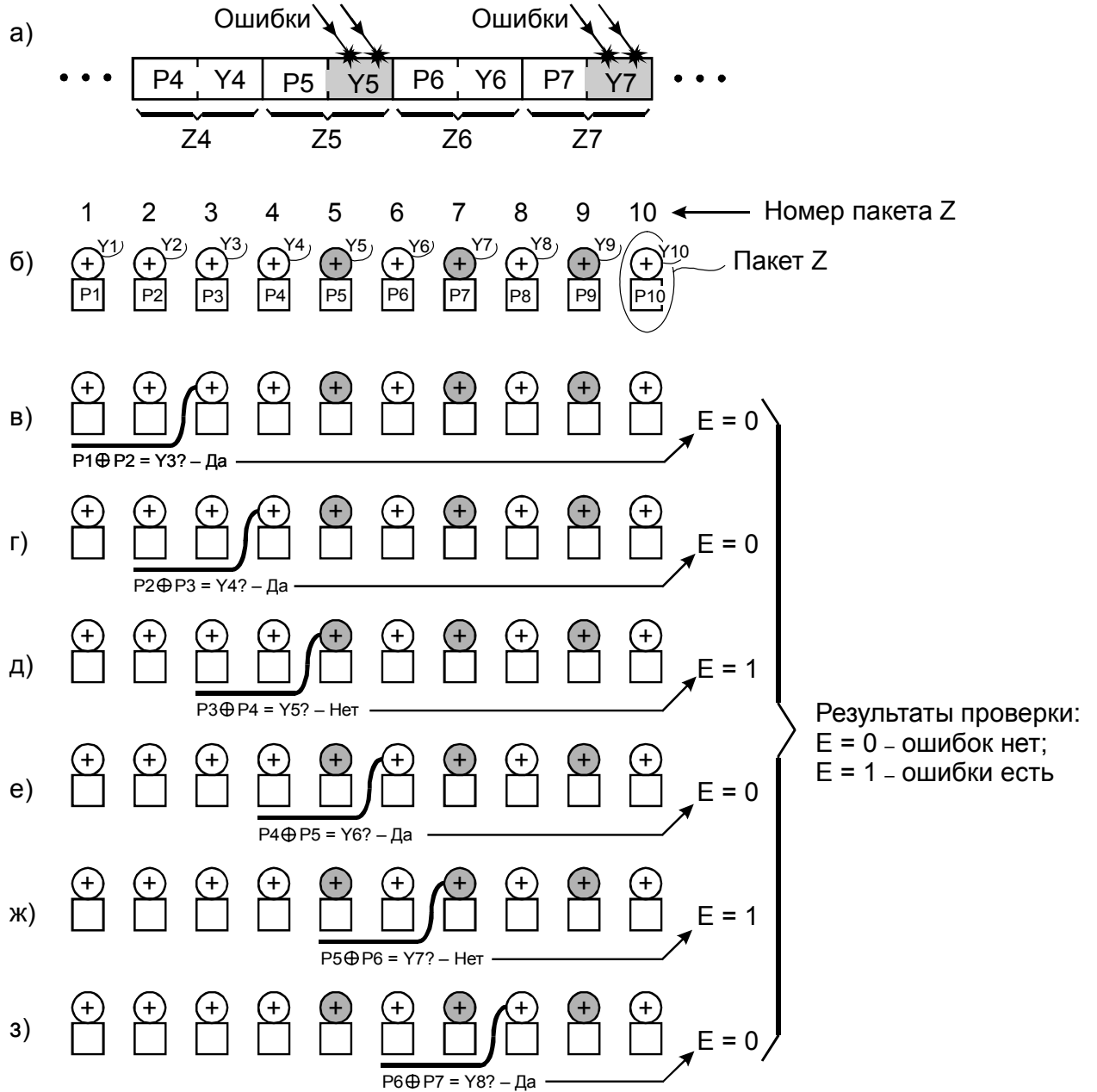


Рис. 7. Приём, размещение (а, б) и анализ (в — з) группы пакетов Z_i устройством-приёмником (третья кодовая ситуация)

2. На этапах продвижения окна через ошибочный пакет P5 (рис. 8, δ — ε) признак E последовательно принимает следующие значения: 0, 1, 1, 0. Эта ситуация подобна показанной на рис. 6. Устройство В делает вывод о том, что ошибки затронули пакет P5 и исправляет его применением одной из двух формул: $P5 = P4 \oplus Y6$ или $P5 = P6 \oplus Y7$.

4. Заключение

В этой статье рассмотрены два метода исправления групповых ошибок, “зона поражения” которых может достигать длины передаваемого по линии связи пакета. В эту зону может попадать либо один пакет, либо части двух смежных пакетов.

Избыточность, вносимая в поток данных при использовании первого метода, точно такая же, как и при обычном использовании циклического избыточного кода CRC. В отличие традиционных методов, предлагается контролировать не строки, а столбцы матрицы, образованной “штабелем” из передаваемых пакетов. Недостатки метода — относительно большая сложность вычислений и значительная ёмкость буферной памяти, что вносит нежелательную задержку передачи потока данных.

Второй метод применим к передаче данных в реальном времени, так как уменьшена задержка потока данных. Вносимая в поток избыточность составляет 100%, но значительно уменьшен объём буферной памяти и существенно снижена сложность вычислений.

ЛИТЕРАТУРА:

1. Пат. США № 5608738 <http://www.uspto.gov>
2. Пат. США № 6226769 B1 <http://www.uspto.gov>
3. С.М.Сухман, А.В.Бернов, Б.В.Шевкопляс. Синхронизация в телекоммуникационных системах. Анализ инженерных решений. — М.: Эко-Трендз, 2003. Электронная версия книги: http://lit.lib.ru/s/shewkopljass_b_w/

Б. Шевкопляс