РАЗРАБОТКА И АНАЛИЗ МЕТОДИКИ ПОСТРОЕНИЯ МНОЖЕСТВА ПЛОТНЫХ ИНФОРМАЦИОННЫХ СОВОКУПНОСТЕЙ ДЛЯ ИСПРАВЛЕНИЯ ПАКЕТОВ ОШИБОК

DOI: 10.36724/2072-8735-2024-18-10-20-26

Manuscript received 05 September 2024;
Accepted 07 October 2024

Исаева Мария Николаевна,
Санкт-Петербургский государственный университет аэрокосмического приборостроения, г. Санкт-Петербург, Россия, **imn@guap.ru**

Ключевые слова: *помехоустойчивое кодирование, декодирование по информационным совокупностям, исправление пакетов ошибок, каналы с памятью, плотные информационные совокупности, линейные коды, граница Рейгера*

В данной статье рассматривается задача исправления одиночных пакетов ошибок случайными линейными кодами. Предполагается, что используется модификация декодирования по информационным совокупностям для исправления пакетов ошибок, в которой количество информационных совокупностей линейно зависит от длины кода. Целью работы является исследование возможности снижения количества информационных совокупностей при таком декодировании, а также предложение процедуры для построения уменьшенного множества информационных совокупностей. В работе рассматривается понятие плотных информационных совокупностей с ограничением их диаметра. Предлагается методика построения множества плотных информационных совокупностей на основе анализа количества безопасных позиций начала пакета, гарантирующая исправление любых одиночных пакетов ошибок в пределах корректирующей способности кода. Проводится оценка количества плотных информационных совокупностей, которые могут быть получены по предложенной методике в зависимости от длины и скорости кода, корректирующей способности, а также заданного диаметра плотных информационных совокупностей. Получена нижняя граница для количества таких плотных информационных совокупностей. Проведена экспериментальная оценка вероятности построения множества плотных информационных совокупностей в зависимости от скорости кода и заданного диаметра. Анализ количества плотных информационных совокупностей, полученных по предложенной методике, для кодов, лежащих на границе Рейгера, показывает, что возможно увеличение диаметра плотных информационных совокупностей при незначительном увеличении их количества, дающее стремление вероятности построения множества плотных информационных совокупностей к значениям, близким к единице. При этом, уменьшение числа плотных информационных совокупностей по сравнению с первоначальным множеством без ограничения диаметра составляет от десятков до сотен раз в зависимости от скорости кода. Полученные результаты позволят существенно снизить сложность исправления пакетов ошибок по информационным совокупностям в пределах корректирующей способности кода в каналах с памятью.

Информация об авторе
Исаева Мария Николаевна, *Санкт-Петербургский государственный университет аэрокосмического приборостроения, Кафедра инфокоммуникационных технологий и систем связи, старший преподаватель, г. Санкт-Петербург, Россия. ORCID: 0009-0007-6228-0617*

Для цитирования:
Исаева М.Н. Разработка и анализ методики построения множества плотных информационных совокупностей для исправления пакетов ошибок // T-Comm: Телекоммуникации и транспорт. 2024. Том 18. №10. С. 20-26.

For citation:
Isaeva M.N. (2024). Development and analysis of a method for constructing dense information sets for error bursts correction. *T-Comm*, vol. 18, no.10 pp. 20-26. (*in Russian*)

Введение

В последнее десятилетие особое развитие получил процесс цифровизации в различных областях деятельности. Внедрение цифровизации требует высоких производительных мощностей и соблюдения ряда условий для хранения и обработки информации. Среди основных задач в рамках обработки информации является её передача по каналам связи. В связи с возникновением различных помех при передаче по каналам связи информация (передаваемая, например, в виде двоичной последовательности) может быть искажена. Для исправления возникших ошибок активно используются методы помехоустойчивого кодирования, таким образом, к полезной информации добавляется избыточность, используемая на приёмной стороне. Среди наиболее широко используемых сегодня помехоустойчивых кодов можно выделить линейные коды, в частности, БЧХ-коды, коды Рида-Соломона, коды с малой плотностью проверок на чётность, полярные коды, а также свёрточные и турбо-коды [1-3].

Исправление возникших ошибок происходит на этапе декодирования принятой двоичной последовательности. Все перечисленные выше классы помехоустойчивых кодов предназначены для борьбы с независимыми ошибками, в то время как зачастую особенности систем передачи, а также физические свойства среды распространения приводят к группированию ошибок. Известным из теории информации результатом является то, что учёт особенностей влияния шума на передаваемую последовательность (в частности, пакетирование ошибок) способен повысить предельно достижимые характеристики передачи. При этом в классической теории кодирования задаче исправления пакетов ошибок уделяется достаточно мало внимания в связи со сложностью анализа каналов с памятью. Некоторые вопросы построения методов кодирования и декодирования при исправлении пакетов рассматривались в [4, 5], в [6-8] изучаются вопросы использования низкоплотностных кодов для борьбы с такими ошибками.

В рамках данной статьи рассматривается задача исправления одиночных пакетов ошибок без учета структуры используемого помехоустойчивого кода, т.е. используются наиболее общие случайные линейные коды. Известный из теории кодирования подход, связанный с декодированием по информационным совокупностям [9] имеет экспоненциальную сложность [10, 11] при исправлении независимых ошибок, что вызвано необходимостью построения экспоненциально большого множества информационных совокупностей. Однако существуют работы, посвященные методам исправления пакетов ошибок с помощью информационных совокупностей, использующих множество информационных совокупностей линейной мощности [12]. Ещё более значительное уменьшение количества требуемых для декодирования пакетов информационных совокупностей может быть получено введением понятия плотной информационной совокупности [13]. Вместе с тем, не разработанным остается вопрос о конкретных методах построения множества плотных информационных совокупностей и их оценке.

В данной статье предлагается методика построения множества плотных информационных совокупностей для исправления одиночных ошибок при использовании случайных линейных кодов, а также проводится оценка минимального количества плотных информационных совокупностей, полученных по данной методике.

Понятие множества плотных информационных совокупностей

Рассмотрим двоичный линейный (n, k)-код, задаваемый $(k\times n)$ порождающей матрицей $\mathbf{G}$. Кодовыми словами такого кода являются двоичные последовательности $\mathbf{a} = (\alpha_1, \ldots, \alpha_n)$. Множество позиций $\chi = \{1 \le j_1 < j_2 < \ldots < j_k \le n\}$ называется информационной совокупностью, если значения $\alpha_{j_1}, \ldots, \alpha_{j_k}$ однозначно определяют кодовое слово $\mathbf{a}$.

Для того, чтобы определить, является ли случайное множество χ из k позиций информационной совокупностью, достаточно рассмотреть k столбцов порождающей матрицы, соответствующих данному множеству позиций. Если составить из них новую матрицу $\mathbf{M}_\chi$ и ее ранг будет равен k, то множество χ – информационная совокупность. Зная значения $\alpha_{j_1}, \ldots, \alpha_{j_k}$, соответствующее кодовое слово может быть восстановлено как $\mathbf{a} = \left(\alpha_{j_1}, \ldots, \alpha_{j_k}\right)\mathbf{M}_\chi^{-1}\mathbf{G}$.

Если при передаче по каналу связи кодового слова ошибочные биты не оказались на позициях информационной совокупности, она называется свободной от ошибок и с её помощью кодовое слово может быть корректно восстановлено. Соответственно, для декодирования по информационным совокупностям необходимо построить такое множество $X = \{\chi_1, \ldots, \chi_T\}$ из T информационных совокупностей, среди которых с высокой вероятностью найдется свободная от ошибок. При исправлении независимых ошибок для обеспечения низкой вероятности ошибки декодирования такое множество будет экспоненциально большим. Однако, при декодировании пакетов ошибок возможно обеспечить наличие не затронутой ошибками информационной совокупности, если построить информационную совокупность для каждого возможного расположения пакета.

В рамках данной статьи будем говорить, что некоторый вектор образует пакет длиной b, если его первый ненулевой элемент отстоит от последнего ненулевого элемента на b позиций. Для линейного кода может быть определено (возможно, с экспоненциальной сложностью) множество корректно исправляемых векторов ошибки [3]. Если все пакеты длиной до некоторого значения $b_{\max}$ входят в это множество, код может исправить все пакеты длиной до $b_{\max}$ и величину $b_{\max}$ назовем корректирующей способностью кода (для пакетов ошибок).

В соответствии с границей Рейгера [3], для любого линейного (n,k)-кода $b_{\max} \le (n - k)/2$. Существует процедура [14], с помощью которой для произвольного линейного кода значение $b_{\max}$ может быть определено с полиномиальной сложностью. В данной статье рассматриваются однократные пакеты размера не более $b_{\max}$, за пределами пакета ошибки отсутствуют. При генерации информационных совокупностей для каждого расположения пакета число информационных совокупностей ограничивается значением $T \le n - b_{\max} + 1$, что обеспечивает полиномиальную сложность декодирования. При этом, сложность декодирования определяется не только числом информационных совокупностей, а и многократном восстановлением кодовых слов по информационным совокупностям, включающем матричные умножения, а также необходимостью хранить матрицы $\mathbf{G}_\chi = \mathbf{M}_\chi^{-1}\mathbf{G}$. Таким образом, актуальной является задача сокращения мощности T

множества X информационных совокупностей, что может быть достигнуто, если одна информационная совокупность может быть использована для нескольких разных расположений пакетов ошибок.

Для решения данной задачи введем понятие плотной информационной совокупности [13]. Пусть информационная совокупность построена для некоторого пакета, расположенного на позициях от b_1 до b_2. Это означает, что номера позиций, вошедших в информационную совокупность, находятся за пределами интервала $[b_1, b_2]$. Перенумеруем позиции информационной совокупности таким образом, что j_1 – первая позиция, большая b_2, j_k – последняя (возможно, с учетом циклического порядка обхода позиций). Назовем диаметром информационной совокупности величину $j_k - j_1 + 1$, если $j_1 < j_k$, и $j_k + n - j_1 + 1$, в противном случае. Будем говорить, что множество X состоит из плотных информационных совокупностей, если диаметр каждой из них не превышает $k + \Delta$ для некоторого заданного Δ. В данной статье приводится методика построения множества плотных информационных совокупностей, а также оценка числа плотных информационных совокупностей, полученных с помощью данной методики.

Вероятность нахождения множества плотных информационных совокупностей для случайных кодов

Задача поиска информационной совокупности аналогична поиску невырожденной матрицы размера $k \times k$. Известно [15], что для случайных матриц вероятность нахождения невырожденной матрицы (нахождения информационной совокупности) для $\Delta = 0$ будет:

$$Q_0 = \prod_{j=1}^{\infty}\left(1-\frac{1}{2^j}\right) \approx 0,288. \qquad (1)$$

Обобщенная формула для $\Delta \geq 0$ выглядит следующим образом:

$$P(\operatorname{rank}(\mathbf{M}_{k,k+\Delta}) = k) = Q_\Delta = \prod_{j=\Delta+1}^{\infty}\left(1-\frac{1}{2^j}\right). \qquad (2)$$

Формулы (1) и (2) фактически соответствуют вероятности нахождения плотной информационной совокупности диаметром $k + \Delta$ в случайной порождающей матрице. Значения вероятностей Q_Δ для некоторых значений дельта приведены в Таблице 1. Данные значения вероятности справедливы для нахождения одной информационной совокупности.

Таблица 1

Вероятность нахождения плотной информационной совокупности

Δ	0	1	2	3	4	5	6	7	8
Q_Δ	0,29	0,58	0,77	0,88	0,94	0,98	0,98	0,99	0,99

Если нам необходимо найти T информационных совокупностей, то формула (2) примет следующий вид:

$$Q_{\Delta_T} = (Q_\Delta)^T = \left(\prod_{j=\Delta+1}^{\infty}\left(1-\frac{1}{2^j}\right)\right)^T,$$

при условии, что нахождение информационных совокупностей – это независимые события.

Приведенные выражения могут использоваться для оценки вероятности построения множества из T плотных информационных совокупностей в случайной матрице при различных значениях Δ.

Методика построения множества плотных информационных совокупностей

Из определения информационной совокупности следует, что для корректного декодирования пакетов ошибок необходимо, чтобы все позиции пакета ошибок не попадали на позиции информационной совокупности. Соответственно, для декодирования пакетов длиной b_{max} необходимо сформировать такое множество информационных совокупностей, чтобы для любого такого пакета нашлась такая информационная совокупность χ, позиции которой не находятся внутри позиций пакета (т.е. данный пакет может быть исправлен данной информационной совокупностью). Будем рассматривать случайный код длиной n, позиции кодовых слов которого пронумерованы от 1 до n.

Назовем позицию начала пакета длиной b_{max} безопасной, если из множества информационных совокупностей, используемого для декодирования, найдется хотя бы одна информационная совокупность χ_i, позиции которой находятся вне позиций пакета. Тогда задача декодирования состоит в том, чтобы построить такой набор информационных совокупностей $X = \{\chi_1, \ldots, \chi_T\}$, для которого все позиции кодового слова будут безопасны. При этом будем стремиться минимизировать количество информационных совокупностей. Так как в качестве корректирующей способности кода рассматривается максимальная длина исправляемого пакета b_{max}, то в дальнейших примерах будем ориентироваться именно на это значение при расположении плотных информационных совокупностей.

Предположим, что для случайного кода нахождение плотной информационной совокупности внутри любых подряд идущих $k + \Delta$ позиций (с учетом цикличности) является достоверным событием. Будем оценивать множество безопасных позиций для пакетов длиной b_{max} при заданных расположениях плотных информационных совокупностей длиной $k + \Delta$. Будем считать, что если позиция $n - b_{max} + 1$ безопасна (т.е. пакет длиной b_{max} находится в конце слова), то позиции от $n - b_{max} + 2$ до n также безопасны.

Рассмотрим информационную совокупность χ_1 на первых $k + \Delta$ позициях (рис. 1, а). Очевидно, что безопасными будут все позиции после χ_1 от $k + \Delta + 1$ до n, при этом все позиции внутри самой информационной совокупности небезопасны. Далее, рассмотрим информационную совокупность χ_2 на последних $k + \Delta$ позициях (рис. 1, б).

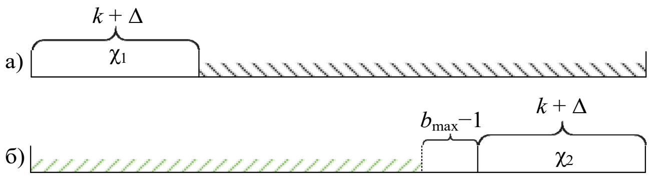

Рис. 1. а) Расположение информационной совокупности в начале кодового слова; б) Расположение информационной совокупности в конце кодового слова

Можно заметить, что теперь небезопасными являются не только позиции внутри χ_2, но также $b_{max} - 1$ позиций перед ее началом. Легко можно увидеть, что выбор двух плотных информационных совокупностей на первых и последних позициях кодового слова максимизирует число безопасных позиций.

Представленный на рисунке 1, а способ выбора информационных совокупностей (если он возможен) позволяет исправлять пакеты даже сверх корректирующей способности, однако, такой способ сложно обобщить на большее количество информационных совокупностей. Поэтому ограничимся рассмотрением исправления пакетов длиной, не превышающей b_{max}, и сформулируем еще один, легко обобщаемый способ построения множества информационных совокупностей.

Так как уже было замечено, что позиции внутри самой информационной совокупности небезопасны, вторая информационная совокупность не должна пересекаться с первой, иначе при их использовании останутся небезопасные позиции. Например, возьмем вторую информационную совокупность на позициях от $k + \Delta + 1$ до $2(k + \Delta)$, т.е. сразу после первой (рис. 2,а). Как было показано на рисунке 1,б, перед второй информационной совокупностью появляются небезопасные позиции, поэтому такое расположение информационных совокупностей не гарантирует исправление пакетов ошибок, начало которых попадает в данный интервал. Очевидно, что для того, чтобы все позиции кодового слова оказались безопасными, необходимо между первой и второй информационными совокупностями обеспечить интервал не меньше, чем $b_{max} - 1$ (рис. 2,б).

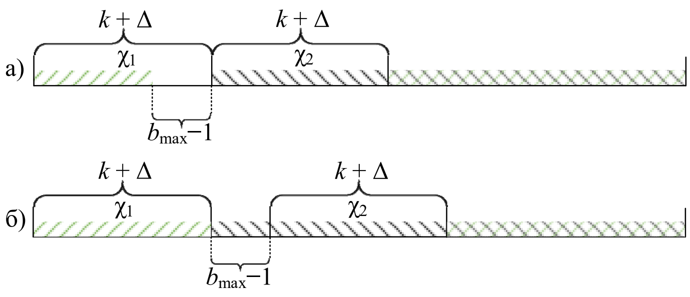

Рис. 2. а) Расположение двух информационных совокупностей без защитного интервала; б) Расположение двух информационных совокупностей с соблюдением защитного интервала

Оценим параметры кода, при которых указанный выбор двух информационных совокупностей гарантирует безопасность всех позиций. Очевидно, для этого должно выполняться неравенство $2k + b_{max} + (2\Delta - 1) \leq n$. Т.к. значение b_{max} в соответствии с границей Рейгера $b_{max} \leq (n - k)/2$, получим $2k + (n - k)/2 + (2\Delta - 1) \leq n$, отсюда

$$k \leq (n - 2(2\Delta - 1))/3. \qquad (3)$$

С учётом скорости кода $R = k/n$, поделим обе части полученного неравенства на n:

$$R \leq 1/3 - 2(2\Delta - 1)/3n.$$

Если значение $2(2\Delta - 1)/3n$ мало, то две плотные информационные совокупности обеспечат безопасность всех кодовых позиций, если скорость кода R не превышает значения, приблизительно равного 1/3.

Рассмотрим случай, когда условие (3) не выполняется, т.е. двух плотных информационных совокупностей недостаточно для обеспечения безопасности всех позиций. Можно заметить, что минимальное число небезопасных позиций в этом случае обеспечивается, если одна информационная совокупность находится в начале слова, а вторая в конце (рис. 3,а).

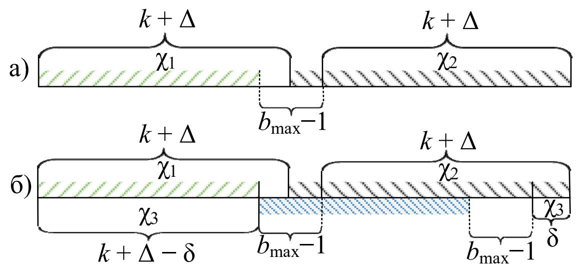

Рис. 3. а) Расположение двух информационных совокупностей, которые не гарантируют безопасность всех позиций; б) Расположение трех информационных совокупностей

Расположим третью информационную совокупность таким образом, чтобы её последняя позиция находилась перед началом небезопасного участка, т.е. выберем информационную совокупность слева от него (рис. 3,б). Согласно рисунку, для того чтобы трех плотных информационных совокупностей было достаточно для обеспечения безопасности всех позиций пакетов ошибок внутри кодового слова, необходимо соблюсти интервал между началом третьей информационной совокупности и концом первой не менее, чем $b_{max} - 1$, иначе возникнут небезопасные позиции.

Приведённые примеры могут быть обобщены до следующей методики. Выберем первую плотную информационную совокупность длиной $k + \Delta$ на первых позициях кодового слова. Начало второй плотной информационной совокупности выберем, выдержав защитный интервал в $b_{max} - 1$ позиций между информационными совокупностями. Аналогичным образом выберем третью плотную информационную совокупность, соблюдая интервал длиной $b_{max} - 1$ после второй. Данная процедура должна продолжаться, пока все позиции кодового слова не окажутся безопасными.

На рисунке 4, а продемонстрирован пример расположения трех информационных совокупностей в соответствии с описанной методикой. С увеличением $k + \Delta$ (рис. 4,б) информационная совокупность χ_3 не обеспечивает безопасность оставшихся позиций, соответственно, требуется построить четвёртую информационную совокупность.

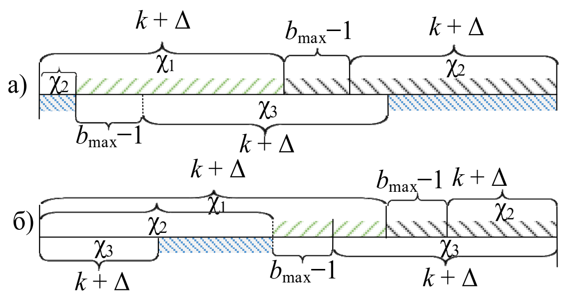

Рис. 4. а) Расположение трех информационных совокупностей в соответствии с методикой; б) Расположение трех информационных совокупностей, при котором не обеспечивается безопасность всех позиций

Оценка количества плотных информационных совокупностей

Для методики, разработанной в предыдущем разделе, оценим количество получаемых плотных информационных совокупностей, в зависимости от параметров кода: длины n и скорости R, а также значения дельта.

Обобщая оценку (3), в соответствии с которой для двух информационных совокупностей необходимо, чтобы выполнялось неравенство $2(k+\Delta)+(b_{max}-1)\le n$, для трех плотных информационных совокупностей можно получить выражение:

$$3(k+\Delta)+2(b_{max}-1)\le 2n.$$

Продолжая обобщение, пусть T – число информационных совокупностей, которые обеспечивают безопасность всех позиций расположения пакетов ошибок, тогда справедливо неравенство:

$$T(k+\Delta)+(T-1)(b_{max}-1)\le (T-1)n. \quad (4)$$

Из выражения (4) получим нижнюю границу для числа плотных информационных совокупностей, построенных по методике, описанной выше:

$$T\ge\left\lceil\frac{n-b_{max}+1}{n-b_{max}+1-(k+\Delta)}\right\rceil. \quad (5)$$

Если предположить, что используются коды, лежащие на границе Рейгера, тогда $b_{max}=(n-k)/2$, что дает границу:

$$T\ge\left\lceil\frac{n+nR+2}{n+nR+2-2(nR+\Delta)}\right\rceil. \quad (6)$$

Выражения (5) и (6) могут использоваться для оценки числа плотных информационных совокупностей в зависимости от параметров кода. Следует учитывать, что в приведенной методике предполагается, что любые $k+\Delta$ последовательные (с учетом цикличности просмотра) позиции кодового слова гарантированно содержат информационную совокупность. С учетом таблицы 1 можно выбрать значения дельты, обеспечивающие высокую вероятность данного события, однако в действительности для заданной матрицы это предположение может не выполняться. В этом случае необходимо увеличивать значение дельта, что может привести к увеличению числа плотных информационных совокупностей в соответствии с (5) и (6).

На рисунке 5 приведена зависимость минимального значения T, полученного по формуле (6), в зависимости от скорости кода, при $n = 1000$ и предположении, что код лежит на границе Рейгера. Значения дельта выбраны равными 0 и 20. Также на графике приведены значения $T_{max}=(n+nR+2)/2$, соответствующие количеству информационных совокупностей без введения свойства плотности, для всех возможных положений пакета и кодов, лежащих на границе Рейгера.

Как можно видеть из графика, при $\Delta = 20$ возрастание T происходит ранее, чем для $\Delta = 0$, однако эта разница незначительна. Таким образом, для не слишком высоких скоростей (на приведенном графике до 0,7) можно увеличивать значение Δ, увеличивая вероятность построения множества плотных информационных совокупностей, без значительного увеличения их числа T.

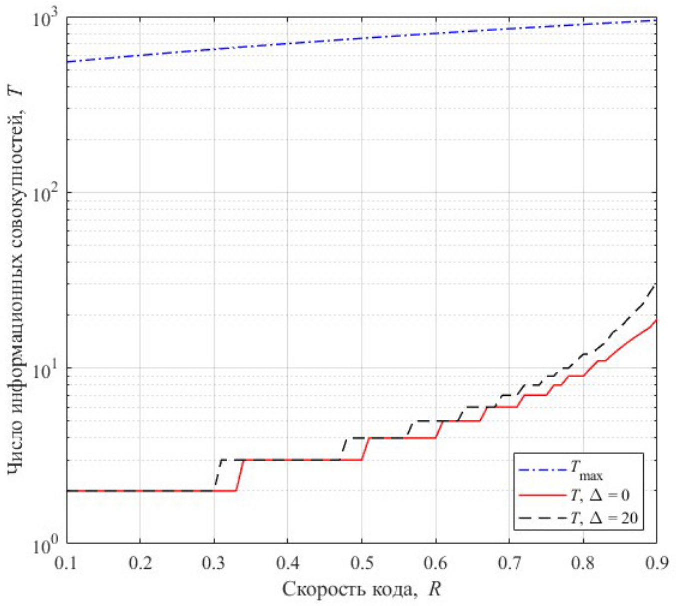

Рис. 5. Зависимость количества плотных информационных совокупностей от скорости кода при $n = 1000$

По сравнению с изначальным количеством информационных совокупностей без учета их диаметра уменьшение количества составляет от нескольких сотен раз для низкоскоростных кодов до десятков раз для кодов с высокой скоростью.

В таблице 2 приведены вероятности нахождения множества плотных информационных совокупностей для некоторых скоростей кодов и значений Δ при длине кода $n = 1000$. Как можно видеть, при минимальном диаметре информационной совокупности ($\Delta = 0$) вероятность построения множества может быть достаточно мала, особенно с ростом скорости кода, однако можно обеспечить высокую вероятность построения увеличением значения Δ при незначительном росте числа T.

Таблица 2

Вероятность нахождения множества плотных информационных совокупностей при $n = 1000$

		R						
		0,1	0,2	0,3	0,4	0,5	0,6	0,7
Δ	0	0,08	0,08	0,08	0,02	0,02	7e–3	6e–4
	5	0,96	0,96	0,96	0,94	0,92	0,90	0,89
	20	0,98	0,98	0,98	0,97	0,96	0,95	0,93

В случае, если для заданного Δ и конкретной порождающей матрицы построить множество плотных информационных совокупностей не удается, а увеличение значения Δ по каким-то причинам является нежелательным можно продолжать построение информационных совокупностей в соответствии с описанной в данной статье методикой, пока все символы кодового слова не окажутся безопасны. Конечно, это приведет к некоторому увеличению числа информационных совокупностей, однако на основании проведенного анализа можно предположить, что такое увеличение будет невелико.

Заключение

В данной статье рассматривается задача исправления одиночных пакетов ошибок случайными кодами.

Для специального класса плотных информационных совокупностей предложена методика их построения с целью минимизации их количества и гарантирования исправления пакетов вплоть до корректирующей способности кода. Построение множества плотных информационных совокупностей носит вероятностный характер. В статье оценена вероятность успешного построения в зависимости от диаметра информационных совокупностей. Проведена оценка количества плотных информационных совокупностей, полученных по данной методике. С помощью предложенной методики удается достичь уменьшения количества информационных совокупностей от десятков до сотен раз в зависимости от скорости кода.

Возможным направлением дальнейших исследований является разработка и анализ методики построения плотных информационных совокупностей для кодов, обладающих специальной структурой, например, квазициклических блочно-перестановочных кодов с малой плотностью проверок на четность.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение № FSRF-2023-0003, «Фундаментальные основы построения помехозащищённых систем космической и спутниковой связи, относительной навигации, технического зрения и аэрокосмического мониторинга».

Литература

1. *Lin S., Li J.* Fundamentals of Classical and Modern Error-Correcting Codes // Cambridge: Cambridge University Press; 2022, 840 p.

2. *Ke X.* Coding Theory in Optical Wireless Communication Systems: Volume II in Optical Wireless Communication Theory and Technology. Singapore: Springer Nature Singapore, 2024. doi: 10.1007/978-981-97-2382-9.

3. *Moon T.K.* Error correction coding: Mathematical methods and algorithms. Wiley, 2020 992 p.

4. *Ovchinnikov A.A., Veresova A.M., Fominykh A.A.* Decoding of linear codes for single error bursts correction based on the determination of certain events // Information and Control Systems, 2022, no. 6, pp. 41-52. doi:10.31799/1684-8853-2022-6-41-52.

5. *Kim C., Kim J.-W., No J.-S.* New Design of Error Control Codes Resilient to Single Burst Error or Two Random Bit Errors Using Constacyclic Codes // IEEE Access, vol. 10, pp. 131101-131108, 2022, doi: 10.1109/ACCESS.2022.3229427.

6. *Nakamura Y., Akamatsu T., Nishikawa M., Okamoto Y.* Performance Evaluation of Burst Error Correction by LDPC Coding and Iterative Decoding System in Magnetic Tape Drive // IEEE Transactions on Magnetics, vol. 59, no. 3, pp. 1-5, March 2023, Art no. 3000205, doi: 10.1109/TMAG.2022.3201898.

7. *Li L., Lv J., Li Y., Dai X., Wang X.* Burst Error Identification Method for LDPC Coded Systems // IEEE Commun. Lett., pp. 1-5, 2024, doi: 10.1109/LCOMM.2024.3391826.

8. *Yang M., Pan Z., Djordjevic I.B.* FPGA-based burst-error performance analysis and optimization of regular and irregular SD-LDPC codes for 50G-PON and beyond // Opt. Express, vol. 31, no. 6, p. 10936, Mar. 2023, doi: 10.1364/OE.477546.

9. *Prange E.* The use of information sets in decoding cyclic codes // IRE Transactions on Information Theory, vol. 8, no. 5, pp. 5-9, September 1962, doi: 10.1109/TIT.1962.1057777.

10. *Barg A., Krouk E., van Tilborg H.C.A.* On the Complexity of Minimum Distance Decoding of Long Linear Codes // IEEE Trans. Inform. Theory. 1999. Vol. 45. No. 5, pp. 1392-1405. doi: 10.1109/18.77114.

11. *Both L., May A.* Optimizing BJMM with nearest neighbors: full decoding in 22/21n and McEliece security // WCC workshop on coding and cryptography. 2017. V. 214.

12. *Исаева М.Н.* Поиск информационных совокупностей при исправлении пакетов ошибок квазициклическими кодами // T-Comm: Телекоммуникации и транспорт. 2023. Том 17. №7. С. 4-12.

13. *Исаева М.Н., Овчинников А.А.* Исправление одиночных пакетов ошибок за пределами корректирующей способности кода с использованием информационных совокупностей // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 1. С. 70-80. doi: 10.17586/2226-1494-2024-24-1-70-80

14. *Veresova A.M., Isaeva M.N., Ovchinnikov A.A.* Estimation of Independent Errors and Bursts Correction Capability of Linear Codes // 2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon), Saint Petersburg, Russian Federation, 2024, pp. 23-27, doi: 10.1109/ElCon61730.2024.10468456.

15. *Berlekamp E.R.* The technology of error correcting codes, Proc. IEEE, vol. 68, pp. 564-593, 1980.

DEVELOPMENT AND ANALYSIS OF A METHOD FOR CONSTRUCTING DENSE INFORMATION SETS FOR ERROR BURSTS CORRECTION

Maria N. Isaeva, *Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia,* **imn@guap.ru**

Abstract

In this article the problem of correction of single error bursts by random linear codes is considered. It is supposed that the modification of decoding on information sets for error bursts correction is used, in which the number of information sets linearly depends on the code length. The purpose of this work is to research the possibility of reduction of quantity of information sets at such decoding, and also to offer the procedure for construction of the reduced amount of information sets. In this work the concept of dense information sets with restriction of their diameter is considered. The method of constructing a set of dense information sets after the analysis of quantity of safe positions of the beginning of a burst, guaranteeing correction of any single error bursts of within the correcting ability of a code is offered. Estimation of the number of dense information sets that can be obtained by the proposed method depending on the length and speed of the code, the corrective ability, and the given diameter of dense information sets is considered. The lower bound for the number of such dense information sets is obtained. Experimental estimation of the probability of building a set of dense information sets depending on the code speed and the given diameter is provided. The analysis of the number of dense information sets obtained by the proposed method for the codes lying on the Reiger bound shows that it is possible to increase the diameter of dense information sets at insignificant increase of their number, which gives the probability of building a set of dense information sets in values close to 1. In this case, the reduction of the number of dense information sets in comparison with the initial set without diameter limitation is from tens to hundreds times depending on the code speed. The obtained results will significantly reduce the complexity of error bursts correction by information sets within the corrective capacity of the code in channels with memory.

Keywords: *error-correcting coding, decoding by information sets, burst error correction, memory channels, dense information sets, linear codes, Reiger bound*

References

[1] S. Lin, J. Li, "Fundamentals of Classical and Modern Error-Correcting Codes," Cambridge: Cambridge University Press. 2022, 840 p.
[2] X. Ke, "Coding Theory in Optical Wireless Communication Systems," Volume II. in Optical Wireless Communication Theory and Technology. Singapore: Springer Nature Singapore, 2024. doi: 10.1007/978-981-97-2382-9.
[3] T.K. Moon, "Error correction coding: Mathematical methods and algorithms," Wiley, 2020. 992 p.
[4] A.A. Ovchinnikov, A. M. Veresova, A. A. Fominykh, "Decoding of linear codes for single error bursts correction based on the determination of certain events", *Information and Control Systems*, 2022, no. 6, pp. 41-52. doi:10.31799/1684-8853-2022-6-41-52
[5] C. Kim, J. -W. Kim and J. -S. No, "New Design of Error Control Codes Resilient to Single Burst Error or Two Random Bit Errors Using Constacyclic Codes", *IEEE Access*, vol. 10, pp. 131101-131108, 2022, doi: 10.1109/ACCESS.2022.3229427
[6] Y. Nakamura, T. Akamatsu, M. Nishikawa and Y. Okamoto, "Performance Evaluation of Burst Error Correction by LDPC Coding and Iterative Decoding System in Magnetic Tape Drive", *IEEE Transactions on Magnetics*, vol. 59, no. 3, pp. 1-5, March 2023, Art no. 3000205, doi: 10.1109/TMAG.2022.3201898
[7] L. Li, J. Lv, Y. Li, X. Dai, and X. Wang, "Burst Error Identification Method for LDPC Coded Systems," *IEEE Commun. Lett.*, pp. 1-5, 2024, doi: 10.1109/LCOMM.2024.3391826.
[8] M. Yang, Z. Pan, and I. B. Djordjevic, "FPGA-based burst-error performance analysis and optimization of regular and irregular SD-LDPC codes for 50G-PON and beyond", *Opt. Express,* vol. 31, no. 6, p. 10936, Mar. 2023, doi: 10.1364/OE.477546.
[9] E. Prange, "The use of information sets in decoding cyclic codes", *IRE Transactions on Information Theory*, 1962. Vol. 8, no. 5, pp. 5-9.
[10] A. Barg, E. Krouk, H.C.A. van Tilborg, "On the Complexity of Minimum Distance Decoding of Long Linear Codes", *IEEE Trans. Inform. Theory*. 1999. Vol. 45, No. 5, pp. 1392-1405. doi: 10.1109/18.771141
[11] L. Both, A. May, "Optimizing BJMM with nearest neighbors: full decoding in 22/21n and McEliece security", *WCC workshop on coding and cryptography.* 2017. Vol. 214.
[12] M.N. Isaeva, "Finding information sets when correcting error bursts with quasi-cyclic codes", *T-Comm*, vol. 17, no.7, pp. 4-12. (in Russian)
[13] M.N. Isaeva, A.A. Ovchinnikov, "Correction of single error bursts beyond the code correction capability using information sets", *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 1, pp. 70-80 (in Russian). doi: 10.17586/2226-1494-2024-24-1-70-80
[14] A. M. Veresova, M. N. Isaeva and A. A. Ovchinnikov, "Estimation of Independent Errors and Bursts Correction Capability of Linear Codes" 2024 *Conference of Young Researchers in Electrical and Electronic Engineering (ElCon),* Saint Petersburg, Russian Federation, 2024, pp. 23-27, doi: 10.1109/ElCon61730.2024.10468456.
[15] E.R. Berlekamp, "The technology of error correcting codes", *Proc. IEEE*, vol. 68, pp. 564-593, 1980.

Information about authors:
Maria N. Isaeva, *Saint-Petersburg State University of Aerospace Instrumentation, senior lecturer of the Department of Infocommunication Technologies and Communication Systems (Department 25), St. Petersburg, Russia. ORCID: 0009-0007-6228-0617*