ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

ИНФОРМАТИКА

INFORMATICS

УДК 621.391.01
https://doi.org/10.29235/1561-2430-2021-57-4-506-512

Поступила в редакцию 25.05.2021
Received 25.05.2021

И. Л. Кузнецова, А. С. Поляков

Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь

ЭФФЕКТИВНЫЙ АЛГОРИТМ ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ ПЕРЕДАВАЕМОЙ ИНФОРМАЦИИ

Аннотация. Рассматривается проблема обеспечения целостности передаваемой информации в современных информационно-коммуникационных системах. Предлагается оптимизированный алгоритм обнаружения и коррекции ошибок в информации, передаваемой по линиям связи, который разработан с учетом результатов ранее проведенных исследований способа коррекции ошибок на основе значений четности координат бинарной матрицы. Разработан легко реализуемый, быстродействующий и эффективный алгоритм обнаружения ошибок, ориентированный на применение бинарных матриц небольших размеров, например, (4 × 8) или (7 × 8) бит. В таких матрицах возможное количество ошибок, появляющихся в них при передаче информации, сравнительно невелико и легко обнаруживается.

Ключевые слова: бинарная матрица, четность координат матрицы, координаты ошибок, интенсивность ошибок, главные диагонали, вспомогательные диагонали

Для цитирования. Кузнецова, И. Л. Эффективный алгоритм обеспечения целостности передаваемой информации / И. Л. Кузнецова, А. С. Поляков // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2021. – Т. 57, № 4. – С. 506–512. https://doi.org/10.29235/1561-2430-2021-57-4-506-512

Irina L. Kuznetsova, Alexander S. Poljakov

United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

THE EFFECTIVE ALGORITHM OF ENSURING THE INTEGRITY OF THE TRANSMITTED INFORMATION

Abstract. The problem of ensuring the integrity of the transmitted information in modern information and communication systems is considered in this paper. An optimized algorithm for detecting and correcting errors in the information transmitted over communication lines is proposed. It was developed on the basis of the results of previous studies of the error correction method based on the parity values of the coordinates of a binary matrix. An easy-to-implement, high-speed and efficient error detection algorithm is proposed which is focused on the use of small binary matrices, for example, (4 × 8) or (7 × 8) bits. In such matrices, the possible number of errors that appear in them during the transfer of information is relatively small and easily detected.

Keywords: binary matrix, matrix coordinate parity, errors, error coordinates, intensity of bit errors, main diagonals, auxiliary diagonals

For citation. Kuznetsova I. L., Poljakov A. S. The effective algorithm of ensuring the integrity of the transmitted information. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2021, vol. 57, no. 4, pp. 506–512 (in Russian). https://doi.org/10.29235/1561-2430-2021-57-4-506-512

Введение. Сформировавшиеся в последние десятилетия информационно-коммуникационные технологии типа RFID-метки, Интернет вещей (Internet-of Things), Умный дом, Умный город (smart sity), беспилотные автомобили и т. п. предусматривают использование доступного, но не самого надежного вида связи – радиосвязи, поэтому существенно обострилась проблема обеспечения безошибочной передачи информации, к примеру от владельца Умного дома к ав-

томатизированным объектам дома. Поскольку любая ошибка в передаваемой информации может привести к непредсказуемым последствиям, например ДТП с беспилотным автомобилем, отключение обогрева квартиры вместо включения обогрева и т. п., то проблема обеспечения целостности информации, т. е. соответствия информации, предназначенной к передаче и полученной на приемной стороне, приобретает особенное значение.

В [1] был предложен способ обнаружения ошибок в передаваемой информации по значениям четности координат бинарной матрицы. Несколько позднее в [2] были представлены результаты экспериментальных исследований, подтвердившие высокую эффективность способа в сравнении с известными опубликованными материалами [3–8].

На основании анализа полученных в [2] результатов установлено, что эффективность способа повышается при уменьшении размеров рассматриваемых бинарных матриц, вследствие чего сделан вывод о целесообразности использования матриц малых размеров, например (4 × 8) или (7 × 8) бит. Действительно, при таких размерах матриц возможное количество ошибок, появляющихся в них при передаче информации, сравнительно невелико. Даже при высокой зашумленности канала связи, например при коэффициенте интенсивности ошибок 10^{-2}, вероятное количество возможных ошибок в матрицах размером (4 × 8) или (7 × 8) находится в пределах 1–3 ошибки. Такое количество ошибок рассматриваемым способом, как правило, обнаруживается всегда. С учетом выявленных особенностей был разработан представленный ниже алгоритм поиска ошибок в матрицах небольших размеров, который оказался существенно проще и эффективнее алгоритма, используемого в [2].

Основные положения рассматриваемого способа. Для понимания особенностей предлагаемого ниже алгоритма кратко представим основные положения способа обнаружения и коррекции ошибок при передаче информации по значениям четности координат бинарной матрицы (фактически это метод кодирования-декодирования информации).

Под *главными диагоналями* понимаются как основная главная диагональ матрицы, так и все параллельные ей диагонали, рассматриваемые как непрерывные цепочки элементов матрицы, начинающиеся с элементов первой строки с номерами 0,1,...,n–1 и проходящие в направлении «сверху – вниз – направо» через все строки матрицы с переходом на крайний левый элемент следующей строки при достижении крайнего правого элемента предыдущей строки. Нумерация элементов новой строки начинается с номера, который был последним в предыдущей строке.

Вспомогательными диагоналями являются основная вспомогательная диагональ матрицы и все параллельные ей диагонали, рассматриваемые как непрерывные цепочки, начинающиеся с элементов первой строки с номерами n–1,…,1,0 и проходящие в направлении «сверху – вниз – налево» через все строки матрицы с переходом на крайний правый элемент следующей строки при достижении крайнего левого элемента предыдущей строки.

В качестве примера на рис. 1 представлена матрица M размером ($r \times n$) бит, $r = 6$, $n = 7$. В левом верхнем углу элементов матрицы указаны номера главных диагоналей, в нижнем правом углу – вспомогательных диагоналей. Направления главных и вспомогательных диагоналей показаны стрелками m и a соответственно. Значками ■ обозначены ошибочные элементы матрицы, т. е. элементы, значения которых изменились при передаче информации.

В каждой матрице производится подсчет значений четности по всем перечисленным выше координатам. Полученные значения четности координат матрицы представляют собой исходную проверочную информацию (ИПИ), которая помещается в отдельный файл, передаваемый вместе с исходным массивом информации.

Полученная на приемной стороне информация снова разбивается на такие же бинарные матрицы, как и на передающей стороне. В полученных матрицах производится вычисление значений четности по всем координатам, которые сравниваются со значениями четности координат, переданными вместе с массивом данных (т. е. с ИПИ). В результате определяются координаты матриц, в которых появились ошибки при передаче (назовем их ошибочные координаты). Выявленные ошибочные значения координат матрицы включаются в списки ошибочных координат SX, SY, SM, SA (ошибочные строки, столбцы, главные и вспомогательные диагонали соответственно).

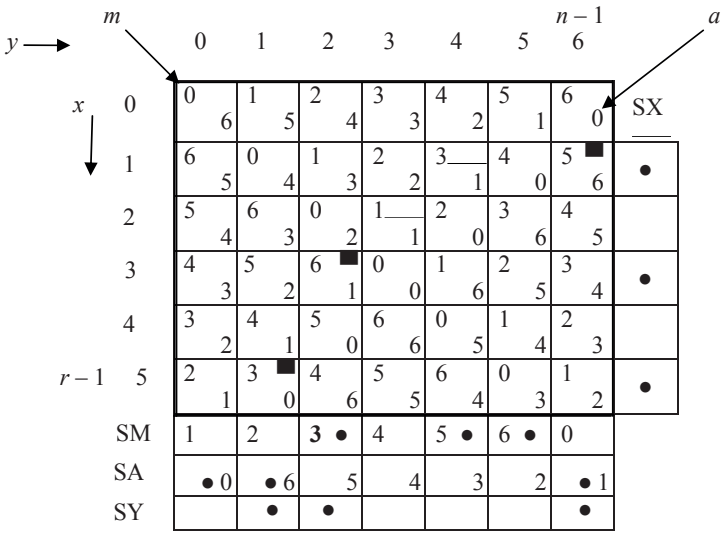

Рис. 1. Матрица *M*

Fig. 1. Matrix *M*

Для матрицы *M* эти списки имеют значения: SX = {1,3,5}; SY = {1,2,6}; SM = {3,5,6}; SA = = {0,1,6}. Строки, столбцы, главные и вспомогательные диагонали, в которых имеются ошибки, отмечены значками ● в столбце SX и строках SM, SA, SY. Поиск ошибок в матрице производится с использованием значений списков ошибочных координат. Для этого на основе двух списков, выбираемых из множества {SX, SY, SM, SA}, формируется список адресов возможных ошибок в виде таблицы *S*, столбцы которой s1, s2, s3, s4 соответствуют спискам ошибочных координат SX, SY, SM, SA, значения строк в двух столбцах таблицы *S,* выбранных в качестве основы, представляют собой все пары значений из этих списков, а значения строк в остальных столбцах таблицы *S* вычисляются по формулам (1)–(6):

$$x(y,m) = (y + n - m) \bmod n, \tag{1}$$

$$x(y,a) = (2n - 1 - y - a) \bmod n, \tag{2}$$

$$y(x,m) = (x - n + m) \bmod n, \tag{3}$$

$$y(x,a) = (2n - 1 - x - a) \bmod n, \tag{4}$$

$$m(x,y) = (n - x + y\ ) \bmod n, \tag{5}$$

$$a(x,y) = (2n - 1 - x - y) \bmod n. \tag{6}$$

При вычислении координат *x, y, m, a* по формулам (1)– (6) приняты следующие соглашения: если вычисленное значение $z < 0$, то $z \bmod n = (n - |z|) \bmod n$; если $z = n$, то $z \bmod n = 0$.

Непосредственно поиск ошибочных элементов в матрице и их корректировка производятся с помощью алгоритма поиска ошибок, блок-схема которого представлена на рис. 2. Исследования характеристик алгоритма производились в соответствии со следующей методикой.

1. Выбирается произвольный массив данных размером 1 024 000 бит, который разбивается на бинарные матрицы размером ($r \times n$) бит. В каждой матрице производится вычисление четности координат матрицы: строк, столбцов, главных и вспомогательных диагоналей. Таким образом, вычисляется исходная проверочная информация (ИПИ). Полученные значения ИПИ присоединяются к исходному массиву данных.

2. С помощью датчика случайных чисел производится имитация помех при передаче информации путем внесения в матрицы ошибок в соответствии с коэффициентом интенсивности ошибок *k*.

3. Массив данных с внесенными ошибками разбивается на матрицы размером ($r \times n$) бит, в которых производится вычисление четности координат матрицы, т. е. вычисление полученной проверочной информации (ППИ).

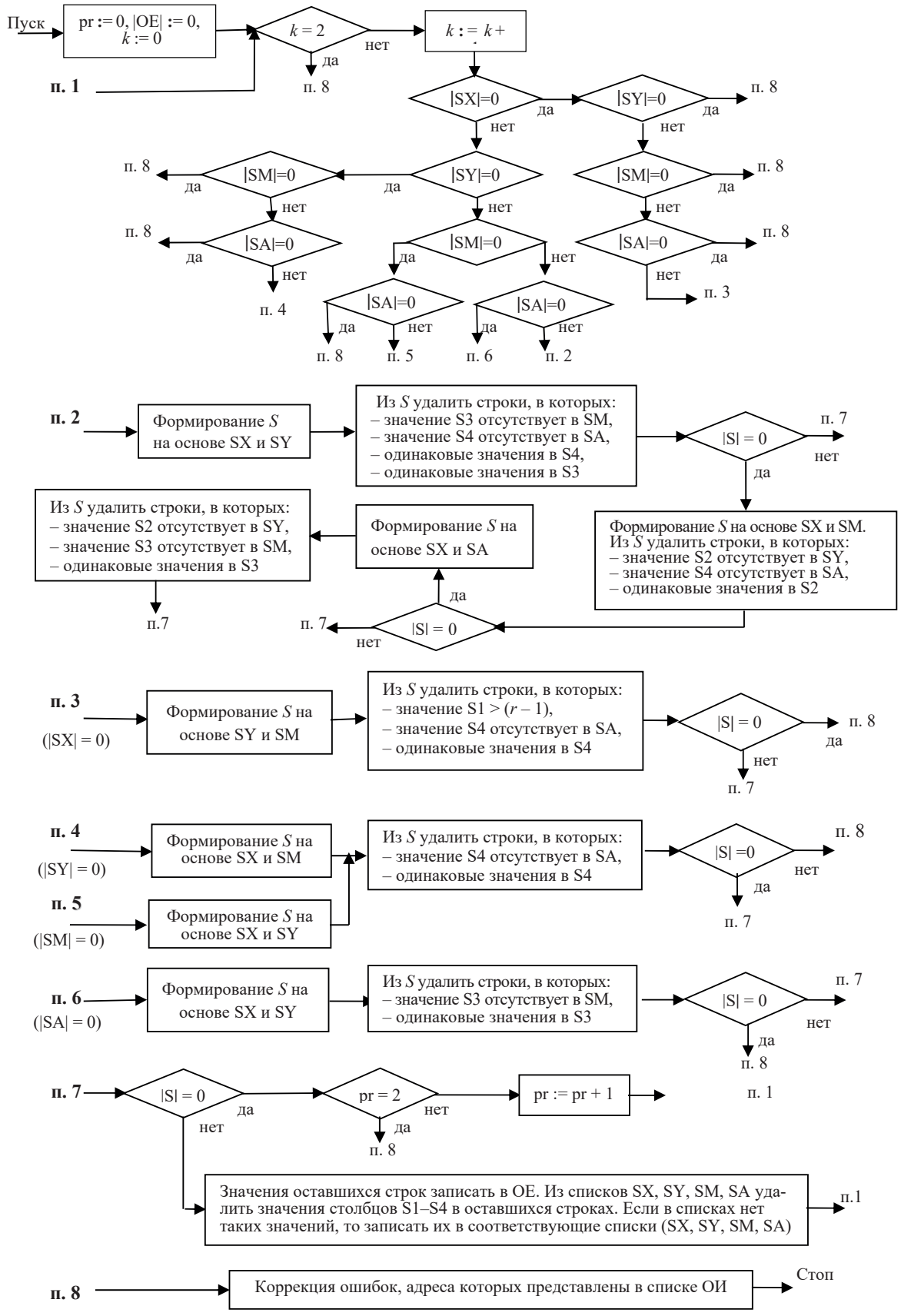
Пуск
pr := 0, |OE| := 0, k := 0
k = 2
нет
да
k := k +
п. 1
п. 8
|SX|=0
|SY|=0
|SM|=0
|SA|=0
п. 3
п. 4
п. 5
п. 6
п. 2
п. 7
п. 2
Формирование S на основе SX и SY
Из S удалить строки, в которых:
– значение S3 отсутствует в SM,
– значение S4 отсутствует в SA,
– одинаковые значения в S4,
– одинаковые значения в S3
|S| = 0
Формирование S на основе SX и SM.
Из S удалить строки, в которых:
– значение S2 отсутствует в SY,
– значение S3 отсутствует в SM,
– одинаковые значения в S2
Формирование S на основе SX и SA
Из S удалить строки, в которых:
– значение S2 отсутствует в SY,
– значение S3 отсутствует в SM,
– одинаковые значения в S3
п. 3
(|SX| = 0)
Формирование S на основе SY и SM
Из S удалить строки, в которых:
– значение S1 > (r – 1),
– значение S4 отсутствует в SA,
– одинаковые значения в S4
п. 4
(|SY| = 0)
Формирование S на основе SX и SM
Из S удалить строки, в которых:
– значение S4 отсутствует в SA,
– одинаковые значения в S4
|S| =0
п. 5
(|SM| = 0)
Формирование S на основе SX и SY
п. 6
(|SA| = 0)
Формирование S на основе SX и SY
Из S удалить строки, в которых:
– значение S3 отсутствует в SM,
– значение S3 отсутствует в SM
п. 7
pr = 2
pr := pr + 1
п. 1
Значения оставшихся строк записать в OE. Из списков SX, SY, SM, SA удалить значения столбцов S1–S4 в оставшихся строках. Если в списках нет таких значений, то записать их в соответствующие списки (SX, SY, SM, SA)
п.1
п. 8
Коррекция ошибок, адреса которых представлены в списке ОИ
Стоп

Рис. 2

Fig. 2

4. Поочередно рассматриваются все матрицы, в которых значения ППИ сравниваются с полученными значениями ИПИ и формируются списки ошибочных координат SX, SY, SM, SA.

5. С использованием значения этих списков с помощью алгоритма поиска ошибок вычисляются адреса ошибочных элементов матриц и производится их корректировка.

Алгоритм реализован на языке C#. Исследование проводилось на компьютере Intel Core 2 со следующими характеристиками: 2,17 Ггц, ОЗУ 2,0 Гбайт, ОС Windows 7.

Результаты исследований на матрицах (4 × 8) и (7 × 8) бит представлены в табл. 1, где использованы следующие обозначения: k – коэффициент интенсивности ошибок в канале связи; v – скорость кода, определяемая как отношение количества элементов в ИПИ к общему количеству передаваемых элементов; t_{alg} – время, затрачиваемое на составление списков ошибочных координат и работу алгоритма поиска ошибок; t_{cod} – время кодирования передаваемой информации, включающее время разбиения исходного массива данных на матрицы и вычисления четности координат матриц; t_{dec} – время декодирования переданной информации, равное сумме времени разбиения массива данных на матрицы, вычисления четности координат матриц, составления списков ошибочных координат и поиска ошибок.

Таблица 1

Table 1

Размер и количество матриц	Интенсивность ошибок в канале связи, k	Количество ошибок в матрице	Обнаружено ошибок, %	t_{cod}, *ms*	t_{dec}, *ms*	t_{alg}, *ms*	v
4 × 8, 32 000	10^{-3}	1	100	280	1100	820	0,53
	10^{-2}	1	100	330	910	580	
	$5 \cdot 10^{-2}$	2	100	350	1000	650	
	$8 \cdot 10^{-2}$	3	70	370	1400	1030	
7 × 8, 18 285	10^{-3}	1	100	200	600	400	0,64
	10^{-2}	1	100	200	600	400	
	$3 \cdot 10^{-2}$	2	98	200	700	500	
	$5 \cdot 10^{-2}$	3	65	200	900	700	

Обсуждение результатов. Согласно представленным в табл. 1 результатам, эффективность рассматриваемого алгоритма поиска ошибок достаточно высокая: даже при интенсивности ошибок, равной 10^{-2}, они обнаруживаются все, а время декодирования составляет менее 1 с. Но необходимо учесть, что поиск ошибок производился с использованием ИПИ, в которую, в соответствии с п. 2 методики исследований, ошибки не вносились. Поэтому характеристики эффективности алгоритма, представленные в табл. 1, являются *«идеальными»*, поскольку не учитывают влияние возмущающих факторов на проверочную информацию кода.

Фактически информационная и проверочная части кода передаются по линии связи одновременно, поэтому в одинаковой степени подвержены искажениям. Исходная проверочная информация, полученная на принимающей стороне (т. е. ППИ), может содержать ошибки, влияние которых на результаты работы алгоритма поиска ошибок неизвестно, очевидно лишь, что количество ошибок, обнаруженных с использованием ППИ (обозначим $E_{ППИ}$) и ИПИ ($E_{ИПИ}$), будет отличаться. Величина отличия зависит от возможности кода нейтрализовать влияние ошибок в проверочной части, т. е. от *устойчивости* кода, его способности противостоять внешним воздействиям, в качестве которых выступают ошибки, появляющиеся в проверочной информации при ее передаче. Обозначим устойчивость кода через U_{cod} и определим ее следующим образом:

$$U_{cod} = E_{ППИ} / E_{ИПИ}. \quad (7)$$

Аналитическим способом сформулировать характеристики устойчивости кода практически невозможно, вследствие непредсказуемости появления ошибок в проверочной части. Точные значения устойчивости кода могут быть получены только экспериментальным способом: путем вычисления значений $U_{ИПИ}$ и $U_{ППИ}$ при различных размерах матриц и различных значениях интенсивности ошибок в канале связи.

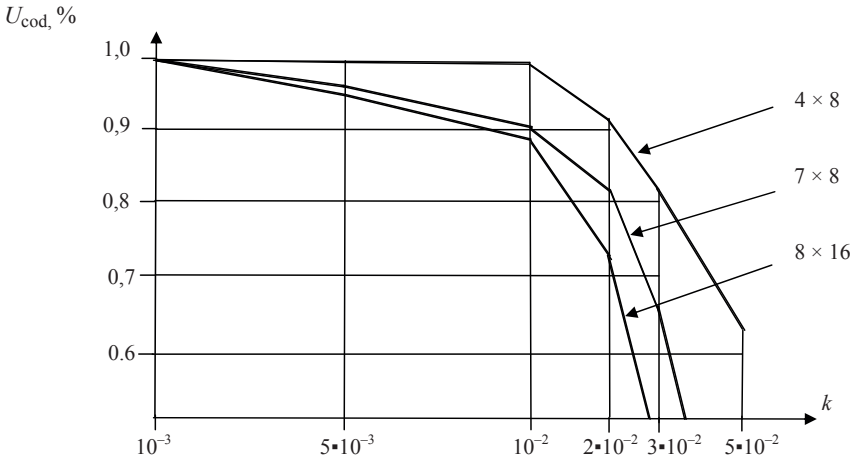

Рис. 3

Fig. 3

Что касается рассматриваемого в данной статье способа, то можно высказать некоторые предположения об устойчивости кода в определенных ситуациях. В частности, структура проверочной части кода допускает появление одной ошибки (что возможно при интенсивности ошибок 10^{-3} или 10^{-2}) в переданной проверочной информации без всяких последствий на эффективность обнаружения ошибок. Результаты экспериментов по исследованию характеристик рассматриваемого способа коррекции ошибок, включая и показатели устойчивости кода в зависимости от размеров матриц и интенсивности ошибок в канале связи, представленные в табл. 2 для матриц размером (4 × 8), (7 × 8) и (8 × 16), подтверждают сформулированные выше выводы.

Таблица 2

Table 2

Размер и количество матриц	k	Количество ошибок		Обнаружено ошибок, %		U_{cod}	Время, *ms*		
		в матрицах	в ППИ	с ИПИ ($E_{ИПИ}$)	с ППИ ($E_{ППИ}$)		t_{cod}	t_{dec}	
								с ИПИ	с ППИ
4 × 8, 32 000	10^{-3}	991	930	100	100	1,00	270	670	610
	$5 \cdot 10^{-3}$	5142	4459	100	100	1,00	260	810	880
	10^{-2}	10 197	9004	100	100	1,00	260	1050	1060
	$2 \cdot 10^{-2}$	20 272	18 129	100	91	0,91	290	1140	1350
	$3 \cdot 10^{-2}$	30 817	26 784	100	82	0,82	310	1330	1200
	$5 \cdot 10^{-2}$	51 323	44 678	100	63	0,63	310	1130	1270
7 × 8, 18 285	10^{-3}	1034	558	100	100	1,00	170	500	660
	$5 \cdot 10^{-3}$	5123	2833	100	98	0,98	170	500	640
	10^{-2}	10273	5637	100	90	0,90	180	540	480
	$2 \cdot 10^{-2}$	20 338	11 481	99	80	0,81	170	790	570
	$3 \cdot 10^{-2}$	30 560	17 168	98	65	0,66	220	690	610
8 × 16 8000	10^{-3}	1003	470	100	100	1,00	130	230	220
	$5 \cdot 10^{-3}$	5126	2235	100	96	0,96	150	250	290
	10^{-2}	10 338	4383	98	87	0,89	140	300	310
	$2 \cdot 10^{-2}$	20 506	8935	90	66	0,73	130	380	420
	$3 \cdot 10^{-2}$	30 720	13 445	76	35	0,46	130	420	490

Характер изменения устойчивости кода в зависимости от интенсивности ошибок в линии связи для нескольких размеров матриц более наглядно показан на рис. 3

Заключение. Анализ представленных результатов показывает, что оптимизированный алгоритм, разработанный на основе способа коррекции ошибок по четности координат матрицы, позволяет обнаруживать все ошибки при их интенсивности 10^{-2} и менее, что обеспечивает возможность эффективного применения предлагаемого способа для обеспечения целостности пе-

редаваемой информации в низкокачественных линиях связи. Наибольшая эффективность рассматриваемого способа достигается при разбиении массива данных на матрицы размером (4 × 8) бит. Размеры матриц, на которые производится разбиение массива данных, несущественно влияют на время, затрачиваемое на сам процесс разбиения и вычисления четности координат. Характеристики предлагаемого алгоритма, а именно быстродействие и скорость кода, обеспечивают более высокий уровень защиты информации от искажений при передаче по линиям связи благодаря эффективной коррекции ошибок.

Список использованных источников

1. Поляков, А. С. Коррекция ошибок при передаче информации по значениям четности координат бинарной матрицы / А. С. Поляков // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2017. – № 2. – С. 101–109.

2. Поляков, А. С. Эффективность способа коррекции ошибок по значениям четности координат бинарной матрицы / А. С. Поляков, И. Л. Кузнецова // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – № 3. – С. 375–382.

3. Конопелько, В. К. Табличные низкоплотные коды, исправляющие модуль и пакет ошибок / В. К. Конопелько // Автоматика и телемеханика. – 1992. – Вып. 4. – С. 155–163.

4. Шиман, Д. В. Свойства и параметры линейных итеративных кодов с двойными диагональными проверками / Д. В. Шиман, Д. М. Романенко // Тр. БГТУ. Сер. 6, Физ.-мат. науки и информатика. – 2007. – Вып. 15. – С. 151–154.

5. Романенко, Д. М. Мажоритарное декодирование двумерных линейных итеративных кодов с объединенными диагональными проверками / Д. М. Романенко, Д. В. Шиман // Тр. БГТУ. Сер. 6, Физ.-мат. науки и информатика. – 2009. – Вып. 17. – С.119–121.

6. HVD: horizontal-vertical-diagonal error detecting and correcting code to protect against with soft errors / M. Kishani [et al.] // Design Automation for Embedded Systems. – 2011. – Vol. 15, № 3/4. – P. 289–310. https://doi.org/10.1007/s10617-011-9078-2

7. Bilal, Y. A refined four-dimensional parity based EDAC and performance analysis using FPGA / Y. Bilal, S. A. Khan, Z. A. Khan // International Conference on Open Source Systems and Technologies (ICOSST 2013). Lahore, Pakistan, 16–18 December 2013. – P. 81–86.

References

1. Polyakov A. S. Error correction when transmitting information by a parity check of binary matrix coordinates. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2017, no. 2, pp. 101–109 (in Russian).

2. Poljakov A. S., Kuznetsova I. L. Efficiency of the error correction method by the parity values of binary matrix coordinates. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, no. 3, pp. 375–382 (in Russian).

3. Konopel'ko V. K. Tabular Low Density Codes Correcting Unit and Error Burst. *Avtomatika i telemekanica = Avtomation and Remote Control*, 1992, no. 4, pp. 155–163 (in Russian).

4. Shiman D. V., Romanenko D. M. Properties and parameters of linear iterative codes with double diagonal checks. *Trudy BGTU. Seriya 6, Fiziko-matematicheskie nauki i informatica = Proceedings of BSTU. Issue 6: Physics and Mathematics. Informatics*, 2007, vol. 15, pp. 151–154 (in Russian).

5. Romanenko D. M., Shiman D. V. Majority decoding of two-dimensional linear iterative codes with combined diagonal checks. *Trudy BGTU. Seriya 6, Fiziko-matematicheskie nauki i informatica = Proceedings of BSTU. Issue 6: Physics and Mathematics. Informatics*, 2009, vol. 17, pp. 119–121 (in Russian).

6. Kishani M., Zarandi H. R., Pedram H., Tajary A., Raji M., Ghavami B. HVD: horizontal-vertical-diagonal error detecting and correcting code to protect against with soft errors. *Design Automation for Embedded Systems*, 2011, vol. 15, no. 3–4, pp. 289–310. https://doi.org/10.1007/s10617-011-9078-2

7. Bilal Y., Khan S. A., Khan Z. A. A refined four-dimensional parity based EDAC and performance analysis using FPGA. *International Conference on Open Source Systems and Technologies (ICOSST 2013). Lahore, Pakistan, 16–18 December 2013*, pp. 81–86.

Информация об авторах

Кузнецова Ирина Леонидовна – главный конструктор проекта, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, г. Минск, Республика Беларусь). E-mail: kiryn@tut.by

Поляков Александр Сергеевич – кандидат технических наук, доцент, ведущий научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, г. Минск, Республика Беларусь). E-mail: alexpolja@tut.by

Information about the authors

Irina L. Kuznetsova – Chief Designer of the Project, United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: kiryn@tut.by

Alexander S. Poljakov – Ph. D. (Engineering), Associate Professor, Leading Researcher, United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: alexpolja@tut.by