УДК 681.3 doi: 10.26102/2310-6018/2018.23.4.040

Э.И. Воробьев, Ю.П. Преображенский

ИССЛЕДОВАНИЕ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ РАЗЛИЧНЫХ ФАЙЛОВ

Воронежский государственный технический университет
Воронежский институт высоких технологий

Задачи, связанные с защитой информации от помех, являются актуальными в различных практических приложениях. Сама информация может быть текстовой, графической, содержать видеофрагменты. Помехи могут быть как непреднамеренные, так и специально создаваемые злоумышленниками. Для обработки и передачи информации на практике используются различные помехоустойчивые коды. В работе рассмотрены характеристики некоторых подобных колов: код Хэмминга, код Рида-Маллера, код БЧХ Боуза-Чоудхури-Хоквингема. Приведены результаты сравнения исправляющих характеристик кодов. Дана иллюстрация закодированного текста и тех бит, которые подверглись искажениям. Приведена графическая зависимость, иллюстрирующая зависимость для трех выбранных кодов числа исправленных ошибок от числа ошибок, которые были добавлены в исходное сообщение. Приведены результаты исследований характеристик избыточности кодов с применением разных файлов. Для тестирования были выбраны коды БЧХ и Рида-Маллера. Анализировались файлы формата txt, midi, wmv и mpeg-2. Составлена таблица по результатам исследований, в которой показано, каким образом размер файла, а также его вид, оказывают влияние на избыточность. Установлено, что код Хэмминга и Рида-Маллера хорошо исправляют одиночные ошибки, код БЧХ подходит для исправления разных ошибок.

Ключевые слова: кодирование, информация, исправляющая способность, ошибка.

Введение. В настоящее время существенно увеличились объёмы, как передаваемой информации, так и принимаемой. Поэтому можно говорить об актуальности проблемы, связанной с сохранением целостности информации, подлежащей обработки и передаче [1].

Например, в памяти компьютера [2] могут возникать ошибки вследствие того, что есть всплески напряжения, относящегося к линиям электропередач, а также могут быть и другие причины. При появлении разных ошибок, чтобы обеспечить борьбу с ними исследователями были предложены разные подходы по кодированию информации. На их базе можно провести обнаружение и исправлении возможных ошибок. Помехоустойчивое кодирование может характеризоваться разными видами.

Для некоторых из них требуется привлекать достаточно сложный математический аппарат, а встречаются и такие, которые являются достаточно простыми. Также есть различие в методах кодирования с точки

зрения и эффективности. Иногда излишняя математизированность материалов, а также недостаточная наглядность ведут к тому, что изучать методы кодирования довольно сложно. В тех случаях, когда в сообщениях есть внутренние корреляционные связи, значение помехоустойчивости в любом коде можно повысить вследствие того, что есть статистические связи между сообщениями. Корреляция говорит о том, что есть зависимость сообщений друг от друга.

В тех случаях, когда связи являются слабыми или неизвестными, их нельзя применять с тем, чтобы повышать помехоустойчивость [3, 4], когда необходимо обеспечивать избыточную форму сообщения. Для решения этой проблемы, например, осуществляют увеличения числа символов по коду сообщения. При этом искусственным образом вводятся корреляционные связи между кодовыми символами. Тогда о помехоустойчивых кодах говорят, как об избыточных. За счёт того, что обеспечивается избыточность в коде, можно достичь для линий связи увеличения энергетической эффективности [5].

При этом по передаваемому сигналу идет обнаружение и исправление ошибок. Частный спектр сужается. Это может быть полезно при проектировании систем связи [6, 7].

При этом необходимо затратить меньшее время на установление связи, лучше характеристики помехозащищенности. Идёт улучшение по корреляционным связям в ансамбле сигналов, за счёт простых средств реализуется разнесённый приём.

В данной работе проведено исследование исправляющей способности некоторых кодов [8] при разных входных данных.

Характеристики кода Хэмминга. Когда идет преобразование одного слова в другое при интервале Хэмминга равном d, то достаточно, чтобы было d битовых ошибок.

Если слово, которое содержит m бит имеет добавление g бит четности, в итоге получается слово, имеющее в длине m+g бит.

Осуществляется нумерация бит, начиная с единицы, эту нумерацию делают слева направо. Биты чётности связаны с номерами, являющимся степенями двоек. Другие биты применяют для данных.

Ещё есть контрольные разряды (дополнительные биты) – их g. Будем считать, что значение общей длины слова равняется n. Кодированное слово содержит в себе m битов, относящихся к данным, есть n-битная единица, а также g контрольных разрядов. По любым двум кодированным словам, например,10101 и 11010 можно сделать вывод о том, какое число бит может быть различными.

В указанном примере мы можем указать, что различных бит-4. Для определения числа различных бит необходимо по двум кодированным словам осуществить логическую операцию, связанную с исключающим или, и провести подсчёт количества бит, которые характеризуются значением 1 для полученного результата. Интервал указания числа битовых позиций, относительно которых наблюдается различие в двух словах, определяет интервал Хэмминга. Необходимо t ошибок в битах, при равенстве такого интервала t для того, чтобы одно из слов преобразовать в другое.

Характеристики кода Рида-Маллера. Для кодов Рида-Маллера характерно то, что их относят к классу линейных кодов. Преобразование осуществляется над $GF(2)$ при простом описании и декодировании. Они проводятся на основе метода простого голосования.

В этой связи можно говорить о том, что коды Рида-Маллера характеризуются важной ролью в кодировании. Для любых целых m и $r < m$ существует код Рида-Маллера длины 2^m, который называется кодом Рида-Маллера r-го порядка длины 2^m.

Особенности кодов БЧХ (Боуза-Чоудхури-Хоквингема).

Коды Боуза-Чоудхури-Хоквингема (БЧХ) представляют собой весьма большой класс кодов [9]. С их помощью могут быть исправлены несколько ошибок.

Порождающий многочлен циклического кода можно представить в виде

$$g(x) = HOK[f_1(x), f_2(x), \ldots, f_r(x)], \quad (1)$$

где $f_1(x), f_2(x), \ldots$ – минимальные многочлены корней $g(x)$.

Коды могут задаваться при помощи порождающих многочленов, которые задаются за счет своих корней.

Если $c(x)$ – является кодовым многочленом, а $e(x)$ – является многочленом ошибок, тогда соответствующий многочлен при коэффициентах из $GF(q)$ представляется как

$$v(x) = c(x) + e(x). \quad (2)$$

Результаты сравнения исправляющих характеристик кодов. Параметры по кодам были определены такие:

в коде **Хемминга** – 9;4 (5 бит информационных, 3 бита контрольных);

в коде **БЧХ** – значение длины информационного слова – 20, значение длины кодового слова – 30, число исправляемых ошибок – 4;

в коде **Рида-Маллера** – значение длины информационного слова – 20, значение длины кодового слова – 32, значение минимального расстояния кода, число ошибок, которые будут исправляться – 4, значение порядка кода Рида-Маллера – 3, определяет длину в кодовом слове – 6. По

результатам исследований, связанных с исправляющими способностями соответствующих кодов проведено обобщение в Таблицах 1, 2, 3. В целях испытаний рассматривались файлы, которые содержали некоторый текст, приведенный внутри данных таблиц.

Таблица 1-Характеристики кода Хемминга

	Текст на входе	Ошибки, которые были добавлены	Ошибки, которые были исправлены	Число битов, которые подверглись искажениям	% битов, подвергнувшихся искажению
1	Cmr	3	3	0	0
2	22	3	0	3(3)	13
3	Cmr	4	0	3(2)	13
4	Cmr	5	2	3(1)	13
5	Cmr	5	1	4(7)	13
6	Cmr	6	6	0	0
7	Cmr	6	2	4(1)	13

Таблица 2 -Характеристики кода БЧХ

	Текст на входе	Ошибки, которые были добавлены	Ошибки, которые были исправлены	Число битов, которые подверглись искажениям	% битов, подвергнувшихся искажению
1	Cmr	3	3	0	0
2	22	3	3	0	0
3	Cmr	4	4	0	0
4	Cmr	5	0	5	17
5	Cmr	5	5	0	0
6	Cmr	6	6	0	0
7	Cmr	6	0	12	46

Таблица 3 -Характеристики кода Рида-Маллера

	Текст на входе	Ошибки, которые были добавлены	Ошибки, которые были исправлены	Число битов, которые подверглись искажениям	% битов, подвергнувшихся искажению
1	Cmr	3	3	0	0
2	22	3	3	0	0
3	Cmr	4	4	0	0
4	Cmr	5	5	3	9
5	Cmr	4	5	0	0
6	Cmr	6	4	6	21
7	Cmr	6	5	3	9

В Таблице 4 можно увидеть текст, который представлен закодированным образом, чтобы продемонстрировать то, каким образом происходил подбор ошибок. В качестве примера, мы взяли код Хемминга.

Это связано с тем, что значение его длины характеризуется наименьшим размером, если проводить сравнение с другими кодами. Те биты, которые подверглись искажению, обозначены при помощи подчеркивания.

Таблица 4 - Иллюстрация закодированного текста

1	11001101101001110010010101011001001000011
2	10000111101001101110001111
3	11001101101001111010001010110011101000011
4	11001101101001111100001010110011101000011
5	11111101101001110011001010110011011110011
6	11001001101011110010001010011001101010011
7	11001101101001111100101010110011101000011

Мы можем наблюдать в Таблице 4, что были рассмотрены разные ошибки: это относится к одиночным, а также последовательным множественным. Для случаев 1, 2, 3, 4, 5 и 7 говорим об одиночных ошибках, для 6 – о множественных.

Рисунок 1 иллюстрирует сравнительный анализ по графикам 3 кодов. Использованы обозначения: x – это количество добавленных ошибок, y – является числом ошибок, которые были исправлены.

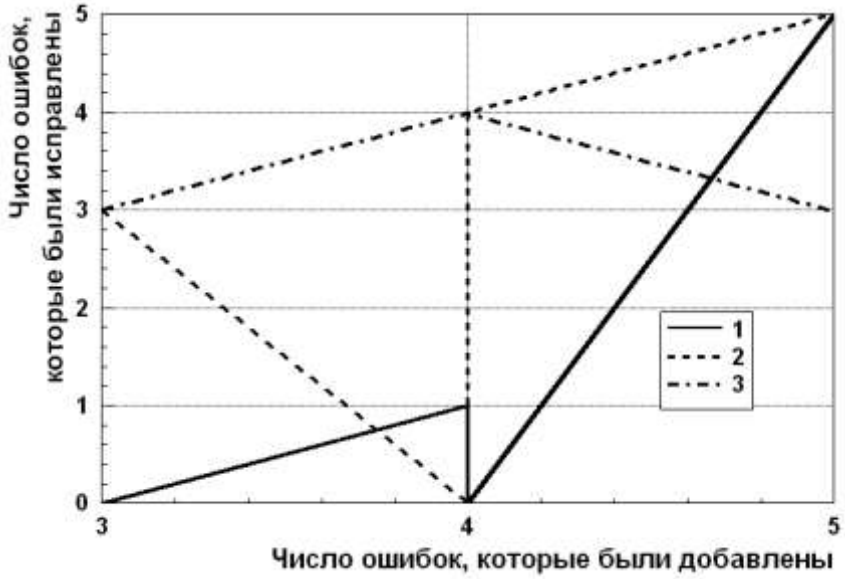

Рисунок 1. – Сравнительный анализ исправляющей способности 3 кодов
1- код Хемминга, 2- код БЧХ, 3 – Код Рида-Маллера

Результаты исследований характеристик избыточности кодов с применением разных файлов.

Для тестирования были выбраны коды БЧХ и Рида-Маллера. Исследовались файлы форматов txt, midi, wmv и mpeg-2.

Мы можем наблюдать в Таблице 5 результаты по проведенным исследованиям, каким образом размер файла, а также вид оказывают влияние на значение избыточности.

Таблица 5-Исследование звукового файла, с форматом midi (.mid)

Код	Значение размера файла (Число бит)	Значение процента избыточности
По исходному файлу	22255	100
Для кода БЧХ	44743	63
Для кода Рида-Маллера	33599	49
Для текстового файла txt		
По исходному файлу	21199	100
Для кода БЧХ	419583	64
Для кода Рида-Маллера	317471	49
Для видеофайла, с форматом wmv		
По исходному файлу	256879	100
Для кода БЧХ	548615	64
Для кода Рида-Маллера	291736	48
Для видеопотока MPEG-2 (При широкоэкранном Program Stream PAL)		
По исходному файлу	569863	100
Для кода БЧХ	2487591	64
Для кода Рида-Маллера	928799	48

Анализ избыточности кодов, если анализируется видеопоток MPEG-2, можно увидеть на Рисунке 2.

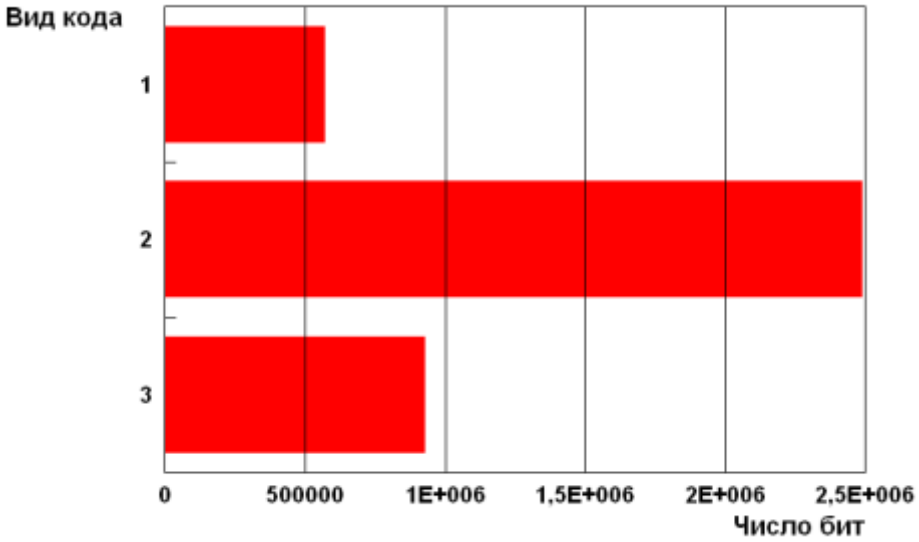

Рисунок 2 – Размер исследуемого файла
1- размер исходного файла, 2- с применением кода БЧХ, 3 – с применением кода Рида-Маллера

Выводы. После проведенных исследований можно сделать такие выводы:

1. За счет кода Хемминга очень хорошо идет исправление одиночных ошибок. Если возникают множественные ошибки, то возможности исправляющей способности по коду резким образом уменьшаются. Также, за счет кода Хемминга будет происходить искажение остальных бит, которые не были связаны с искажением.
2. За счет кода БЧХ хорошо исправляются различные ошибок, проблемы могут появиться только, если есть множественные ошибки, и если они имеют общее число более чем 3. При этом мы можем наблюдать и заметное искажение информации.
3. За счет кода Рида-Маллера очень хорошо идет исправление одиночных ошибок (причем искажений меньше, чем для кода БЧХ), а также множественных - не более 3.
4. Значение избыточности по коду при исследовании видеофайла растет прямо пропорционально с размером исходных файлов. Это справедливо для обоих кодов.
5. Рост размеров файла порядка 45 раз (при сравнении mid и mpeg-2 файлов) обуславливает рост в избыточности для кода БЧХ на 0,7% и на 0,95%, когда применяется код РМ.
6. За счет кода БЧХ очень хорошо исправляются одиночные ошибки по формату MPEG-2, причем неважно, где находится искажённый бит. Но за счет него плохо идет исправление множественных ошибок. Обратная картина справедлива для кода Рида-Маллера.

Множественные ошибки он обрабатывает очень хорошо, а вот при одиночных значение эффективности заметным образом уменьшается.

ЛИТЕРАТУРА

1. Головинов С.О. Проблемы управления системами мобильной связи / С.О. Головинов, А.А. Хромых // Вестник Воронежского института высоких технологий. 2012. № 9. С. 13-14.
2. Шапаев А.В. Распространение радиоволн в городских условиях / А.В. Шапаев, О.Ю. Клишина // Вестник Воронежского института высоких технологий. 2018. № 3(26). С. 14-18.
3. Шапаев А.В. Оценка степени затухания сигналов мобильной связи в городских условиях А.В. Шапаев, О.Ю. Клишина // Вестник Воронежского института высоких технологий. 2018. № 3(26). С. 19-23.
4. Львович И.Я. Основы информатики / И.Я.Львович, Ю.П.Преображенский, В.В.Ермолова // Воронеж, Издательство: Воронежский институт высоких технологий (Воронеж), 2014, 339 с.
5. Кульнева Е.Ю. О характеристиках, влияющих на моделирование радиотехнических устройств / Е.Ю. Кульнева, И.А. Гащенко // Современные наукоемкие технологии. 2014. № 5-2. С. 50.
6. Мишин Я.А. О системах автоматизированного проектирования в беспроводных сетях / Я.А.Мишин // Вестник Воронежского института высоких технологий. 2013. № 10. С. 153-156.
7. Воронов А.А. Обеспечение системы управления рисками при возникновении угроз информационной безопасности / А.А. Воронов, И.Я. Львович, Ю.П. Преображенский, В.А. Воронов // Информация и безопасность. 2006. Т. 9. № 2. С. 8-11.
8. Блейхуд Р. Теория и практика кодов, контролирующих ошибки / Р.Блейхуд // Пер. с англ. - М.: Мир, 1986. - 576 с.
9. Жулябин Д.Ю. Особенности практической реализации помехоустойчивых кодов / Д.Ю.Жулябин, С.Е.Малявин, Д.А. Скроготунов // Вестник Воронежского института высоких технологий. 2018. № 3(26). С. 54-57.

E. I. Vorobyev, Y. P. Preobrazhenskiy

THE INVESTIGATION OF ERROR-CORRECTING CODING OF THE VARIOUS FILES

Voronezh State Technical University
Voronezh Institute of High Technologies

Problems related to the protection of information from interference are relevant in a variety of practical applications. The information itself can be text, graphic, contain video clips. Interference can be unintentional as well as specially created by cyber criminals. To process and transmit information in practice, various noise-resistant codes are used. The paper discusses the characteristics of some similar stakes: Hamming code, code, reed-Muller, BCH code Bose-Chowdhury-Hoquinghem. The results of comparison of characteristics-correcting codes are shown. An illustration of the encoded text and those bits that have been distorted. A graphical dependency is given to illustrate the dependence of the number of corrected errors on the number of errors that were added to the original message for the three selected codes. The results of studies of the characteristics of redundancy of codes with different files are shown. The BCH and reed-Muller codes were chosen for testing. Txt, midi, wmv and mpeg-2 files were analyzed. A table on the results of research, which shows how the file size, as well as its appearance, have an impact on redundancy. It is established that Hamming and reed-Muller code fix single errors well, BCH code is suitable for correcting various errors.

Keywords: coding, information, correcting ability, error.

REFERENCES

1. Головинов С.О. Проблемы управления системами мобильной связи / С.О. Головинов, А.А. Хромых // Вестник Воронежского института высоких технологий. 2012. No. 9. pp. 13-14.
2. Шапаев А.В. Распространение радиоволн в городских условиях / А.В. Шапаев, О.Ю. Клишина // Вестник Воронежского института высоких технологий. 2018. No. 3(26). pp. 14-18.
3. Шапаев А.В. Оценка степени затухания сигналов мобильной связи в городских условиях А.В. Шапаев, О.Ю. Клишина // Вестник Воронежского института высоких технологий. 2018. No. 3(26). pp. 19-23.
4. Львович И.Я. Основы информатики / И.Я.Львович, Ю.П.Преображенский, В.В.Ермолова // Воронеж, Издательство: Воронежский институт высоких технологий (Воронеж), 2014, 339 p.
5. Кульнева Е.Ю. О характеристиках, влияющих на моделирование радиотехнических устройств / Е.Ю. Кульнева, И.А. Гащенко // Современные наукоемкие технологии. 2014. No. 5-2. pp. 50.

6. Мишин Я.А. О системах автоматизированного проектирования в беспроводных сетях / Я.А.Мишин // Вестник Воронежского института высоких технологий. 2013. No. 10. pp. 153-156.
7. Воронов А.А. Обеспечение системы управления рисками при возникновении угроз информационной безопасности / А.А. Воронов, И.Я. Львович, Ю.П. Преображенский, В.А. Воронов // Информация и безопасность. 2006. Vol. 9. No. 2. pp. 8-11.
8. Блейхуд Р. Теория и практика кодов, контролирующих ошибки / Р.Блейхуд // Пер. с англ. - М.: Мир, 1986. - 576 р.
9. Жулябин Д.Ю. Особенности практической реализации помехоустойчивых кодов / Д.Ю.Жулябин, С.Е.Малявин, Д.А. Скроготунов // Вестник Воронежского института высоких технологий. 2018. No. 3(26). pp. 54-57.