УДК 519.725 DOI: 10.14529/ctcr170318

О НЕКОТОРЫХ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТАХ НАД СПИСОЧНЫМ ДЕКОДЕРОМ

В.Д. Кряквин, К.В. Крыжановский

Южный федеральный университет, г. Ростов-на-Дону

Исследована принципиальная возможность осуществления успешного декодирования сообщений, количество ошибок в которых априори превосходит исправляющую способность пары (RS-код, GS-декодер), для некоторых типов кодов Рида – Соломона. Для проверки гипотезы о существовании такой возможности была построена модель и разработан специальный алгоритм, основанный на обработке стираний, проведены вычислительные эксперименты. Также было проанализировано изменение средней мощности выходного списка GS-декодера при использовании упомянутого алгоритма. Установлено, что с использованием разработанного алгоритма частота успешных декодирований возрастает, при этом статистически значимого изменения средней мощности выходного списка не наблюдается.

Ключевые слова: алгебраическое кодирование, RS-код, GS-декодер, исправляющая способность, декодирование.

Введение

Обмен информацией всегда задействует некоторый канал, обеспечивающий связь между отправителем и получателем. Зачастую канал находится в изменяющейся среде, которая неизбежно влияет тем или иным образом на его свойства, тем самым внося в сигнал, идущий по этому каналу, искажения и ошибки. Перед непосредственной отправкой сообщения по каналу связи отправляющая сторона применяет кодер и при необходимости модулятор. Принимающая же сторона пропускает при необходимости полученный сигнал через демодулятор и затем направляет его в декодер для восстановления и дальнейшего прочтения сообщения. При этом максимальное число ошибок, которое может исправить конкретная связка кодер-декодер, называется исправляющей способностью. На сегодняшний день широчайшее распространение получили коды из семейства кодов Рида – Соломона (РС-коды, RS-коды). РС-коды нашли свое применение во многих областях, вот лишь некоторые из них: защита данных на носителях информации, спутниковая связь (в частности, например, в сетях WiMAX), передача данных в оптических каналах связи. Для РС-кодов существует несколько типов декодеров (методов декодирования), из которых стоит особо выделить так называемый списочный декодер. Его коренное отличие от всех остальных декодеров заключается в том, что результатом его декодирования является не единственное сообщение, а список сообщений, в котором при выполнении определенных условий гарантированно содержится верное сообщение. Это фундаментальное отличие позволяет списочному декодеру исправлять большее количество ошибок, чем другие декодеры. Первый полиномиальный алгоритм, решающий задачу списочного декодирования для РС-кодов, был создан в 1997 году М. Суданом [1], а позднее модифицирован В. Гурусвами и М. Суданом [2] с целью снятия ограничений по эффективному применению, связанных с некоторыми характеристиками исходных РС-кодов. Этот алгоритм впоследствии стал известен как списочный декодер Гурусвами – Судана (GS-декодер). Далее слово «списочный» будет иногда опускаться.

В настоящей работе проверена гипотеза о наличии возможности успешного декодирования сообщений, количество ошибок в которых заведомо превышает исправляющую способность связки (RS-код, GS-декодер) для некоторых типов РС-кодов. Также проверялась возможность уменьшения объема выходного списка GS-декодера. Исследование проводилось при помощи специального алгоритма обработки стираний, который был разработан и реализован в виде отдельной надстройки к декодеру Гурусвами – Судана. Стиранием считается такая ошибка, позиция которой в сообщении априори известна, но чаще всего ошибки носят случайный характер, поэтому мы не можем знать позицию ошибки заранее, но можем попытаться «угадать» ее и, таким

образом, впоследствии получить возможный прирост исправляющей способности. В частности, сравнивались результаты декодирования, которые были получены с использованием упомянутого выше алгоритма и без него.

Основным достоинством используемой схемы, в которую входит разработанный алгоритм, по сравнению с классической является возможность осуществления декодирования сообщений, количество ошибок в которых превышает исправляющую способность исходной связки кодер-декодер, не прибегая при этом к смене кодера.

Математическая постановка задачи

Приведем описание процедуры кодирования, изложенной в [3]. Пусть $\mathbb{F}_q$ – конечное поле, $q = p^s$, где p – простое число, s – натуральное. Выберем длину кода n и размерность кода k, такие, что $k \le n \le q$. Из $\mathbb{F}_q$ выберем некоторые $\alpha_0 \dots \alpha_{n-1}$, такие, что: 1) $\alpha_i \ne \alpha_j$, $i \ne j$; 2) $\alpha_i \ne (0)\ \forall\ i$, где (0) – нуль в $\mathbb{F}_q$. В дальнейшем будем называть эти $\alpha_0 \dots \alpha_{n-1}$ базой кода. Сообщение – вектор $m = (m_0, \dots, m_{k-1})$, $m_i \in \mathbb{F}_q$; после составления сообщения формируется информационный полином $m(x)$: $m(x) = m_0 + m_1x + m_2x^2 + \dots + m_{k-1}x^{k-1}$. Процесс кодирования сообщения m заключается в вычислении кодового вектора $c = (c_0, \dots, c_{n-1})$ по правилу $c_i = m(\alpha_i)$. Кодовый вектор передается по каналу связи, в котором к нему добавляется квазислучайная аддитивная ошибка $e = (e_0, \dots, e_{n-1})$, в результате из канала приходит вектор $v = (v_0, \dots, v_{n-1})$, где $v \in F_q^n$, и $v_i = c_i + e_i$, $i = 0, \dots, n-1$. Декодирование, согласно [3], осуществляется следующим образом: пусть $\mathbb{F}_q$ – поле Галуа мощности q, $\mathbb{F}_q[x]$ – кольцо полиномов переменной x над полем $\mathbb{F}_q$, $\mathbb{F}_q[x, y]$ – кольцо полиномов двух переменных x и y над полем $\mathbb{F}_q$, $(\alpha_0, \dots, \alpha_{n-1})$ – база кода, $(v_0, \dots, v_{n-1})$ – пришедшее из канала слово. На первом шаге происходит построение многочлена от двух переменных $Q(x, y) \in \mathbb{F}_q[x, y]$, который должен удовлетворять следующим приведенным в [3] условиям: 1) $Q(\alpha_i, v_i) = 0,\ \forall\ i = 0,\ \dots,\ n-1$; 2) $Q(x, y) \ne 0$. На втором шаге производится факторизация полинома $Q(x, y)$ на множители вида: $y - m(x)$, $\deg\big(m(x)\big) < k$ и $|\{i : m(\alpha_i) = v_i\}| \ge t$ (или $|\{i : m(\alpha_i) \ne v_i\}| \le t_m$), где, согласно [4], $t_m = n - t$ – управляющий параметр, задающий радиус локации декодера. Множество таких $m(x)$ формирует выходной список декодера. В [1] М. Судан предложил искать такой полином $Q(x, y)$ в следующем виде:

$$Q(x, y) = \sum_{j=0}^{l} \sum_{k=0}^{m+(l-j)d} x^k\, y^j, \tag{1}$$

где $d = k - 1$, $l = \left\lfloor \sqrt{2(n+1)/d} \right\rfloor$, $m = \lceil d/2 \rceil - 1$. Подставив в (1) элементы базы и компоненты пришедшего по каналу слова и учтя требования, накладываемые на полином $Q(x, y)$ выше, получим СЛАУ относительно q_{kj} – коэффициентов полинома $Q(x, y)$:

$$\sum_{j=0}^{l} \sum_{k=0}^{m+(l-j)d} {\alpha_i}^k\, {v_i}^j,\ i = 0,\ \dots,\ n-1. \tag{2}$$

В [4] приведено значение верхней границы исправляющей способности списочного декодера Гурусвами–Судана. Согласно ей GS-декодер способен исправлять вплоть до t_{GS} ошибок, где $t_{GS} = n - 1 - \left\lfloor \sqrt{(k-1)n} \right\rfloor$. В рамках приведенных исходных данных была поставлена задача анализа возможности осуществления успешного декодирования пришедшего из канала сообщения, количество ошибок в котором превышало бы фактическую исправляющую способность для исследуемых классов РС-кодов, а также сокращения мощности выходного списка GS-декодера.

Метод

В рамках исследования упор был сделан на методику обработки стираний. Если $v = (v_0, \dots, v_{n-1})$ – пришедшее по каналу слово, и мы точно знаем, что в v_1, например, находится неверное значение, то вектор v преобразуется в вектор $v' = (v_0, v_2, \dots, v_{n-1})$. Так, если стираний произошло τ штук, то из v удаляются τ элементов, содержащие эти стирания. При декодировании подобное выкалывание элементов означает для нас укорочение исходного РС-кода. Укорочение или выкалывание – операция, состоящая в удалении одного проверочного символа α_i (элемента базы кода). Длина кода n уменьшается на единицу, размерность k остается неизменной.

Соответственно, кодовое расстояние $d = n - k + 1$, так же как и длина, уменьшается на единицу. Таким образом, если в пришедшем из канала слове $v = (v_0, \dots, v_{n-1})$ мы детектировали τ стираний в некоторых элементах $v_{j_0}, v_{j_1}, \dots, v_{j_{\tau-1}}$, то для исходного кода мы удаляем $\alpha_{j_0}, \alpha_{j_1}, \dots, \alpha_{j_{\tau-1}}$ из базы и учитываем это при декодировании. Также стоит упомянуть о том, что укороченный подобным образом РС-код так же является РС-кодом. Итак, система (2) трансформируется согласно принципу, изложенному выше:

$$\sum_{j=0}^{l'} \sum_{k=0}^{m+(l'-j)d} (\alpha_i')^k \, (v_i')^j, \; i = 0, \; \dots, \; n-1-\tau \, ,$$

где: τ – число стираний; $d = k - 1$; $l' = \left\lceil \sqrt{(2(n+1-\tau))/d} \right\rceil$; $m = \lceil d/2 \rceil - 1$; $v' = (v_0', \dots, v_{n-1-\tau}')$ – вектор, полученный из вектора $v = (v_0, \dots, v_{n-1})$ выкалыванием элементов, которые содержат стирания; $\alpha' = (\alpha_0', \dots, \alpha_{n-1-\tau}')$ – новая база, полученная из $\alpha = (\alpha_0, \dots, \alpha_{n-1})$ выкалыванием соответствующих элементов. Теперь с помощью описанного выше метода продемонстрируем, как можно добиться декодирования за границей исправляющей возможности декодера, предварительно введя следующее основное предположение. Предположим, что существуют два РС-кода: RS-1 длины n и размерности k, RS-2 длины $n-1$ и размерности k, такие, что: 1) RS-2 получен из RS-1 укорочением на один проверочный символ; 2) t_{GS}[RS-2] $\geq t_{GS}$[RS-1], т. е. $n - 2 - \left\lfloor \sqrt{(k-1)(n-1)} \right\rfloor \geq n - 1 - \left\lfloor \sqrt{(k-1)n} \right\rfloor$. Пусть выполнено основное предположение. Пусть исправляющая способность равна в точности t_{GS}[RS-1]. Закодируем кодом RS-1 некоторое сообщение и внесем аддитивную ошибку e веса $w(e) = t_{GS}$[RS-1] + 1, таким образом, в векторе $v = (v_0, v_1, ..., v_{n-1})$ содержится t_{GS}[RS-1] + 1 ошибок. Пусть мы узнали или угадали хотя бы одну позицию ошибочного элемента в векторе v, тогда мы можем произвести процедуру укорочения исходного кода и получить код RS-2. Стало быть, после проведения такой операции у нас остался вектор v', полученный из v выкалыванием одного неверного элемента и содержащий теперь t_{GS}[RS-1] ошибок, и код RS-2, способный в силу второй части основного предположения восстановить исходное сообщение при условии, что вес ошибки не превосходит t_{GS}[RS-2]. Таким образом, в силу того, что t_{GS}[RS-2] $\geq t_{GS}$[RS-1], код RS-2 восстановит невосстановимое для RS-1 сообщение. Следовательно, используя этот принцип для кода RS-1, мы осуществляем декодирование за границей его исправляющей способности. Покажем, что основное предположение выполняется, т. е. существуют РС-коды, удовлетворяющие его условиям. Эта задача сводится к вопросу существования таких параметров n и k, при которых выполняется условие

$$n - 2 - \left\lfloor \sqrt{(k-1)(n-1)} \right\rfloor \geq n - 1 - \left\lfloor \sqrt{(k-1)n} \right\rfloor. \qquad (3)$$

Легко убедиться, что RS(64,5) и RS(63,5) удовлетворяют этим условиям. Для них t_{GS}[RS(63,5)] = = t_{GS}[RS(64,5)] = 47. Также этим условиям удовлетворяют коды (RS(64,10), RS(63,10)), …, (RS(128,19), RS(127,19)), …, (RS(256,32), RS(255,32)) и другие. Таким образом, факт существования таких кодов доказан.

Теперь приведем детальное описание численных экспериментов и разработанного для них алгоритма обработки стираний. Число t_{GS} для каждого РС-кода является верхней границей его исправляющей способности при применении GS-декодера. Точное максимальное количество гарантированно исправляемых ошибок для каждого РС-кода можно установить лишь экспериментально. Проверим гипотезу о том, что, используя описанный выше метод, порой можно добиться верного декодирования за границей корректирующей способности произвольного кода, для которого выполнено условие (3), проведя следующие численные эксперименты.

Численный эксперимент № 1: пусть есть некоторый РС-код RS-1, длины n и размерности k, известна его исправляющая способность $t_{RS\text{-}1}$, причем $t_0 \leq t_{RS\text{-}1} \leq t_{GS}$[RS-1]. Кодом RS-1 закодировано некоторое сообщение m. В канале на вектор c наложилась аддитивная ошибка e, такая, что $w(e) = t_{RS\text{-}1} + 1$. Теперь попытаемся произвести декодирование двумя способами. Способ первый: подадим пришедшее по каналу искаженное кодовое слово в GS-декодер для кода RS-1 и получим на выходе некоторый результат-список. Назовем его NC_LIST. Способ второй: проведем n процедур укорочения для исходного кода RS-1 с глубиной укорочения $\tau = 1$, каждый раз «вычеркивая» один элемент из вектора v (вычеркиваемые элементы выделены) и один элемент из базы α:

$$(\boxed{v_0}, v_1, \dots, v_{n-1}) \to v'_0 = (v_1, \dots, v_{n-1}), \alpha' = (\alpha_1, \dots, \alpha_{n-1});$$
$$(v_0, \boxed{v_1}, \dots, v_{n-1}) \to v'_1 = (v_0, v_2, \dots, v_{n-1}), \alpha' = (\alpha_0, \alpha_2, \dots, \alpha_{n-1});$$
$$\vdots$$
$$(v_0, v_1, \dots, \boxed{v_{n-1}}) \to v'_{n-1} = (v_0, \dots, v_{n-2}), \alpha' = (\alpha_0, \dots, \alpha_{n-2}).$$

К каждому v'_i будем применять GS-декодер согласно описанному ранее методу и получим множество списков, назовем его $\{\text{C\_LIST}_i\}$, $i = 0, \dots, n-1$. Теперь построим результирующий список, именуемый *UN_LIST*, который будет представлять собой объединение списков C_LIST_i: $\text{UN\_LIST} = \bigcup_{i=0}^{n-1} \text{C\_LIST}_i$ (под объединением списков здесь и далее понимается объединение множеств *y*-корней, которые их составляют). Сравним списки UN_LIST и NC_LIST, и тем самым оценим эффективность обоих способов. Результаты такого опыта описаны в разделе «Результаты».

Численный эксперимент № 2: пусть у нас есть некоторый РС-код RS-1, длины n и размерности k, для него известна его исправляющая способность $t_{\text{RS-1}}$, причем $t_0 \le t_{\text{RS-1}} \le t_{\text{GS}}$[RS-1]. Кодом RS-1 было закодировано некоторое сообщение m. В канале на кодовое слово c наложилась аддитивная ошибка e, такая, что $w(e) = t_{\text{RS-1}}$, т. е. код RS-1 в сочетании с GS-декодером может ее исправить. Для кода RS-1 была проделана процедура, описанная в численном эксперименте № 1, и сформированы списки NC_LIST и UN_LIST. Теперь построим список INT_LIST как пересечение списков NC_LIST и UN_LIST: $\text{INT\_LIST} = \text{NC\_LIST} \cap \text{UN\_ LIST}$ (под пересечением списков здесь и далее понимается пересечение множеств *y*-корней, которые их составляют). Сравним объемы списков INT_LIST и NC_LIST, в частности, выясним, можно ли получить таким способом список, количество элементов которого меньше, чем в исходном списке NC_LIST, полученном для неукороченного кода. Результаты численного эксперимента № 2 представлены в разделе «Результаты».

Результаты

В результате численных экспериментов было выявлено, что использование специального алгоритма обработки стираний приводит к повышению частоты успешных декодирований (48 % успешных декодирований против менее 1 % для исследуемого кода) за границей исправляющей способности пары (RS-код, GS-декодер). В частности, для кода RS(128, 19) эта зависимость проиллюстрирована на рис. 1 и 2 соответственно (верхние перекрестия на графиках – успешное декодирование, нижние перекрестия – сообщение не было восстановлено). Вес вносимой ошибки – 65 при $t_{\text{RS}(128,19)} = 64$; $t_{\text{RS}(128,19)}$ – максимальное количество ошибок, которое может гарантированно исправить исследуемый код при применении GS-декодера, значение получено экспериментальным путем.

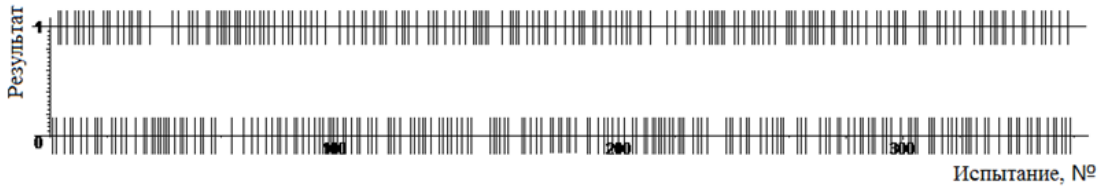

Рис. 1. Распределение результатов декодирования с использованием специальной обработки стираний (1 – успешное декодирование, 0 – сообщение не было восстановлено)

Рис. 2. Распределение результатов декодирования без использования специальной обработки стираний (1 – успешное декодирование, 0 – сообщение не было восстановлено)

Также наряду с RS(128, 19) исследовались и коды меньшей длины (RS(64, 5) и RS(16, 2)), результаты численных экспериментов для которых не имеют качественных отличий от результатов, полученных для RS(128,19). Численные эксперименты, направленные на уменьшение объема выходного списка GS-декодера, показали, что используемый алгоритм обработки стираний не позволяет значительно его (объем списка) сократить. Среднее уменьшение объема выходного списка составило менее 1 %. Полученные результаты являются новыми.

Заключение

Численные эксперименты показали, что использование описанной схемы и разработанного алгоритма приводит к повышению частоты успешных декодирований за границей исправляющей способности пары (RS-код, GS-декодер), а это значит, что корректирующая способность списочного декодера несколько увеличилась. Стоит отметить, что таким образом можно получить выигрыш в исправляющей способности связки кодер-декодер, не прибегая к смене кодера, что особенно важно, если параметры исходного кода ограничены на физическом уровне устройства, реализующем функции кодера. Численный анализ возможности уменьшения конечного объема выходного списка декодера показал, что статистически значимых изменений с применением специальной обработки стираний достичь не удается.

Литература

1. Sudan, M. Decoding of Reed Solomon codes beyond the error correction bound / M. Sudan // J. Compl. – 1997. – Vol. 13. – P. 180–193.

2. Guruswami, V. Improved decoding of Reed-Solomon and algebraic-geometry codes / V. Guruswami, M. Sudan // IEEE Transactions on Information Theory. – 1999, September. – Vol. 45. – P. 1757–1767. DOI: 10.1109/18.782097

3. Sudan, M. Lectures «Algorithmic Introduction to Coding Theory» / M. Sudan. – 2001.

4. McEliece, R.J. The Guruswami-Sudan Decoding Algorithm for Reed-Solomon Codes / R.J. McEliece // IPN Progress Report 42-153. – May 15, 2003. – P. 1–60.

Кряквин Вадим Донатович, канд. физ.-мат. наук, доцент кафедры алгебры и дискретной математики, Институт математики, механики и компьютерных наук им. И.И. Воровича, Южный федеральный университет, г. Ростов-на-Дону; vadkr@math.rsu.ru.

Крыжановский Константин Викторович, аспирант кафедры алгебры и дискретной математики, Институт математики, механики и компьютерных наук им. И.И. Воровича, Южный федеральный университет, г. Ростов-на-Дону; ccxq@yandex.ru.

Поступила в редакцию 30 апреля 2017 г.

DOI: 10.14529/ctcr170318

ABOUT SOME NUMERICAL EXPERIMENTS ON A LIST DECODER

***V.D. Kryakvin**, vadkr@math.rsu.ru,*
***K.V. Kryzhanovskiy**, ccxq@yandex.ru*
Southern Federal University, Rostov-na-Donu

This paper deals with the fundamental possibility of implementing successful decoding of messages the number of errors in which a priori exceeds the pair's correcting ability (RS-code, GS-decoder), for some types of Reed-Solomon codes. To test the hypothesis of the existence of such a possibility, a model was constructed and a special algorithm based on the processing of erasures

was developed, and computational experiments were carried out. The change in the average power of the output list of the GS decoder using this algorithm was also analyzed. It is shown that with the use of the proposed algorithm the frequency of successful decoding increases, while there is no statistically significant change in the average power of the output list.

Keywords: algebraic coding, RS-code, GS-decoder, error-correcting capability, decoding.

References

1. Madhu Sudan. Decoding of Reed Solomon Codes Beyond the Error Correction Bound. *J. Compl*., 1997, vol. 13, pp. 180–193.

2. Guruswami V., Sudan M. Improved Decoding of Reed-Solomon and Algebraic-Geometry Codes. *IEEE Transactions on Information Theory*, 1999, September, vol. 45, pp. 1757–1767. DOI: 10.1109/18.782097

3. Madhu Sudan. *Lectures "Algorithmic Introduction to Coding Theory"*, 2001.

4. McEliece R.J. The Guruswami-Sudan Decoding Algorithm for Reed-Solomon Codes, *IPN Progress Report* 42-153, May 15, 2003, pp. 1–60.

Received 30 April 2017

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Кряквин, В.Д. О некоторых численных экспериментах над списочным декодером / В.Д. Кряквин, К.В. Крыжановский // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2017. – Т. 17, № 3. – С. 153–158. DOI: 10.14529/ctcr170318

FOR CITATION

Kryakvin V.D., Kryzhanovskiy K.V. About Some Numerical Experiments on a List Decoder. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2017, vol. 17, no. 3, pp. 153–158. (in Russ.) DOI: 10.14529/ctcr170318