УДК 621.391

ШТОМПЕЛЬ Н.А., к.т.н., доцент (УкрГАЖТ)

Оценка вычислительной сложности методов кодирования кодами с малой плотностью проверок на четность

Проведена оценка вычислительной сложности методов кодирования кодами с малой плотностью проверок на четность. Показано, что улучшенные методы кодирования на основе преобразованной проверочной матрице и представления проверочной матрицы с помощью графа Таннера имеют линейную вычислительную сложность.

Ключевые слова: *коды с малой плотностью проверок на четность, кодирование, вычислительная сложность.*

Постановка проблемы и анализ литературы

Энергетическая эффективность кодов с малой плотностью проверок на четность (МППЧ-кодов) с большой длиной кодового слова практически достигает границы Шеннона при сохранении приемлемой вычислительной сложности декодирования [1]. Для классических линейных блоковых кодов (например, кодов БЧХ, кодов Рида-Соломона) процесс кодирования обычно менее сложный, чем процесс декодирования [2]. МППЧ-коды также относятся к классу линейных блоковых кодов, однако их ключевым отличием является разреженная проверочная матрица, что в общем случае приводит к порождающей матрице плотно заполненной ненулевыми символами. В связи с этим процесс кодирования на основе порождающей матрицы обладает сравнительно высокой вычислительной сложностью. Для уменьшения вычислительной сложности кодирования МППЧ-кодами применяются методы, основанные на следующих идеях: использовании алгебраических свойств проверочной матрицы; приведении произвольной (случайной) проверочной матрицы к заданной форме; представлении проверочной матрицы в виде графа Таннера. Хотя алгебраические МППЧ-коды часто имеют линейную сложность кодирования, более эффективными являются случайные МППЧ-коды [3], поэтому актуальной задачей является рассмотрение особенностей улучшенных с вычислительной точки зрения методов кодирования данными кодами.

Цель статьи

Исследование особенностей стандартного и улучшенных методов кодирования МППЧ-кодами и оценка их вычислительной сложности.

Основная часть

Рассмотрим (n,k) МППЧ-код, заданный случайной разреженной проверочной матрицей H размера $(n-k)\times n$. Совершая элементарные операции над строками и столбцами порождающей матрицы, ее можно привести к следующей канонической форме:

$$H_{sys}=[P\ I_{(n-k)}], \quad (1)$$

где P – матрица размера $(n-k)\times k$;

$I_{(n-k)}$ – единичная матрица размера $(n-k)\times(n-k)$.

С учетом того, что пространство строк порождающей матрицы ортогонально пространству строк проверочной матрицы, т.е. $GH^T=0$, соответствующую порождающую матрицу (n,k) МППЧ-кода представим следующим образом:

$$G_{sys}=[I_k\ P^T], \quad (2)$$

где I_k – единичная матрица размера $k\times k$;

P^T – матрица размера $k\times(n-k)$.

Тогда (1) и (2) задают эквивалентный (n,k) МППЧ-код в систематической форме, а отображение информационного слова $m=[m_1\ m_2\ \dots m_k]$ в кодовое слово c можно осуществить непосредственно путем его умножения на порождающую матрицу G_{sys}

$$c=mG_{sys}=[c_1\ c_2\ \dots c_k\ c_{k+1}\dots c_n]=[m\ p], \quad (3)$$

где m – информационная часть кодового слова длиной k;

p – проверочная часть кодового слова длиной $r = n - k$.

Таким образом, процесс кодирования состоит из двух частей: предварительных операций для нахождения порождающей матрицы (2) и непосредственно нахождения кодового слова (3). Вычислительная сложность первого этапа кодирования составляет $O(n^3)$ операций, но это не является ограничением, т.к. он выполняется предварительно. Более существенный недостаток порождающей матрицы (2) заключается в том, что она не является разреженной, поэтому второй этап процесса кодирования обладает квадратичной вычислительной сложностью, т.е. требует $O(k \times r)$ или, в более общем случае, $O(n^2)$ операций. Данный факт не позволяет применять «прямой» метод кодирования для МППЧ-кодов с большой длиной кодового слова n.

Для уменьшения вычислительной сложности кодирования произвольными МППЧ-кодами возможно применение метода на основе преобразования случайной формы проверочной матрицы H в приближенную нижнюю треугольную форму [3]. Для сохранения свойства разреженности проверочной матрицы допускается только перестановка строк и столбцов, в результате данных операций получается матрица следующего вида:

$$H_t = \begin{bmatrix} A & B & T \\ C & D & E \end{bmatrix}, \qquad (4)$$

где T – матрица, имеющая нижнюю треугольную форму, размера $(r-g)\times(r-g)$ (это означает, что T содержит единицы на диагонали от верхнего левого угла до нижнего правого угла, а на всех позициях выше диагонали – нули);

A, B – матрицы с размерами $(r-g)\times k$ и $(r-g)\times g$ соответственно (если исходная матрица H имеет полный ранг);

C, D, E – матрицы с размерами $g \times k$, $g \times g$ и $g \times (r-g)$ соответственно.

Тогда g строк матрицы H_t, оставшихся в матрицах C, D, E, соответствуют разнице между приближенной и идеальной нижними треугольными формами исходной матрицы H, при этом, чем меньше g, тем меньше вычислительная сложность кодирования МППЧ-кодом.

После получения приближенной нижней треугольной формы проверочной матрицы (4), применяется исключение Гаусса-Жордана для преобразования матрицы E в нулевую матрицу

$$\tilde{H} = \begin{bmatrix} I_{(r-g)} & 0 \\ -ET^{-1} & I_g \end{bmatrix} H_t = \begin{bmatrix} A & B & T \\ \tilde{C} & \tilde{D} & 0 \end{bmatrix}, \qquad (5)$$

где $\tilde{C} = -ET^{-1}A + C$, $\tilde{D} = -ET^{-1}B + D$, $\tilde{E} = -ET^{-1}T + E = 0$.

Т.к. матрица T имеет нижнюю треугольную форму, то при использовании исключения Гаусса-Жордана (5) влияние оказывается только на матрицы $\tilde{C}$ и $\tilde{D}$; остальные составляющие проверочной матрицы (4) остаются разреженными.

Далее для осуществления кодирования с использованием полученной матрицы $\tilde{H}$ кодовое слово c делится на три части, так что $c = [m\ p^1\ p^2]$, где $p^1 = [p_1^1\ p_2^1 \dots p_g^1]$ – вектор, содержащий g первых проверочных символов, $p^2 = [p_1^2\ p_2^2 \dots p_{r-g}^2]$ – вектор, содержащий остальные $(r-g)$ проверочные символы.

Кодовое слово c МППЧ-кода должно удовлетворять проверочному уравнению $c\tilde{H}^T = 0$, следовательно:

$$Am + Bp^1 + Tp^2 = 0, \qquad (6)$$

$$\tilde{C}m + \tilde{D}p^1 + 0\,p^2 = 0. \qquad (7)$$

Поскольку матрица E является нулевой, то вектор p^1 зависит только от информационной части m и поэтому может быть определен независимо от вектора p^2. Если матрица $\tilde{D}$ является обратимой, то вектор p^1 можно найти из (7)

$$p^1 = \tilde{D}^{-1}\tilde{C}m. \qquad (8)$$

Если матрица $\tilde{D}$ не является обратимой, тогда необходимо соответствующим образом переставить

столбцы матрицы $\tilde{H}$. Сохраняя минимально допустимый размер g, можно обеспечить вычислительную сложность произведения (8) на уровне $O(n+g^2)$ операций.

После того как вектор p^1 найден, вектор p^2 определяется из (6)

$$p^2 = -T^{-1}(Am + Bp^1). \quad (9)$$

Разреженность матриц A, B и T используется для обеспечения низкой сложности вычисления (9). Т.к. матрица T имеет нижнюю треугольную форму, то фактически вектор p^2 может быть найден с использованием обратной подстановки, что требует $O(n)$ операций.

Таким образом, общая вычислительная сложность данного метода кодирования МППЧ-кодами составляет $O(n+g^2)$ операций.

Другой подход к уменьшению вычислительной сложности кодирования заключается в использовании метода декодирования МППЧ-кодов для канала со стиранием символов на основе итеративного обмена сообщениями между переменными v_i и проверочными q_j вершинами графа Таннера, построенного на основе проверочной матрицы H (где $i=1,2,\ldots,n$, $j=1,2,\ldots,r$) [4].

Для осуществления кодирования (n,k) МППЧ-кодом среди переменных вершин выбираются k вершин, соответствующих известным символам информационного слова m, а остальные r вершин представляют собой неизвестные (стертые) символы проверочной части кодового слова (3). Для нахождения стертых символов используется стандартный для МППЧ-кодов метод декодирования «обмен сообщениями», имеющий линейную вычислительную сложность. Следовательно, данный метод кодирования также требует $O(n)$ операций, но при этом может возникать проблема существования «множества остановки», приводящая к невозможности нахождения проверочной части кодового слова p.

Выводы

«Прямой» метод кодирования МППЧ-кодами обладает квадратичной вычислительной сложностью, что не позволяет использовать произвольные МППЧ-коды с большой длиной кодового слова на практике. Улучшенные методы кодирования на основе преобразованной проверочной матрице и представления проверочной матрицы с помощью графа Таннера, имеющие линейную вычислительную сложность, позволяют снизить данное ограничение.

Литература

1 MacKay, D.J.C. Good error correcting codes based on very sparse matrices / D.J.C. MacKay // IEEE Transaction on Information Theory. – 1999. – March. – vol. 45, №2. — P. 399-432.

2 Блейхут, Р. Теория и практика кодов, контролирующих ошибки: пер. с англ. [Текст] / Р. Блейхут. – М.: Мир, 1989. – 448 с.

3 Richardson, T.J. Efficient encoding of low-density parity-check codes / R.L. Urbanke, T.J Richardson // IEEE Transactions on Information Theory. – 2001. – February. – vol. 47, №2. – P. 638-656.

4 Haley, D. Iterative encoding of low-density parity-check codes / D. Haley, A. Grant, J. Buetefuer // Proceedings IEEE Global Communications Conference (GLOBECOM), November, 2002, vol. 2. – 2002. – P. 1289-1293.

Штомпель М.А. Оцінка обчислювальної складності методів кодування кодами з малою щільністю перевірок на парність. Проведена оцінка обчислювальної складності методів кодування кодами з малою щільністю перевірок на парність. Показано, що покращені методи кодування на основі перетвореної перевірочної матриці та представлення перевірочної матриці за допомогою графа Таннера мають лінійну обчислювальну складність.

Ключові слова: коди з малою щільністю перевірок на парність, кодування, обчислювальна складність.

———————————————

Shtompel N.A. Estimation of the computational complexity of methods of encoding by means of low-density even parity check codes. The estimation of the computational complexity of methods of encoding by means of low-density even parity check codes has been carried out. It is shown that the improved methods of encoding based on the transformed check matrix and check matrix representation using a Tanner graph have a linear computational complexity.

Key words: low-density even parity check codes, encoding, computational complexity.

Рецензент д.т.н., профессор Краснобаев В.А. (Полтавский национальный технический университет им. Ю.Кондратюка)

Поступила 14.08.2013 г.