УДК 519.688

Владимиров С.М. [1]

[1] Московский физико-технический институт

Использование кодов с малой плотностью проверок на чётность

Теорема Шеннона утверждает, что при определённых условиях вероятность ошибки декодирования (то есть невозможность декодером исправить ошибку передачи) можно уменьшить, выбрав большую длину ключевого слова. Однако, данная теорема (и работа вообще) не показывает, как можно выбрать большую длину, а точнее как эффективно организовать процесс кодирования и декодирования информации с большой длиной ключевых слов. Если предположить, что в кодере и декодере есть некие таблицы соответствия между входным блоком информации и соответствующим кодовым словом, то такие таблицы будут занимать очень много места. Для двоичного симметричного канала без памяти (если говорить упрощённого, то на вход кодера поступает поток из нулей и единиц) количество различных блоков составляет 2n, где n — количество бит (нулей или единиц) которые будут преобразовываться в одно кодовое слово. Для 8 бит это 256 блоков информации, каждый из которых будет содержать в себе соответствующее кодовое слово. Причём кодовое слово обычно большей длины, так как содержит в себе дополнительные биты для защиты от ошибок передачи данных. Поэтому одним из способов кодирования является использование проверочной матрицы, которые позволяют за одно математическое действие (умножение строки на матрицу) выполнить декодирование кодового слова. Аналогичным образом каждой проверочной матрице соответствует порождающая матрица, аналогичным способом одной операцией умножения строки на матрицу генерирующей кодовой слово.

Таким образом, для сравнительно коротких кодовых слов кодеры и декодеры могут просто содержать в памяти все созможные варианты, или даже реализовывать их в виде полупроводниковой схемы. Для большего размера кодового слова эффективнее хранить порождающую и проверочную матрицу. Однако, при длинах блоков в несколько тысяч бит хранение матриц размером, соответственно, в мегабиты, уже становится неэффективным. Одним из способов решения данной проблемы становится использования кодов с малой плотностью проверок на чётность, когда в проверяющей матрице количество единиц сравнительно мало, что позволяет эффективнее организовать процесс хранения матрицы или же напрямую реализовать процесс декодирования с помощью полупроводниковой схемы.

Первой работой на эту тему стала работа Роберта Галлагера «Low-Density Parity-

Check Codes» 1963 года (основы которой были заложены в его докторской диссертации 1960 года). Поэтому часто LDPC-коды назвают кодами Галлагера. В работе учёный описал требования к таким кодам, описал возможные способы построения и способы их оценки. В настоящий момент (2007) данные коды являются наиболее эффективными.

Однако, из-за сложности в реализации кодеров и декодеров эти коды были позабыты[1]. Позже с развитием телекоммуникационных технологий снова возрос интерес к передачи информации с минимальными ошибками. Несмотря на сложность реализации по сравнению с турбо-кодом, отсутствие преград к использованию (незащищённость патентами) сделало LDPC-коды привлекательными для телекоммуникационной отрасли, и фактически стали стандартом де-факто. В 2003 году LDPC-код, вместо турбо-кода, стал частью стандарта DVB-S2 спутниковой передачи данных для цифрового телевидения. Аналогичная замена произошла и в стандарте DVB-T2 для цифрового наземного телевизионного вещания.

LDPC-коды описаываются проверочной матрицей, содержащей в основном нули и относительно малое количество единиц. В частности, у (n, j, k) LDPC-кода, проверочная матрица имеет n строк (что соответствует длине кодового слова), каждый столбец матрицы содержит малое фиксированное количество единиц j, и каждая строка — k единиц.

Обычно рассматриваются матрицы больших размеров. Например, в работе Галлагера в разделе экспериментальных результатов используются «малые» количества строк n=124, 252, 504 и 1008 строк (число столбцов проверочной матрицы немного больше). На практике применяются матрицы с большим количеством элементов — от 10 до 100 тысяч строк. Однако количество единиц в строке или в столбце остаётся достаточно малым, обычно меньшим 10. Замечено, что коды с тем же количество элементов на строку или столбец, но с большим размером, обладают лучшими характеристиками.

Важной характеристикой матрицы LDPC-кода является отсутсвие циклов определённого размера. Под циклом длины 4 понимают образование в проверочной матрице прямоугольника, в углах которого стоят единицы. Отсутствие цикла длины 4 можно также определить через скалярное произведение столбцов (или строк) матрицы. Если каждое попарное скалярное произведение всех столбцов (или строк) матрицы не более 1, это говорит об отсутствии цикла длины 4. Циклы большей длины (6, 8, 10 и т. д.) можно определить, если в проверочной матрице построить граф, вершинами которого являются единицы, а рёбра — все возможные соединения вершин, параллельные сторонам матрицы (то есть вертикальные или горизонтальные линии). Минимальный цикл в этом графе и будет минимальным циклом в проверочной матрице LDPC-кода.

Хотя LPDC-код, как и любой линейный блоковый код описывается проверочной

матрицей, при использовании на практике предпочитают описание в виде двудольного графа Танстола, который содержит информацию о парах индексов строк и столбцов, на пересечении которых есть единица. Также, если используются специальные случаи построения LDPC-кодов, могут использоваться и специальные способы задания матриц.

В настоящее время используются два принципа построения проверочной матрицы кода. Первый основан на генерации начальной проверочной матрицы с помощью псевдослучайного генератора. Второй - использование специальных техник, основанных, например, на группах и конечных полях.

В своей работе Галлагер предпочёл с помощью генератора псевдослучайных чисел создать начальную проверочную матрицу небольшого размера с заданными характеристиками, а далее увеличить её размер, дублируя матрицу и используя технику пермутации (перемешивания строк и столбцов) для избавления от циклов определённой длины.

В 2003 году Джеймсом МакГованом и Робертом Вильямсоном был предложен способ удаления циклов из матрицы LDPC-кода, в связи с чем стало возможным в начале сгенерировать матрицу с заданными характеристиками (n, j, k), а затем удалить из неё циклы. Так происходит в схеме Озарова-Вайнера[2].

СПИСОК ЛИТЕРАТУРЫ

1. *David J.C. MacKay* Information theory, inference and learning algorithms, CUP. — 2003 — ISBN 0-521-64298-1
2. *Косолапов Ю.В.* О применении схемы озарова-вайнера для защиты информации в беспроводных многоканальных системах передачи данных // Информационное противодействие угрозам терроризма: Научно-практический журнал. — 2007. — № 10. — С. 111-120.
3. *David J.C. MacKay, Radford M.* Neal Near Shannon Limit Performance of Low Density Parity Check Codes.
4. *Pearl J.* Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo, 1988.
5. *Shannon C.E.* A Mathematical Theory of Communication // Bell System Technical Journal. — 1948. — Т. 27. — С. 379-423, 623–656.
6. *Gallager R. G.* Low Density Parity Check Codes. — Cambridge: M.I.T. Press, 1963. — P. 90.