О ДЕКОДИРОВАНИИ НЕКОТОРЫХ ВИДОВ ОШИБОК В ДВУМЕРНЫХ КОДАХ-ПРОИЗВЕДЕНИЯХ

Липницкий В.А., Сергей А.И.

Военная академия Республики Беларусь
Гродненский государственный университет имени Янки Купалы

В статье описывается подход к решению задачи точного декодирования двумерных кодов-произведений, построенных на основе двоичных линейных кодов. Формулируется и доказывается лемма, которая служит вспомогательным инструментом для построения алгоритма исправления любых конфигураций ошибок в пределах теоретической корректирующей способности кода-произведения.

Идея кодов-произведений заключается в построении мощных помехоустойчивых кодов на основе более простых базовых кодов.

Кодом-произведением $C = C_1 \times C_2$ называется код, словами которого являются все двумерные матрицы со строками, являющимися словами кода C_1 и столбцами, являющимися словами кода C_2. При этом минимальное расстояние кода C равно произведению минимальных расстояний кодов сомножителей [1]. Способ построения кодов-произведений с высокими корректирующими возможностями описан в [2].

Одним из главных преимуществ кодов-произведений является способность справляться с ситуациями, когда ошибки возникают с высокой частотой в коротком промежутке, т. е. исправлять так называемые пакетные ошибки, нередко встречающиеся на практике. В таких случаях количество реально исправленных ошибок сильно превосходит теоретическую корректирующую способность кода.

Для практических применений скорость часто оказывается более важным параметром, чем качество декодирования, поэтому широкое распространение получили так называемые блочные турбокоды. Для расшифровки турбокодов применяют, как правило, простую итеративную процедуру, которая представляет собой компромисс между скорость работы и качеством дешифровки. Существуют конфигурации ошибок, которые итерационный метод не способен исправить, хотя корректирующая способность кода позволяет это сделать.

Что касается точного декодирования, то существуют подходы, требующие использования базовых кодов специального вида. Например, в [1] описана схема точного исправления ошибок в кодах-произведениях, построенных на основе БЧХ-кодов.

В данной статье исследуется возможность точного декодирования кодов-произведений, сконструированных на основе более широкого класса кодов, а именно двоичных линейных кодов.

Рассмотрим сначала пример построения двумерного кода-произведения. В качестве базовых кодов C_1 и C_2 возьмем классический код Хэмминга (7, 4, 3), исправляющий однократную ошибку. Т. е. в рассматриваемом случае используется один и тот же код и для строк, и для столбцов.

Таким образом, получившийся код-произведение имеет 16 информационных разрядов. Перед началом кодирования биты передаваемого сообщения записываются в виде матрицы 4×4. Далее строки этой матрицы независимо кодируются с помощью базового кода C_1, после чего к столбцам полученной 4×7 матрицы применяется аналогичная процедура, т. е. каждый столбец кодируется с использованием базового кода C_2. В результате получится матрица 7×7, которая и будет кодовым словом, соответствующим передаваемому сообщению.

Можно доказать, что построенный код имеет минимальное расстояние, равное 9, т. е. исправляет четырехкратные ошибки [3].

В общем случае, разумеется, базовые коды для строк и столбцов могут быть различными. Например, если взять за основу коды с параметрами (17, 9, 5) и (15, 11, 3), то полученный в результате код-произведение будет иметь параметры (15×17, 11×9, 3×5) или (255, 99, 15), т. е. позволит исправлять семикратные ошибки [2].

Хотя минимальное расстояние кода-произведения теоретически вычисляется легко, процесс декодирования представляет собой сложную задачу. Вызвано это в первую очередь богатым разнообразием конфигураций возникающих ошибок [4]. Алгоритм декодирования также должен учитывать, что базовые коды могут неверно декодировать или вовсе не обнаружить ошибку в соответствующей строке/столбце, т. к. количество ошибок в отдельно взятой строке/столбце может превышать декодирующую способность базового кода.

Итеративные процедуры декодирования, обычно применяемые для исправления ошибок в кодах-произведениях, многократно повторяют операции поочередного исправлениях ошибок во всех строках, потом во всех столбцах и т. д. При этом, например, при декодировании строк никак не учитывается информация, известная об ошибках в столбцах, необходимая для того, чтобы результаты декодирования по строкам и по столбцам были согласованы друг с другом.

Далее приведена лемма, позволяющая учитывать информацию об ошибках в столбцах при декодировании строк и наоборот.

Лемма 1. Пусть C - линейный код с параметрами [n, k, d]. Для него справедливо следующее свойство:

Пусть известно, что ошибки могли возникнуть только в $s \le d$ заранее известных позициях $i_1, i_2, ..., i_s$ принятого вектора-слова длины n. Рассмотрим бинарный вектор v длины n, в котором в позициях $i_1, i_2, ..., i_s$ стоят единицы, а остальные элементы нулевые.

Если $v \notin C$, тогда синдромы всех возможных 2^s ошибок различны, т. е. можно однозначно исправить любые ошибки веса $\le s$.

Если $v \in C$ (и, следовательно, $s = d$), тогда имеется 2^{d-1} различных синдромов ошибок. Каждый синдром встречается ровно 2 раза: синдром вектора ошибок e совпадает с синдромом вектора $e + v$. Другими словами, вектор ошибок в таком случае определяется с точностью до инвертирования позиций $i_1, i_2, ..., i_d$.

Доказательство.

Пусть H – проверочная матрица кода C.

1. $s \le d - 1$. В этом случае утверждение Леммы 1 напрямую следует из того, что любые $d-1$ матрицы H линейно независимы [1] и образуют базис.

2. $s = d$ и $v \notin C$. Поскольку в матрице H любые $d-1$ столбцов линейно независимы, то столбцы $i_1, i_2, ..., i_d$ могут быть линейно зависимыми только в случае, если $Hv^T = 0$. Но это невозможно, т. к. $v \notin C \Rightarrow Hv^T \neq 0$. Поэтому столбцы $i_1, i_2, ..., i_d$ линейно независимы и все 2^d синдромов различны.

3. $s = d$ и $v \in C$. Пусть A – матрица, образованная столбцами $i_1, i_2, ..., i_d$ матрицы H.

Рассмотрим следующую систему уравнений над $GF(2)$: $Ax^T = y$, где y – произвольный синдром ошибки.

Ранг матрицы A равен $d-1$, поэтому имеется $d-1$ базисных переменных и одна свободная. Таким образом, система имеет ровно два решения. При этом, поскольку $v \in C \Rightarrow Hv^T = 0 \Rightarrow A(1 \ 1 \ \dots \ 1)^T = 0$, значит, если $Ax^T = 0$, то и $A(x + (1 \ 1 \ \dots \ 1))^T = 0$, т. е. кодовое слово определяется с точностью до инвертирования позиций $i_1, i_2, ..., i_d$.

Лемма доказана.

Продемонстрируем возможности применения леммы 1 на примере вышеупомянутого кода [255, 99, 15] ([15 x 17, 11 x 9, 3 x 5]). Этот код позволяет исправить вплоть до 7-ми произвольных ошибок. Заметим, что если ошибки возникли не более чем в 5 столбцах, и при этом в каждом из этих столбцов ошибка была зарегистрирована, то с помощью Леммы 1 можно точно декодировать исходное сообщение по строкам, даже если количество ошибок в отдельно взятой строке выходит за пределы декодирующей возможности кода C_1.

Нетрудно видеть, что существуют конфигурации ошибок, в которых одного только применения Леммы 1 недостаточно для декодирования сообщения. Разбор таких частных случаев является необходимым условием для построения алгоритма точного декодирования кодов-произведений, но выходит за рамки данной статьи.

Такими случаями являются:

- сильно разреженные ошибки, возникающие сразу во многих строках и многих столбцах;
- существование строки/столбца, вектор ошибок в котором представляет собой кодовое слово, а, следовательно, ошибка в этой строке/столбце не будет зарегистрирована.

В статье рассмотрена Лемма 1 – вспомогательный инструмент, которым можно пользоваться при расшифровке кодов-произведений. Использование леммы вместе с разбором нескольких частных случаев позволяет построить алгоритм точного декодирования кодов-произведений [63, 12, 9] ([7 x 9, 4 x 3, 3 x 3]) и [255, 99, 15] ([15 x 17, 11 x 9, 3 x 5]).

Список использованных источников:

6. Блейхут Р. Теория и практика кодов, контролирующих ошибки / Р. Блейхут. – М. : Мир, 1986. – 576 с.

7. Липницкий В.А. Тензорные произведения кодов Хэмминга. – 11-я Белорусская математическая конференция: Тезисы докладов Междунар. науч. Конф. Минск, 5 – 9 ноября 2012 г. – Часть 4. – Мн.: Институт математики НАН Беларуси, 2012. – С. 62 – 63.

8. Мак-Вильямс, Ф. Дж. Теория кодов исправляющих ошибки / Ф. Дж. Мак-Вильямс, Н. Дж. А. Слоэн. – М. : Связь, 1979. – 744 с.

9. Липницкий В.А., Сергей А.И., Спичекова Н.В. Классификация точечных образов. История и современность. // XI-я Белорусско-российская НТК «Технические средства защиты информации», г. Минск, 5 – 6 июня 2013 г. Тезисы докладов. – Мн.: БГУИР, 2013. – С. 42.