Код Рида-Маллера

Илья Коннов

Факультет компьютерных наук Высшая Школа Экономики

13 февраля 2022 г.

Код Рида-Маллера

2022-02-13

1. Если вы смотрите презентацию, то на сером фоне справа иногда видны некоторые ценные комментарии, для которых поля слайда оказались слишком узки. Если вы читаете pdf-ку, то эти комментарии уже находятся в самом подходящем для них месте в тексте (а в внешних полях видны заголовки слайдов). Если вы смотрите мой доклад и видите этот текст, то что-то пошло серьёзно не так. Да, у этого одного файла есть три разные версии. По любым вопросам: ReedMuller@sldr.xyz или t.me/iliago или vk.com/iliago.

Введение

Описан Дэвидом Маллером (автор идеи) и Ирвингом Ридом (автор метода декодирования) в сентябре 1954 года. Обозначаются как $\mathrm{RM}(r,m)$, где r — ранг, а 2^m — длина кода. Кодирует сообщения длиной $k = \sum_{i=0}^r C_m^i$ при помощи 2^m бит.

Традиционно, считается что коды бинарные и работают над битами, т.е. \mathbb{Z}_2 .

Соглашение: сложение векторов $u,v\in\mathbb{Z}_2^n$ будем обозначать как $u \oplus v = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n).$



Булевы функции и многочлен Жегалкина

Всякую булеву функцию можно записать при помощи таблицы истинности

$$\begin{array}{c|ccc} x & y & f(x,y) \\ \hline 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ \end{array}$$

И при помощи многочлена Жегалкина:

$$f(x,y) = xy + x + y + 1$$



Многочлены Жегалкина

В общем случае, многочлены будут иметь следующий вид:

$$f(x_1,x_2,...,x_m) = \sum_{S \subseteq \{1,\dots,m\}} c_S \prod_{i \in S} x_i$$

Например, для m=2:

$$f(x_1, x_2) = c_1 \cdot x_1 x_2 + c_2 \cdot x_1 + c_3 \cdot x_2 + c_4 \cdot 1$$

Всего $n=2^m$ коэффициентов для описания каждой функции.



Функции небольшой степени

Рассмотрим функции, степень многочленов которых не больше r:

$$\{f(x_1,x_2,...,x_m)\mid \deg f\leq r\}$$

Каждую можно записать следующим образом:

$$f(x_1, x_2, ..., x_m) = \sum_{\substack{S \subseteq \{1, ..., m\} \\ |S| < r}} c_S \prod_{i \in S} x_i$$

В каждом произведении используется не больше rпеременных.

Сколько тогда всего коэффициентов используется?
$$k = C_m^0 + C_m^2 + \ldots + C_m^r = \sum_{i=0}^r C_m^i$$

Код Рида-Маллера ∟_{Введение} 2022-02-13

Функции небольшой степени

- 1. Замечу, что при $S=\varnothing$, мы считаем, что $\prod_{i\in S}x_i=1$, таким образом всегда появляется свободный член.
- 2. Если говорить несколько проще, то для составления многочленов мы сложим сначала одночлены (x+y+z+...), затем произведения одночленов (xy+yz+xz+...) и т.д. вплоть до r множителей (поскольку мы работаем в поле \mathbb{Z}_2 , здесь нету x^2,y^2,z^2 , т.к. $a^2=a$). Тогда легко видеть, почему k именно такое: мы складываем все возможные перестановки сначала для 0 переменных, потом для одной, двух, и так до всех r

Идея кодирования

Пусть каждое сообщение (длины k) — коэффициенты многочлена от m переменных степени не больше r.Тогда мы можем его представить при помощи 2^m бит, подставив все возможные комбинации переменных.

Таким образом получим таблицу истинности, из которой позднее сможем восстановить исходный многочлен, а вместе с ним и сообщение.

Зафиксировав в таблице порядок строк, можно выделить вектор значений, который и будет кодом.

Код Рида-Маллера

2022-02-13

∟Идея кодирования



- 1. Их 2^m , поскольку рассматриваем многочлены только над \mathbb{Z}_2 от m
- 2. Вектор значений обозначается $\operatorname{Eval}(f)$ столбец таблицы истинности, содержащий значения функции. Имеет смысл только при зафиксированном порядке строк в таблице. У меня он везде самый обычный, как в примере выше.



Пример

r = 1 (степень многочлена), m = 2 (переменных). Это RM(1,2).

 \blacksquare Тогда наш многочлен: $f(x,y)=c_1x+c_2y+c_3.$

lacktriangle Сообщение: 101, тогда f(x,y) = x + 0 + 1.

■ Подставим всевозможные комбинации:

$$\begin{array}{c|cccc} x & y & f(x,y) \\ \hline 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ \end{array}$$

■ Получили код: $\mathrm{Eval}(f) = 1100$.

Код Рида-Маллера 2022-02-13

□Пример



1. Здесь и далее я для краткости и удобства записываю битовые векторы не как $(1 \quad 0 \quad 0 \quad 1)$, а как 1001 при помощи нескучного шрифта.



Декодирование когда потерь нет

- Мы получили код: 1100
- Представим таблицу истинности.

x	y	f(x,y)
0	0	1
0	1	1
1	0	0
1	1	0

■ Подстановками в $f(x,y) = c_1 x + c_2 y + c_3$ получим СЛАУ.

$$\left\{ \begin{array}{cccc} & c_3 &=& 1 \\ & c_2 &+& c_3 &=& 1 \\ c_1 &+&& c_3 &=& 0 \\ c_1 &+& c_2 &+& c_3 &=& 0 \end{array} \right.$$

 $c_1 = 1, c_2 = 0, c_3 = 1,$ исходное сообщение: 101.

Код Рида-Маллера 2022-02-13 -Кодирование

___Декодирование когда потерь нет



1. Теперь покажем, как можно декодировать когда потерь нет. Этот пример — продолжение предыдущего.

Коды 0-го порядка

Для случая $\mathrm{RM}(0,m)$ нужна функция от m аргументов, степени не выше 0.

 $\quad \blacksquare \ f(x_1,x_2,...,x_m)=0$

ование
$$g(x_1, x_2, ..., x_m) = 1$$
 Таблица истинности:

 $r \perp f(r, \dots, r) = g(x_1, \dots)$

	x_1	x_2	•••	x_m	$J(x_1,, x_m)$	$g(x_1,, x_m)$
2^m	(0	0		0	0	1
	0	0	•••	1	0	1
	ĺ		٠.			
	(1	1		1	0	1

Вывод: это 2^m -кратное повторение символа

- Сообщение 0 даст код 00...0
- Сообщение 1 даст код 11...1

Код Рида-Маллера

2022-02-13

└─Коды 0-го порядка



- 1. Отдельно стоит рассмотреть вариант кода при r=0, он нам в будущем пригодится для доказательств.
- 2. Таких функций существует всего лишь две, поскольку мы можем влиять лишь на свободный член. Все остальные коэффициенты обнуляются из-за требования $\deg f \leq 0$.
- 3. Здесь число строк, как и в любой другой таблице истинности, равно m , а колонки с значениями никак не зависят от аргументов функций. Получается две колонки – одна с нулями, другая с единицами.



Доказательство линейности

Пусть C(x) кодирует сообщение $x \in \mathbb{Z}_2^k$ в код $C(x) \in \mathbb{Z}_2^m$.

$$C(x) = (p_x(a_i) \mid a_i \in \mathbb{Z}_2^m)$$

где $p_x(a_i)$ — соответствующий сообщению x многочлен. Причём p_x берёт в качестве своих коэффициентов биты из x. Поскольку многочлены степени не выше r образуют линейное пространство, то $p_{(x \oplus y)} = p_x + p_y$.

$$C(x\oplus y)_i=p_{(x\oplus y)}(a_i)=p_x(a_i)+p_y(a_i)=C(x)_i+C(y)_i$$

т.е.
$$\forall x,y \quad C(x\oplus y) = C(x) + C(y)$$
, ч.т.д.

2022-02-13

Код Рида-Маллера

Свойства и параметры кода

—Доказательство линейности



- 1. Хотим показать, что этот код является линейным, т.е. что его кодовые слова образуют линейное пространство, и у нас есть изоморфизм из пространства сообщений (\mathbb{Z}_2^k) в пространство слов (\mathbb{Z}_2^m) . Для этого необходимо немного формализовать всё описанное раньше.
- 2. Пояснение: перебираем все векторы $a_i\ (2^m\ {
 m штук})$, подставляем каждый в p_x в качестве переменных и таким образом получаем вектор значений (длины 2^m). Именно он и называется кодом.
- 3. Напомню, что базис пространства многочленов выглядит примерно так: 1, x, y, z, xy, yz, xz (для трёх переменных, степени не выше 2). Чтобы преобразовать сообщение в многочлен, мы берём каждый бит сообщения и умножаем его на соответствующий базисный вектор. Очевидно, такое преобразование будет изоморфизмом. Именно поэтому $p_{(x+y)}=p_x+p_y.$ Обратите внимание, что сообщение x это не просто число (\mathbb{Z}_{2^k}) и мы рассматриваем его биты, а реально вектор битов (\mathbb{Z}_2^k) . У него операция сложения побитовая.
- 4. Здесь я использую запись $C(x)_i$ для i-го элемента вектора C(x). Поскольку i произвольное, то и весь вектор получился равен. Таким образом, этот код действительно линейный и к нему применимы уже известные теоремы!



Последствия линейности

 \blacksquare Существует порождающая матрица G.

$$C(x) = x_{1 \times k} G_{k \times n} = c_{1 \times n}$$

2 Минимальное растояние будет равно минимальному весу Хемминга среди всех кодов.

$$d = \min_{\substack{c \in C \\ c \neq 0}} w(c)$$

Корректирующая способность:

$$t = \left| \frac{d-1}{2} \right|$$

- 1. Так можно кодировать сообщения x в коды c. Но искать её мы не будем, обойдёмся одними многочленами, это интереснее. 2. Вес Хэмминга вектора — количество в нём ненулевых элементов.
- 3. Доказательство очень просто: минимальное расстояние вес разности каких-то двух различных кодов, но разность двух кодов тоже будет кодом, т.к. мы в линейном пространстве. Значит достаточно найти минимальный вес, но не учитывая нулевой вектор, т.к. разность равна нулю тогда и только тогда, когда коды равны.
- 4. Однако мы ещё не знаем как выглядят наши коды (как выглядят таблицы истинности функций степени не больше r?). А значит не можем ничего сказать про минимальное расстояние.

Конструкция Плоткина: многочлены

Хотим понять как выглядят кодовые слова.

- Код таблица истинности функции $f(x_1,...,x_m)\in \mathrm{RM}(r,m)$, причём $\deg f\leq r.$
- Разделим функцию по x_1 : $f(x_1,...,x_m) = g(x_2,...,x_m) + x_1 h(x_2,...,x_m). \label{eq:force}$
- lacksquare Заметим, что $\deg f \leq r$, а значит $\deg g \leq r$ и $\deg h \le r - 1$.

Код Рида-Маллера

2022-02-

-Свойства и параметры кода

Конструкция Плоткина

Конструкция Плоткина: многочлены

- 1. Порядок очевидно не больше r, потому что это условие для включения в пространство кодов $\mathrm{RM}(r,m)$.
- 2. Теперь у нас есть две функции от меньшего числа аргументов. Очевидно, так можно сделать всегда, когда m > 1



Конструкция Плоткина: таблица истинности

Ранее: $f(x_1,...,x_m) = g(x_2,...,x_m) + x_1 h(x_2,...,x_m).$

lacktriangle Заметим, что таблица истинности f состоит из двух

 $\operatorname{Eval}(f) = \left(\frac{\operatorname{Eval}^{[x_1=0]}(f)}{\operatorname{Eval}^{[x_1=1]}(f)}\right)$

lacksquare Причём $\operatorname{Eval}^{[x_1=0]}(f)=\operatorname{Eval}(g)$, а $\mathrm{Eval}^{[x_1=0]}(f) \oplus \mathrm{Eval}^{[x_1=1]}(f) = \mathrm{Eval}(h).$

частей: при $x_1 = 0$ и при $x_1 = 1$.

■ Таким образом, $\mathrm{Eval}(f) = (\mathrm{Eval}(g) \mid \mathrm{Eval}(g) \oplus \mathrm{Eval}(h)).$

Код Рида-Маллера

2022-02-13 Свойства и параметры кода

—Конструкция Плоткина

Конструкция Плоткина: таблица истинности

 $Eval(f) = \frac{\left(Eval^{\sigma_1 - it}(f)\right)}{\left(Eval^{\sigma_1 - it}(f)\right)}$

 $\operatorname{End}^{[p_g-b]}(f) = \operatorname{Eval}(g)$, \mathfrak{s} $(f) \oplus \operatorname{Eval}^{[p_g-b]}(f) = \operatorname{Eval}(h)$.

- 1. Теперь рассмотрим те же функции, но со стороны их таблиц истинности. Нам же интересны именно коды, а они как раз очень тесно связаны с этими таблицами.
- 2. Про обозначения: $\mathrm{Eval}(f)$ таблица для всей функции (вектор значений, елси точнее), $\mathrm{Eval}^{|x_1=0|}(f)$ кусок таблицы при $x_1=0$, $\mathrm{Eval}^{[x_1=1]}(f)$ — кусок таблицы при $x_1=1$. Они нам после этого доказательства больше не понадобятся.
- 3. Это всё следует из ранее полученного утверждения. Если мы подставим $x_1=0$, то останется только g — первое равенство очевидно. Если же мы рассмотрим $\mathrm{Eval}^{[x_1=1]}(f)$, то получим $\mathrm{Eval}(g+h)$, но если туда прибавить ещё раз $\mathrm{Eval}(g)$, то останется только $\mathrm{Eval}(h)$ (поскольку 1+1=0 в \mathbb{Z}_2) — получили второе равенство.
- 4. Палочка по центру конкатенация векторов.



Конструкция Плоткина: вывод

Если дана $f(x_1,...,x_m)$, причём $\deg f \leq r$, то можно её разделить:

 $f(x_1,...,x_m) = g(x_2,...,x_m) + x_1 h(x_2,...,x_m)$

Также известно, что

 $\text{Eval}(f) = (\text{Eval}(g) \mid \text{Eval}(g) \oplus \text{Eval}(h)).$

Заметим, что $\mathrm{Eval}(f)$ – кодовое слово (как и для g,h).

 $c=\mathrm{Eval}(f)\in\mathrm{RM}(r,m)$ $(\tau.\kappa. \deg f \leq r)$

 $u = \text{Eval}(g) \in \text{RM}(r, m - 1)$ (τ . κ . $\deg g \leq r$) $v = \text{Eval}(h) \in \text{RM}(r-1, m-1)$ (т.к. $\deg h \le r-1$)

Утверждение: Для всякого кодового слова $c \in \mathrm{RM}(r,m)$ можно найти $u \in \mathrm{RM}(r,m-1)$ и $v \in \mathrm{RM}(r-1,m-1)$, такие что $c = (u \mid u + v)$.

2022-02-

Код Рида-Маллера

-Свойства и параметры кода

—Конструкция Плоткина

└Конструкция Плоткина: вывод

$$\label{eq:local_constraints} \begin{split} & \operatorname{Torgat} : \\ & \operatorname{Torgat} : \\ & \varepsilon = \operatorname{End}(f) \in \operatorname{RM}(r, m) \qquad (\tau \times \deg f \le r) \\ & \varepsilon = \operatorname{End}(g) \in \operatorname{RM}(r, m-1) \qquad (\tau \times \deg g \le r) \\ & - \operatorname{End}(g) \in \operatorname{RM}(r-1, m-1) \qquad (\tau \times \deg g \le r) \\ & - \operatorname{End}(g) \in \operatorname{RM}(r-1, m-1) \qquad (\tau \times \deg g \le r-1) \end{split}$$

- 1. Теперь собираем всё это в одно важное утверждение.
- 2. Причём мы уже знаем, что $\deg g \leq r$ и $\deg h \leq r-1$, если $\deg f \leq r$
- 3. Напомню, что ${
 m RM}(r,m)$ включает в себя все функции (их таблицы истинности, если точнее) от m аргументов и степени не выше r.Очевидно, наши годятся.
- 4. Что здесь важно отметить оба наших новых кодовых слова u,vполучились «меньше», чем исходное c. \exists то позволяет, во-первых, устраивать индукцию по m, чем мы скоро и займёмся. Во-вторых, это позволяет легко строить большие

порождающие матрицы, но мы этим не будем заниматься.

Минимальное расстояние

Хотим найти минимальное расстояние для кода $\mathrm{RM}(r,m)$

$$d = \min_{c \in C, c \neq 0} w(c)$$

Предположим, что $d = 2^{m-r}$ и докажем по индукции. **База:** $\mathrm{RM}(0,m)$ — единственный бит потворён 2^m раз.

Очевидно, $w(\underbrace{{\tt 11}...{\tt 1}})=2^m=2^{m-0}\geq 2^{m-r}.$

Гипотеза: Если $v\in \mathrm{RM}(r-1,m-1)$, то $w(v)\geq 2^{m-r}.$

$$\begin{split} w(c) &= w((u \mid u \oplus v)) \overset{(1)}{=} w(u) + w(u \oplus v) \geq \\ &\overset{(2)}{\geq} w(u) + (w(v) - w(u)) = w(v) \overset{IH}{\geq} 2^{m-r} \blacksquare \end{split}$$

2022-02-

Код Рида-Маллера

-Свойства и параметры кода

—Минимальное расстояние

—Минимальное расстояние

1. Случай $\mathrm{RM}(0,m)$ мы разбирали раньше, но я напомню. Здесь длина сообщения равна $k=\sum_{i=0}^r C_m^i=C_m^0=1$, а длина кода $n=2^m$. Причём мы просто берём один бит (соответсвует функции $f(x_1,...,x_m)=0$ или $f(x_1,...,x_m)=1$) и повторяем его 2^m раз (в таблице истинности).

Замечу, что не рассматриваю второй случай $w(\mathtt{00...0})$, поскольку он нам не нужен для расчёта минимального расстояния. Вариант с нулевым вектором явно выкидывается, см. определение d выше.

2. Теперь немного объяснений.

Переход (1): $w((x \mid y)) = w(x) + w(y)$. Вес это всего лишь число ненулевых элементов, поэтому нет разницы как мы будем группировать части вектора.

Переход (2): $w(u\oplus v)\geq w(v)-w(u).$ Если у нас в v стоит w(v) бит, то прибавив к нему u, мы сможем изменить (обнулить) не больше w(u) бит. Возможно появится больше единиц, но нас интересует нижняя граница.

Переход (ІН): предположение индукции в чистом виде.



Свойства и параметры

 $r \leq m$

Для бинарного кода RM(r, m):

- \blacksquare Длина кода: 2^m
- lacksquare Длина сообщения: $k = \sum_{i=0}^r C_m^i$
- Минимальное расстояние: $d=2^{m-r}$
- \blacksquare Корректирующая способность: $t=2^{m-r-1}-1$
- lacktriangle Существует порождающая матрица G для кодирования
- lacktriangle Проверочная матрица H совпадает с порождающей для RM(m-r-1,m)

Код Рида-Маллера Свойства и параметры кода 2022-02-13

Свойства и параметры

- 1. Теперь можно подвести итоги исследования свойств. 2. , поскольку $t=\left\lfloor\frac{d-1}{2}\right\rfloor=\left\lfloor\frac{2^{m-r}}{2}-\frac{1}{2}\right\rfloor=\left\lfloor2^{m-r-1}-0.5\right\rfloor=2^{m-r-1}-1$ 3. , она позволяет делать так: C(x)=xG. Но я, как обычно, её избегаю. Рекомендую почитать «Коды Рида-Маллера: Примеры исправления ошибок», если интересно.
- 4. , но это я это доказывать не собираюсь. Но его можно найти в «Reed-Muller Codes: Theory and Algorithms», раздел Duality.



Как линейный код

Этот код является линейным кодом, к нему применимы все обычные (и неэффективные методы):

- Перебор по всему пространству кодовых слов в поисках ближайшего.
- С использованием синдромов: $s = rH^{T}$.

Код Рида-Маллера 2022-02-13 ∟Как линейный код 1. Этот способ применим ко всем кодам, но никто в здравом уме им не 2. Здесь s — синдром, r — полученное сообщение, H — проверочная матрица. Этот метод обычен для линейных кодов. 3. Эти способы нужно иметь ввиду, но о них было рассказано и без меня, так что я их пропущу.



Синдромы и как их использовать

Пусть у нас в полученном сообщении r есть ошибка e. Тогда r = v + e, где v — кодовое слово, которое крайне легко можно декодировать. Получается, что $s=rH^T=(v+e)H^T=vH^T+eH^T=eH^T$, поскольку $vH^T=0$ (есть такое свойство). Мы можем перебрать всевозможные ошибки (e), для каждой посчитать синдром и записать всё это в таблицу. Тогда чтобы восстановить сообщение, нужно посчитать синдром, по таблице найти ошибку и исправить её.

Код Рида-Маллера

2022-02-13

Синдромы и как их использовать

- 1. Я не стал включать это в презентацию, но вообще-то говоря метод полезный, так что пусть будет здесь.
- 2. Источник: https://ru.wikipedia.org/wiki/Линейный_код



Пример

Ранее: 101 кодируется как 1100 при помощи ${
m RM}(1,2)$

- t=1
- t = 0

Код Рида-Маллера 2022-02-13 -Алгоритм Рида □Пример

- 1. Теперь начинаем нормальный алгоритм декодирования, придуманный Ридом (тем самым). Именно из-за алгоритма декодирования Рида включили в соавторы кода Рида-Маллера. 2. (см. самый первый пример).

TODO



- https://arxiv.org/pdf/2002.03317.pdf великолепный обзор, очень рекомендую.
- http://dha.spb.ru/PDF/ReedMullerExamples.pdf очень хорошо и подробно, но используется подход через матрицы, а не через полиномы, а это не весело.
- 13 https://en.wikipedia.org/wiki/Reed-Muller_code кратко, чётко, понятно, но не описано декодирование.
- 4 https://ru.bmstu.wiki/Коды_Рида-Маллера в целом всё есть, но написано очень непонятно;

Код Рида-Маллера ∟_{Источники}

2022-02-13

- 1. Бонусный раздел, который не включён в основную презентацию, но может быть очень полезен.