

Реализация Layered Schedule для Information Bottleneck *

Мельников Игнат^{1,2}, Крещук Алексей^{1,3}

¹ Институт Проблем Передачи Информации Российской Академии наук

² Московский Физико-Технический Институт (Национальный
Исследовательский Университет)

³ Высшая Школа Экономики (Национальный Исследовательский Университет)

Аннотация Для проектирования протоматриц МПП-кодов (малой плотности проверок), используется Дискретная эволюция плотностей (Discrete Density Evolution). Также DDE позволяет оптимизировать выбор декодера. Наиболее гибким (и современным) подходом к оптимизации декодера является Information Bottleneck (IB). Для более эффективного декодирования используется Layered Schedule, но такой метод для IB требует реализовать обратную операцию.

1 Введение

Во многих современных стандартах систем связи можно найти протоматрицы МПП-кодов, исправляющих ошибки, вносимые каналом. Такие коды просто кодируются и декодируются, а самая сложная задача — их проектирование, то есть выбор хорошей протоматрицы. На практике для этого часто используют Дискретную эволюцию плотностей (Discrete Density Evolution) [1], которая предсказывает распределения случайных величин на выходе декодера, а значит и вероятность ошибочного декодирования.

Но Дискретная эволюция плотностей позволяет оптимизировать не только выбор матрицы, но и выбор декодера. Наиболее гибким (и современным) подходом к оптимизации декодера является Information Bottleneck. Information Bottleneck (IB) - это мощный универсальный инструмент из области машинного обучения, основная цель которого сохранить как можно больше релевантной информации, содержащейся в наблюдении за случайной величиной, когда применяется сжатие этой переменной с потерями. X – пространство входов в канала, Y – пространство выходов из канала, в дополнение к ним вводится третья переменная T , которая является сжатым представлением Y . Переменные $X \rightarrow Y \rightarrow T$ образуют цепь Маркова. Ключевая идея заключается в том, чтобы сжать $Y \rightarrow T$ и, таким образом, минимизировать $I(Y; T)$, чтобы сохранить максимально возможную взаимную информацию $I(T; X)$ [2]. Полученная задача оптимизации также называется вариационным принципом IB и может быть выражена, как минимизация

* Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ

функционала \mathcal{L} :

$$\mathcal{L}[p(t|y)] = I(T; Y) - \beta I(T; X). \quad (1)$$

Применяемый метод оптимизации - это оптимизация Лагранжа. Следовательно β обозначает множитель Лагранжа связанный с информационным ограничением на $I(T; X)$. Крайний случай $\beta = 0$ указывает на максимальное сжатие, а $\beta \rightarrow \infty$ означает максимальное сохранение релевантной информации. Оптимизация же декодера происходит разбиением операции проверочного узла на серию частичных операций. Принцип работы рассматриваемого узла проиллюстрирован на рисунке 1. Каждый черный квадрат соответствует частичной операции, которая принимает два целых числа и выводит одно целое число.

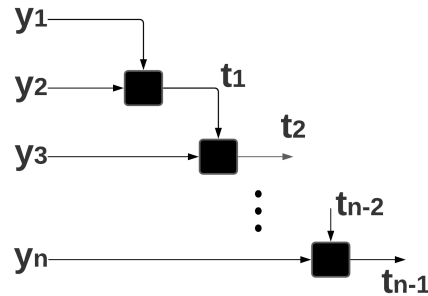


Рис. 1: Иллюстрация разделения работы контрольного узла.

В кодах с малой четностью проверок (англ. low-density parity-check (LDPC))[1] декодеры реализованные на базе ИВ обрабатывают только целые числа без знака и используют только простые таблицы поиска в качестве узловых операций, что делает их гораздо проще и быстрее [2] [3].

Обычно в каждой итерации декодирования все узлы переменных и впоследствии все проверочные узлы отправляют сообщения всем своим соседям (flooding schedule). Альтернативный, более эффективный подход заключается в последовательном обновлении сообщений проверочных узлов (Layered Schedule)[4][5], что позволяет экономить существенную часть памяти, так как имеет смысл хранить информацию только о входах и выходах определенного узла. Однако такой подход требует реализовать обратную операцию для ИВ.

2 Обзор модели

В данной работе рассматривается выход Y из AWGN-канала (Additive white Gaussian noise), помимо того он принимается симметричным относительно

передаваемого бита. Для максимального сохранения релевантной информации выберем довольно большой β из формулы (1). Для задачи оптимизации будем использовать Modified Sequential Information Bottleneck Algorithm [1], который при $\beta \rightarrow \infty$ сводится к минимизации разности взаимных информаций:

$$I(X; Y) - I(X; T). \quad (2)$$

3 Разработка алгоритма

Для построения обратных таблиц будем также минимизировать функционал \mathcal{L} (1), что сводится к оптимизации (2). Далее вычислим следующие взаимные информаций:

$$I(z_2 = f^{-1}(t, y_1); x = b_2) = \sum_{z_2, b_2} P(z_2, b_2) \log_2 \frac{P(z_2, b_2)}{P(z_2)P(b_2)}, \quad (3)$$

где

$$P(z_2, b_2) = \sum_{y_1, y_2: z_2 = f^{-1}(f(y_1, y_2), y_1)} P(y_1)P(y_2, b_2).$$

y_1 и y_2 – входы прямого преобразования и выходы AWGN-канала; z_2 – выход после обратного преобразования (рис. 2); b_2 – значение второго бита.

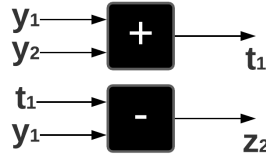


Рис. 2: y_1 и y_2 – входы прямого преобразования и выходы AWGN-канала; z_2 – выход после обратного преобразования.

Взаимная информация (3) покажет сколько информации сохранилось после прохода через канал, а также после прямого и обратного преобразования. Сравнение с взаимной информацией после прохода только через канал покажет эффективность такого метода:

$$I(y_2; x = b_2) = \sum_{y_2, b_2} P(y_2, b_2) \log_2 \frac{P(y_2, b_2)}{P(y_2)P(b_2)} \quad (4)$$

Второй проверкой эффективности станет рассмотрение взаимной информации оставшейся о первом бите b_1 после обратного преобразования, :

$$I(z_2 = f^{-1}(t, y_1); x = b_1) = \sum_{z_2, b_1} P(z_2, b_1) \log_2 \frac{P(z_2, b_1)}{P(z_2)p(b_1)}, \quad (5)$$

где

$$P(z_2, b_1) = \sum_{y_1, y_2: z_2 = f^{-1}(f(y_1, y_2), y_1)} P(y_1)P(y_1, b_1).$$

Если численно значения (3) и (4) близки к другу другу это означает, что информация о сохраняемом бите в процессе преобразований почти не меняется. В свою очередь близость (5) к нулю означает качество удаления информации о вычитаемом бите.

4 Результаты

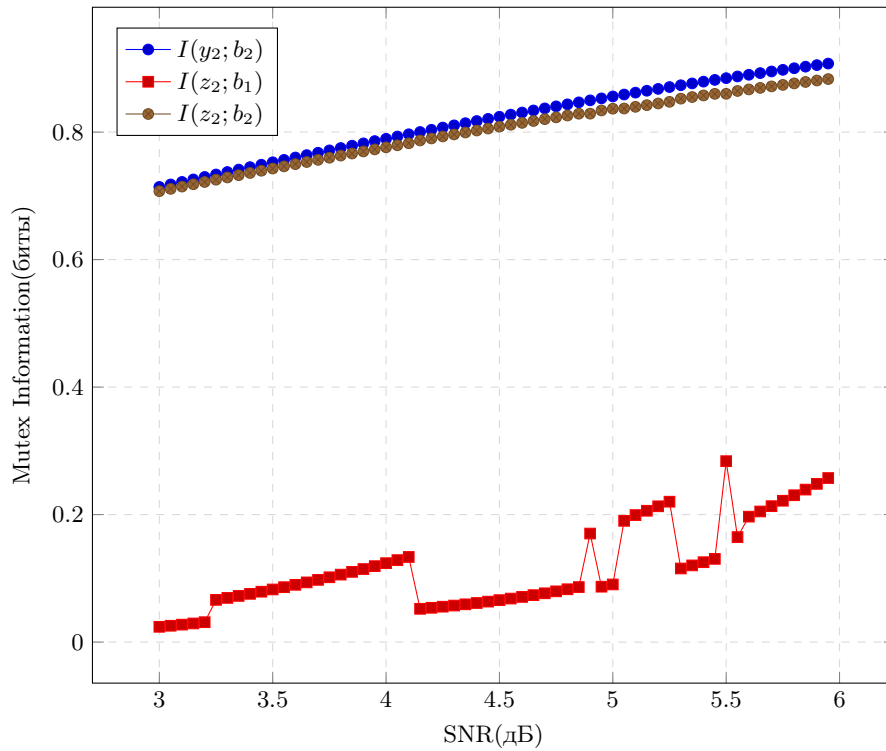


Рис. 3: Поведение взаимных информации. $I(y_2; b_2)$ – взаимная информация после прохода только через канал; $I(z_2; b_2)$ – взаимная информация после прохода через канал, а также после прямого и обратного преобразования; $I(z_2; b_1)$ – взаимная информация оставшаяся о первом бите после обратного преобразования.

Графики $I(y_2; b_2)$ и $I(z_2; b_2)$ на рис. 3, взаимной информации после прохода только через канал и после прохода через канал, прямого и обратного преобразования соответственно, находятся очень близко друг к другу, что говорит о сохранении релевантной информации. Лишь при большом отношении сигнал/шум (ОСШ; англ. signal-to-noise ratio, сокр. SNR) наблюдается их расхождение.

График же остаточной информации $I(z_2; b_1)$ на рис. 3, взаимной информации оставшаяся о вычитаемом бите после обратного преобразования, в начале достаточно близок к нулю, однако с увеличением ОСШ быстро растет. После каждой итерации уровень шума фактически уменьшается, поэтому на первых итерациях предложенный алгоритм должен работать хорошо, а на последних будет оставаться слишком много лишней информации и точность декодера будет падать.

5 Вывод

Discrete Density Evolution на основе Information Bottleneck и Layered Schedule по отдельности уже показали свою эффективность. Реализация обратной операции может позволить сэкономить довольно много памяти для декодера, что существенно понизит его цену.

Предложенный в данной работе метод хорошо себя показывает в качестве сохранения прошедшей информации. Однако алгоритм хорошо работает только на первых итерациях. Количество лишней информации растет вместе с количеством итераций, что уменьшает эффективность алгоритма на последних повторениях.

Список литературы

- [1] В. М. Kurkoski, К. Yamaguchi и К. Kobayashi, «Noise thresholds for discrete LDPC decoding mappings,» в *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, New Orleans, LA, USA: IEEE, 2008.
- [2] J. Lewandowsky и G. Bauch, «Information-optimum LDPC decoders based on the information bottleneck method,» *IEEE Access*, т. 6, с. 4054–4071, 2018.
- [3] M. Stark, J. Lewandowsky и G. Bauch, «Information-bottleneck decoding of high-rate irregular LDPC codes for optical communication using message alignment,» en, *Appl. Sci. (Basel)*, т. 8, № 10, с. 1884, окт. 2018.
- [4] E. Sharon, S. Litsyn и J. Goldberger, «Efficient serial message-passing schedules for LDPC decoding,» *IEEE Trans. Inf. Theory*, т. 53, № 11, с. 4076–4091, нояб. 2007.
- [5] D. E. Hocevar, «A reduced complexity decoder architecture via layered decoding of LDPC codes,» в *IEEE Workshop on Signal Processing Systems, 2004. SIPS 2004*, Austin, Texas, USA: IEEE, 2004.