Применение идей фонтанного кодирования при передаче малого числа битовых пакетов

Иванов Е.Р.

10 октября 2024 г.

Аннотация. В статье изучена возможность применения идей фонтанного кодирования при передаче малого числа битовых пакетов: предложена постановка задачи, допускающая использование стохастического подхода; предложены методы ее решения; проведены соответствующие эксперименты и сделаны выводы о применимости подхода.

Ключевые слова: теория кодирования, фонтанный код

1 Введение

В конце прошлого века стали активно развиваться технологии, способные передавать большое количество информации, например, телевидение и интернет. Их отличительной чертой также стала не слишком высокая цена ошибки декодирования. Именно эти две особенности процесса передачи информации оказались решающими в вопросе применимости хорошо изученных алгебраических кодов.

Для решения многих задач в новых условиях было бы достаточно, чтобы ошибки декодирования происходили «не слишком часто» (например, утеря или искажения одного из кадров видеоряда), а размеры передаваемых битовых последовательностей были достаточно большими. В качестве решения в начале текущего столетия были предложены фонтанные коды [1, 2]. В таком случае источник работает как разбрызгивающий пакеты цифровой фонтан, а для декодирования получателю достаточно лишь собрать в «ведро» определенное количество любых «капель».

2 Постановка задачи

В качестве модели канала связи использовался BEC (Binary Erasure Channel) [2]. Вероятность утери пакета далее обозначается τ.

Была поставлена следующая задача: используя подходы и идеи фонтанного кодирования, предложить систематический код, снижающий уровень ошибок в канале до наперед заданного числа π (хорошим результатом будем считать улучшение на порядок), используя допустимую избыточность не более 20%. Дополнительным ограничением на течение процесса является то, что количество пакетов, высылаемых при одних и тех же параметрах, невелико.

Систематичность кода кажется противоречивым ограничением для фонтанных кодов, которые были спроектированы независящими от порядка посылки символов. Результат, полученный М. Лаби в [3], сформулирован в предельной форме, в ходе доказательства встречаются переходы, верные при достаточно больших значениях параметров, поэтому на практике LT-коды оказываются полезными при значениях k порядка тысяч, что делает их неприменимыми в поставленной задаче. Однако, можно ослабить одно из преимуществ фонтанных кодов, заключающееся в независимости результата от пропускной способности канала связи. Сделаем это, дополнительно предположив, что канал достаточно надежный ($\tau \leq 7\,\%$), допустим это как априорное знание.

3 Изученный подход

Имеющимися степенями свободы являются алгоритм декодирования и распределение, по которому высылаются избыточные пакеты. Возможные различные стратегии распределения сложности между ними. Мы остановились на следующем варианте: использование сложного алгоритма декодирования – LT-процесса [3] – и малопараметрического распределения.

Выбор возможных распределений степеней кодирующих символов осуществлялся исходя из следующего наблюдения. Удачные распределение идейно выглядят следующим образом: чаще всего генерируются символы одной и той же степени, которые постепенно расходуются для поддержания размера очереди. Перечислим выбранные распределения:

- Биномиальное с параметрами k и p
- Пуассоновское с параметром λ
- Степень d с вероятностью 1

Сделаем несколько замечаний об использовании упомянутых распределений. В области значений случайной величины с биномиальным распределением есть ноль, что соответствует кодирующему символу нулевой степени; с этим недостатком можно, например, бороться с помощью исключения этого значения и последующей перенормировки, однако в данной работе это не применялось. Область значений случайной величины с пуассоновским распределением составляет весь натуральный ряд и ноль; для борьбы с возможным символом степени ноль к реализации случайной величины прибавлялась единица, а в случае равенства значению большего k она приравнивалась k.

4 Эксперименты

Каждый эксперимент однозначно определяется следующими параметрами: n – число посылаемых кодирующих символов, k – число пакетов, τ – процент потерь в канале и метод кодирования.

В качестве оценки $\hat{\pi}$ для величины π – достигнутой пропускной способности канала – использовалось выборочное математическое ожидание:

$$\hat{\pi} = \frac{1}{N_{\text{exps}}} \sum_{j=1}^{N_{\text{exps}}} \left[1 - \frac{N^j_{\text{recovered}}}{k} \right],$$

где N_{exps} – число проведенных экспериментов, $N^j_{\text{recovered}}$ – число восстановленных пакетов в j реализации.

n	k	Метод	$\hat{\pi}$
12	10	b	0,0123
24	20	d	0,0083
36	30	d	0,0058
48	40	d	0,0041
60	50	d	0,0030
72	60	d	0,0024

Таблица 1: Лучшие результаты методов; $\tau = 0{,}05$; избыточность 20%

n	k	Метод	$\hat{\pi}$
12	10	d	0,0221
24	20	d	0,0173
36	30	b	0,0132
48	40	d	0,0100
60	50	d	0,0087
72	60	d	0,0069

Таблица 2: Лучшие результаты методов; $\tau = 0{,}07$; избыточность 20%

Результаты экспериментов приведены на рисунках 1 и в таблицах 1-2. По оси абсцисс отложено математическое ожидание степени каждого кодирующего символа, по оси ординат – вычисленная оценка $\hat{\pi}$. Запись b в таблице означает, что лучшим оказалось биномиальное распределение, d – вырожденное.

5 Выводы

Опишем сделанные из экспериментов выводы:

- для всех распределений верно, что существует единственное значение параметра, при котором оно показывает наилучший результат;
- наиболее удачным оказался опыт использования вырожденного распределения (кодирующий символ имеет степень d с вероятностью 1); оказалось, что вносимая другими распределениями неопределенность в условиях малого количества пакетов скорее мешает, чем помогает;
- до максимума значения растут быстрее, чем падают после; это объяснимо тем, что при малой средней степени генерируемые пакеты скорее

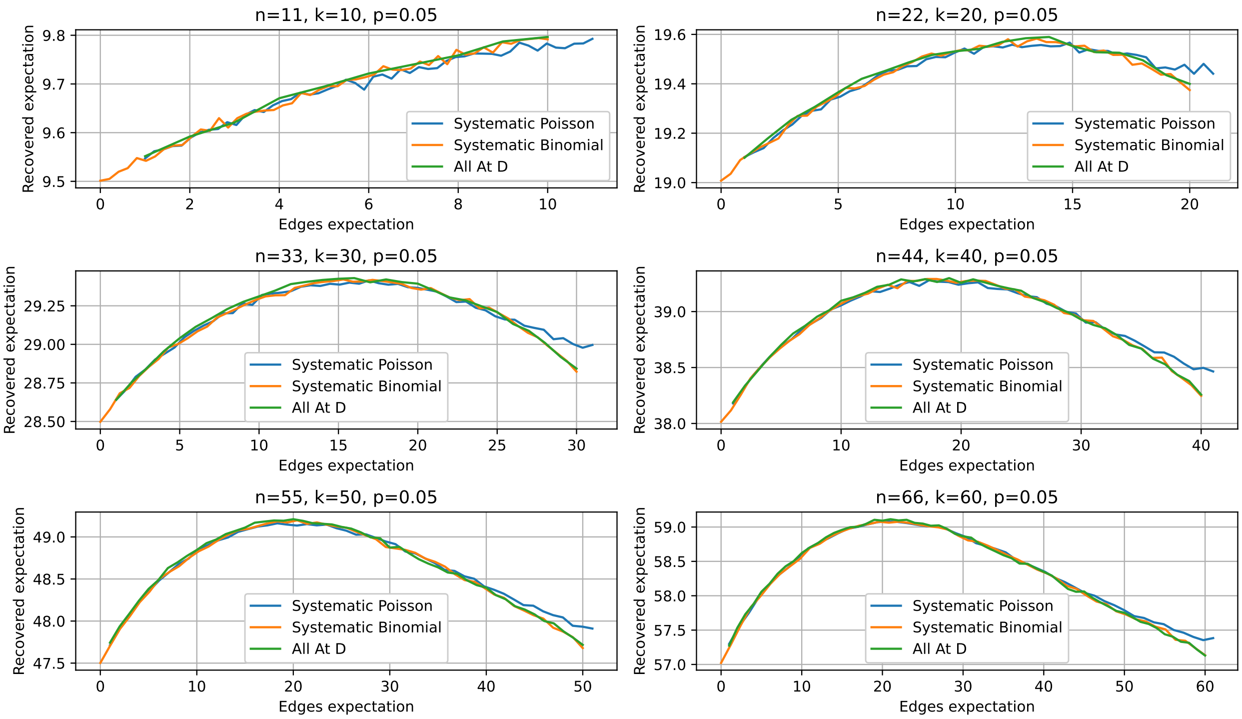

Рис. 1: $n = 1{,}1k; \tau = 0{,}05\%$

неинформативны, так как с большой вероятностью полностью «покрываются» дошедшими систематическими;

- при больших средних значениях степеней лучше работает пуассоновское распределение; это объяснимо тем, что оно обладает наибольшей дисперсией, а значит при больших средних значениях степеней кодирующие символы меньших степеней будут генеироваться чаще, чем в других распределениях.

Литература

1. MacKay, D. J. C. Fountain codes. *IEE Proceedings - Communications* **152,** 1062—1068. ISSN: 1359-7019 (дек. 2005).
2. Kythe, D. K. & Kythe, P. K. *Algebraic and Stochastic Coding Theory* ISBN: 978-1-46650562-9 (CRC Press, Boca Raton, FL, USA, 2017).
3. Luby, M. в *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.* 19 (IEEE). ISBN: 978-0-7695-1822.