

Сергей Пищулов. Работа №3

[Ссылка на архиватор](#)

Кодирование:

1. BWT с суффиксным массивом за  $O(n)$
2. MTF1
3. RLE0 возвращает сжатый список чисел и список счетчиков нулей
4. Сжатый RLE список кодируем динамическим Хаффманом
5. Список счетчиков нулей кодируем гамма-кодом
6. Результат - конкатенация
  - a. 128 бит - запись длины кода Хаффмана
  - b. Код Хаффмана
  - c. Гамма-код списка счетчиков

Сравнение результатов при использовании разных кодов

С гамма-кодом коэффициент сжатия = 0.29

С дельта-кодом коэффициент сжатия = 0.32

Схема с переполнением

80% выборки не превосходит 4. Максимальное кодируемое значение = 2200

Возьмем схему 2+4+16. Коэффициент сжатия = 0.311

Запуск

```
compressor.exe story.txt story.bin --encode
```

```
compressor.exe story.bin story_uncompressed.txt --decode
```

Файл Jack London - The Sea-Wolf размера 565 кБ удалось сжать до 166 кБ. Коэффициент сжатия - 0.29