

# Теория кодирования и сжатия информации

## Лабораторная работа №9

Гущин Андрей, 431 группа, 1 подгруппа

2022 г.

### 1 Сравнение коэффициента сжатия данных

Исходный размер, байт	Коэффициент сжатия данных							
	Код Хаффмана	Код Фано	Алг. Шеннона	Алп. код Хаффмана	Стопка книг	Алг. Гилберта-Мура	LZ77	LZ78
2	0.28571	0.28571	0.15385	0.5	0.4	0.15385	0.25	0.33333
33	0.3587	0.3587	0.17188	0.6	0.45205	0.16837	0.25	0.33333
2739	1.37087	1.34993	1.23267	1.43403	0.87873	1.06825	15.5625	2.23227
330	6.875	6.875	6.11111	5.07692	7.17391	3.47368	20.625	3.05556
59	0.35119	0.34911	0.17251	0.5619	0.4958	0.16905	0.25	0.33333
7958	1.01466	0.98102	1.25481	1.4075	1.30096	1.08464	1.07193	1.03579
138245	1.14738	1.12422	1.51266	1.61462	1.32241	1.27212	1.31687	1.31643
574426	1.11119	1.12263	1.57949	1.69144	1.41842	1.31906	1.38212	1.41673
2752	7.88539	7.88539	7.81818	7.95376	7.90805	7.81818	57.33333	12.3964
2814	7.64674	7.64674	6.76442	7.64674	7.62602	3.66406	50.25	8.01709

### 2 Сравнение скорости сжатия данных

Исходный размер, байт	Скорость сжатия данных, Кб/Сек							
	Код Хаффмана	Код Фано	Алг. Шеннона	Алп. код Хаффмана	Стопка книг	Алг. Гилберта-Мура	LZ77	LZ78
2	0.603018	0.715146	0.839344	0.061427	0.360548	0.431158	0.402022	0.381485
33	12.11943	14.134477	14.08	0.95642	13.245272	12.395048	11.545913	13.466972
2739	1115.8687	960.11501	982.9969	14.834143	932.34804	1053.4219	852.1148	987.4092
330	129.533301	139.46347	137.89838	9.618171	120.24553	140.288531	138.8474	136.0386
59	21.931573	21.791163	24.00079	1.522494	23.908192	23.849206	23.8374	23.762439
7958	2439.2702	2572.4858	2591.5064	8.141355	2292.5846	2635.5084	356.2595	629.98334
138245	16610.0	16047.029	14122.042	6.929059	7049.944	13377.705	177.1079	162.4453
574426	20524.878	17966.3016	18472.2615	9.0	9697.55	17911.865	170.0799	154.0666
2752	985.848522	1037.5729	911.10507	68.337438	1006.53558	821.1095	738.9178	789.0928
2814	1164.8452	1175.8971	1145.8538	63.616386	1141.20237	1044.0347	306.5708	1086.552

### 3 Выводы

Все алгоритмы увеличивают размер файла при сжатии за счёт метаданных, если сам файл достаточно малого размера.

Можно заметить, что алгоритмы, основанные на создании кода по частотным данным символов в файле, показывают лучший результат на файлах, которые в основном состоят из одного очень часто встречающегося файла. Хуже всего они себя показывают на файлах с равномерным или близком к равномерному распределению символов.

Алгоритмы семейства LZ\* показывают значительно лучшие результаты, чем частотные алгоритмы сжатия. Как и частотные алгоритмы, они показывают себя лучше всего, если в файле преобладает какой-либо один символ, либо некоторый повторяющийся набор символов. При этом кодирование происходит значительно медленнее остальных алгоритмов.