

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Сибирский государственный университет телекоммуникаций и
информатики»
(СибГУТИ)

Кафедра прикладной математики и кибернетики

Теория информации
Лабораторная работа №4

Выполнил: студент гр. ИП-711

Мартасов Илья Олегович

Проверила:

Доцент кафедры ПМиК

Мачикина Елена Павловна

Новосибирск 2021 г.

Содержание

Расчет теоретических значений.....	3
Результаты работы	4
Вывод.....	5

Расчет теоретических значений

Алфавит источника – 2 символа (0 и 1):

$$P(\text{символ алфавита}) = 1/2$$

$$H = (-1/2 * \log_2[1/2]) * 2 = 1$$

Результаты работы

Метод кодирования	Название текста	Оценка избыточности кодирования	Оценка энтропии выходной посл-ти (частоты символов)	Оценка энтропии выходной посл-ти (частоты пар символов)	Оценка энтропии выходной посл-ти (частоты троек символов)
Метод Шеннона	Война и Мир. Том 1	0.468611	0.995808	0.995671	0.995486
	И свет во тьме светит	0.497425	0.992384	0.99229	0.992204
	Утро помещика	0.465188	0.995013	0.994554	0.993836
Метод Фано	Война и Мир. Том 1	0.0584472	0.99438	0.994379	0.99395
	И свет во тьме светит	0.102073	0.9882	0.987874	0.987482
	Утро помещика	0.0652784	0.994117	0.994117	0.994071
Метод Хаффмана	Война и Мир. Том 1	0.0390473	0.994831	0.994827	0.994722
	И свет во тьме светит	0.0379084	0.994533	0.994309	0.994185
	Утро помещика	0.0357293	0.995563	0.995521	0.995442
Метод Гилберта-Мура	Война и Мир. Том 1	1.39917	0.994902	0.994053	0.991262
	И свет во тьме светит	1.42236	0.994038	0.992992	0.99178
	Утро помещика	1.39613	0.994584	0.994152	0.992938

Вывод

При анализе текстов, закодированных разными кодами, получилось, что самым малоизбыточным кодом является код Хаффмана, а оценки энтропий для каждого метода практически близки к максимально возможной энтропии. Это связано с тем, что после сжатия текста передается по сути та же информация, но каждый символ в закодированной последовательности несет больше информации, чем каждый символ в исходной.