

Университет ИТМО, факультет программной инженерии и компьютерной техники
Двухнедельная отчётная работа по «Информатике»: аннотация к статье

Дата прошедшей лекции	Номер прошедшей лекции	Название статьи/главы книги/видеолекции	Дата публикации (не старше 2021 года)	Размер статьи (от 400 слов)	Дата сдачи
11.09.2024	1	АНАЛИЗ ОРГАНИЗАЦИИ ЗАЩИТЫ ИНФОРМАЦИИ В СЕТЯХ 4G/LTE	2022	762	25.09.2024
25.09.2024	2	СЖАТИЕ ДАННЫХ. АЛГОРИТМ ХАФФМАНА	2022	2477	09.10.2024
	3				
	4				
	5				
	6				
	7				

Выполнил(а) Софьин Вячеслав Евгеньевич, № группы P3107, оценка
Фамилия И.О. студента не заполнять

Прямая полная ссылка на источник или сокращённая ссылка (bit.ly, tr.im и т.п.)

<https://cyberleninka.ru/article/n/szhatie-dannyh-algoritm-haffmana/>

Теги, ключевые слова или словосочетания (минимум три слова)

Сжатие данных без потерь, алгоритм Хаффмана, теория графов

Перечень фактов, упомянутых в статье (минимум четыре пункта)

1. Метод Хаффмана неравномерный и префиксный (удовлетворяет условию Фано)
2. Частые последовательности должны кодироваться наименьшим количеством символов, редкие – большим количеством символов.
3. Сначала обработка данных для подсчёта вероятности символов, потом построение дерева Хаффмана для задания кода символам.
4. Построение дерева - сортируется (в порядке возрастания) список символов по частоте их появления, объединяем два наименее частых символа в дерево с общим родителем и это дерево снова кладем в сортированный список (частота нового дерева равна сумме частот всех символов, причём каждый раз её пересчитывать не надо, можно использовать данные прошлого расчёта), повторяем до того, пока в списке не будет одно дерево, заметим, что оно бинарное.
5. После построения дерева строятся коды символов, они равны конкатенации 0 и 1 в зависимости от пути от корня дерева до листов дерева (исходных символов), как пример, 0 левая ветка, 1 – правая.
6. В результате получаем связный список кодов символов, однозначно кодирующий и декодирующий последовательность символов.
7. Заметим, что дерево строится неоднозначно, а значит кодовое расстояние может получиться разным.

Позитивные следствия и/или достоинства описанной в статье технологии (минимум три пункта)

1. Уменьшает размер данных
2. Увеличивается скорость передачи данных (чтение и запись)
3. Уменьшение использования оперативной памяти

Негативные следствия и/или недостатки описанной в статье технологии (минимум три пункта)

1. Неэффективен для маленьких файлов (много вычислений для малого уменьшения объёма)
2. Нужно хранить дерево кодов
3. Декодер должен знать дерево кодов

Ваши замечания, пожелания преподавателю или анекдот о программистах¹

GPTClient gpt =GPTClient()

anekdot = gpt.run(“\

Расскажи смешной анекдот о программистах, запрещаю использовать упоминания о GPT”)

print(anekdot)

¹ Наличие этой графы не влияет на оценку