Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

«МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ (НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»

Институт №3 «Системы управления, информатика и электроэнергетика»

Кафедра 307 «Цифровые технологии и информационные системы»

ОТЧЁТ

О выполнении задания по дисциплине

«Программирование на языках высокого уровня»

«Отчёт о разработке хеш-функции и создании программы для хеширования строк»

Выполнил:

Студент: Барденков Т. Д.

Группы М3О-212Б-22

Подпись: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Проверил преподаватель:

Павлов О. В.

Подпись: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Москва 2023 г.

**Цель работы**

Целью работы является создание алгоритма хеширования строк, его анализ и написание программы, создающей хеш-строки по созданному алгоритму.

**Описание разработанной хеш-функции**

Первым шагом функция создает переменную длиной 8 байт, хранящую хеш-код, и присваивает ей значение, равное 2^64, деленному целочисленно на 3. Далее для каждого символа строки выполняется преобразование значения хеш-кода:

1. Значение хеш-кода сдвигается влево на количество бит, равному остатку деления хеш-кода на 3.
2. Хеш код умножается на число, соответствующее порядковому номеру символа строки в таблице ASCII.

**Блок-схема алгоритма хеш-функции**

****

**Скорость сравнения**

Ниже представлены таблица и график зависимости скорости сравнения хеш-кодов относительно длины строки, скорости сравнения строк от длины строки и количества коллизий от длины строки. Данные основаны на среднем арифметическом значении выборки из 10 значений для каждой зависимости.

Таблица 1 — Средние значения выборки

|  |  |  |  |
| --- | --- | --- | --- |
| Длина строки | Скорость сравнения строк | Скорость сравнения хеш-кодов | Количество коллизий |
| 3 | 168 | 41.6 | 370 |
| 4 | 169.7 | 45.3 | 81 |
| 5 | 171.8 | 42.1 | 11.6 |
| 6 | 173.9 | 43 | 4.3 |
| 7 | 173.3 | 42.2 | 0.7 |

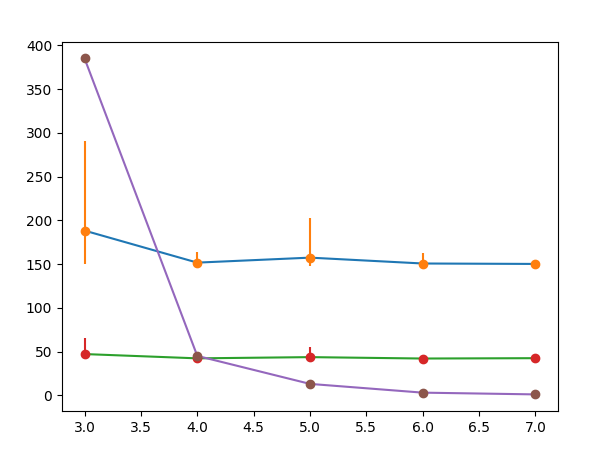


Рисунок 1 — Скорости сравнения строк, хеш-кодов и количество коллизий

Синей кривой на графике обозначены скорости сравнения строк, зеленой – хеш-кодов. Количество коллизий обозначено фиолетовой кривой.

Погрешностями на графиках являются максимальные и минимальные значения, полученные в ходе измерений.

Из полученных данных можно сделать следующие заключения:

1. Скорость сравнения хеш-кодов не изменяется по мере увеличения длины строки.
2. Скорость сравнения строк слегка увеличивается по мере увеличения длины строки.
3. Сравнение хеш-кодов значительно быстрее сравнения строк.
4. Количество коллизий уменьшается экспоненциально по мере увеличения длины строки.

**Выводы**

Мы научились писать простую хеш-функцию и нашли разницу в скорости сравнения строк и целых чисел. Полученные данные свидетельствуют о следующем:

1. Существуют одинаковые значения хеш-кодов для разных строк.
2. Количество коллизий уменьшается по мере увеличения размера входных данных.
3. Сравнения целых чисел и строк имеют разную зависимость.