Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Российский химико-технологический университет имени Д.И. Менделеева»

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №9

Выполнил студент группы КС-36: Золотухин А.А.
Ссылка на репозиторий: https://github.com/MUCTR-IKT-CPP/ZolotukhinAA_36_ALG
Принял: Крашенников Роман Сергеевич
Дата сдачи: 05.05.2025

Москва
2025

Оглавление

Описание задачи

В рамках лабораторной работы необходимо реализовать алгоритм хэширования: *SHA1*.

Дополнительный вариант: *Luffa*, *SHA3*.

Для реализованной хеш-функции провести следующие тесты:

- провести сгенерировать 1000 пар строк длинной 128 символов отличающихся друг от друга 1,2,4,8,16 символов и сравнить хеши для пар между собой, проведя поиск одинаковых последовательностей символов в хешах и подсчитав максимальную длину такой последовательности. Результаты для каждого количества отличий нанести на график, где по оси x кол-во отличий, а по оси у максимальная длинна одинаковой последовательности;
- провести $N = 10^i$(i от 2 до 6) генерацию хешей для случайно сгенерированных строк длинно 256 символов, и выполнить поиск одинаковых хешей в итоговом наборе данных, результаты привести в таблице, где первая колонка это N генераций, а вторая таблица наличие и кол-во одинаковых хешей, если такие были;
- провести по 1000 генераций хеша для строк длинной n (64, 128, 256, 512, 1024, 2048, 4096, 8192)(строки генерировать случайно для каждой серии), подсчитать среднее время и построить зависимость скорости расчета хеша от размера входных данных.

Описание метода/модели

Хеш-функция - структура данных, реализующая интерфейс ассоциативного массива.

Для такой таблицы значений, как правило, реализуются следующие операции:

- **Вставка(k, v)** - вставить в таблицу элемент с ключом;
- **Удаление(k)** - удалить из таблицы элемент с ключом;
- **Поиск(k)** - найти в таблице элемент с ключом;

При этом выделяется 2 вида таких таблиц:

- **Set** - таблица, хранящая в себе множество значений, каждое из которых само по себе является ключом и значением одновременно;
- **Map** - таблица, хранящая в себе пары ключ-значение в явном виде.

Хэширование - применение особого алгоритма к некоторым входным данным произвольного типа, преобразующего их в битовую строку установленной длины, строка часто выводится в форме шестнадцатеричного числа.

Алгоритм, который применяется для хеширования, называют **хеш-функцией**.

Результат работы алгоритма хеширования называют **хешем**

Криптографическая хеш-функция - это специальный класс хеш-функций, который имеет различные свойства, необходимые для криптографии.

Лавинный эффект - ситуация, в которой при генерации хеша для двух слабо отличающихся друг от друга изначальных набора данных, результат будет отличаться колоссально.

Соль - дополнительные данные, которые вносятся в хеш-функцию и генерируются случайным образом.

Соль может быть:

- *Статическая* - одинаковая на каждый хеш;
- *Динамическая* - разная для каждого хеша.

SHA-1 - Secure Hash Algorithm 1 (1995). Использует 160 бит для хранения результата.

SHA-3 - Secure Hash Algorithm 3.

Исходное сообщение разбивается на блоки по 512 бит в каждом. Последний блок дополняется до длины, кратной 512 бит. Сначала добавляется 1 (бит), а потом — нули, чтобы длина блока стала равной $512 - 64 = 448$ бит. В оставшиеся 64 бита записывается длина исходного сообщения в битах (в big-endian формате). Если последний блок имеет длину более 447, но менее 512 бит, то дополнение выполняется следующим образом:

сначала добавляется 1 (бит), затем — нули вплоть до конца 512-битного блока; после этого создается ещё один 512-битный блок, который заполняется вплоть до 448 бит нулями, после чего в оставшиеся 64 бита записывается длина исходного сообщения в битах (в big-endian формате). Дополнение последнего блока осуществляется всегда, даже если сообщение уже имеет нужную длину.

Выполнение задачи

SHA-1 реализован на языке *Java*. Построение графиков проводились с помощью программы *GNUplot*.

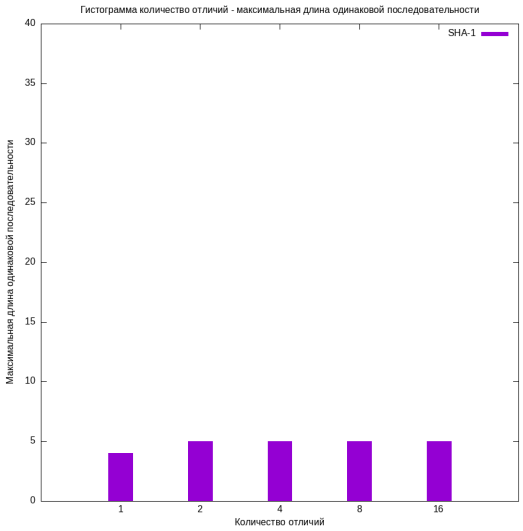

Рис. 1: Гистограмма количество отличий - максимальная длина одинаковой последовательности

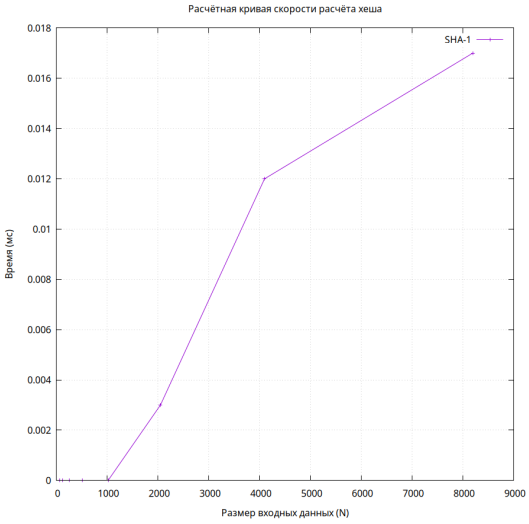

Рис. 2: Расчётная кривая скорости расчёта хеша.