Алгоритм Рабина-Карпа – это алгоритм используемый для поиска подстроки в тексте с помощью хэш-функции. Вместо старого и наивного способа поиска, алгоритм Р-К фильтрует подстроки основной строки, которые равны хэшу искомой строки.

Принцип работы:

Берется последовательность символов и проверяется на возможность наличия искомой строки. Если возможность найдена, выполняется сопоставление символов.

Время работы:  
Время для подсчета хэша O(m). В худшем случае время работы алгоритма – O(nm), в среднем O(n+m), где m – длина подстроки, а n – длина первоначальной строки.

Плюсы:

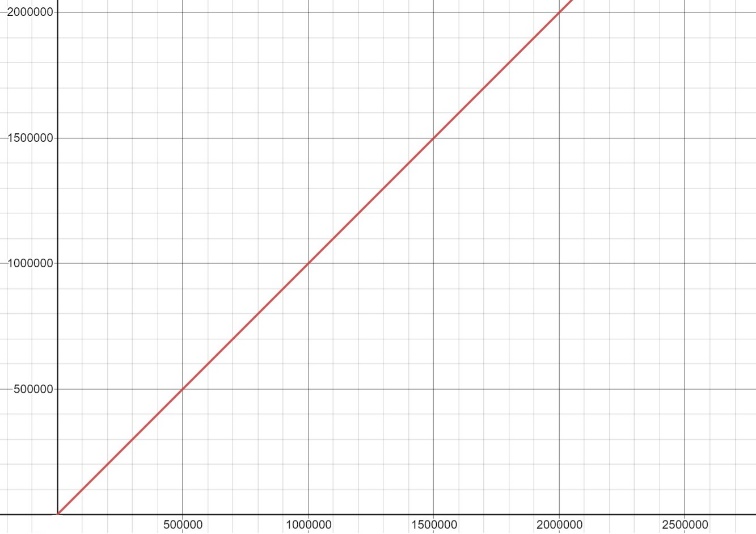
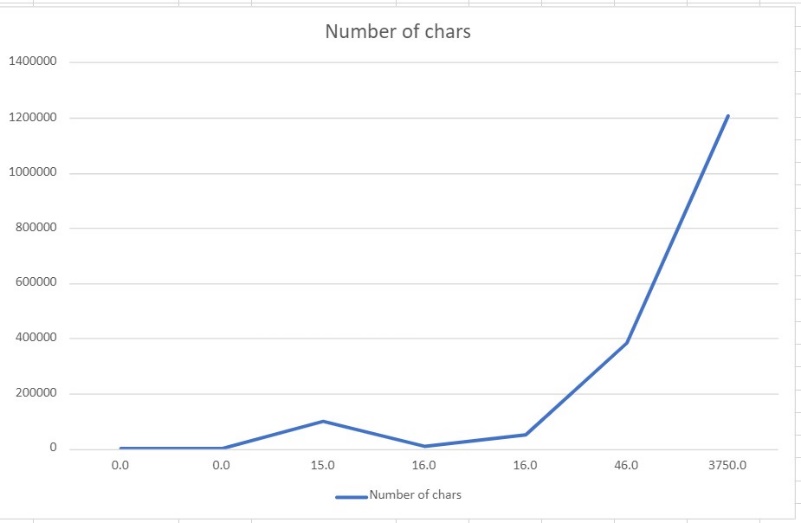
-эффективнее, чем посимвольное сравнение

Минусы:

-от сложности хэш-функции зависит точность работы алгоритма, для получение 100% результата, необходимо проверять, каждую необходимую последовательность посимвольно.

Выводы:

Алгоритм хороший, но не полностью точный, опять же из-за использования хэш-функций. Для получения максимальной точности следует использовать улучшенный алгоритм, в котором происходит посимвольная проверка найденных последовательностей.

График слева, зависимость времени от кол-ва символов. График справа, зависимость числа итераций от кол-ва символов.