Частотный анализ как один из наиболее эффективных методов вскрытия простых шифров

Грошков П. В.

Грошков Павел Викторович / Groshkov Pavel Viktorovich - студент,
кафедра информационной безопасности,
факультет микроприборов и технической кибернетики,
Национальный исследовательский университет
Московский институт электронной техники (МИЭТ), г. Зеленоград

Аннотация: *в данной статье я на примере разберу один из наиболее эффективных методов вскрытия криптографических шифров – метод частотного анализа.*
Ключевые слова: *шифрование, расшифрование, биграмма, частотный анализ.*

Еще с древних времен у людей появилась возможность тайно обмениваться информацией между друг другом при помощь шифрования текста, причем чаще всего использовались самые примитивные шифры замены, в которых каждая буква передаваемого сообщения заменялась на какой-то заведомо известный обоим переговаривающимся сторонам символ. В данной статье я попробую рассказать, почему такие алгоритмы шифрования, как обычная замена, являются легко вскрываемыми и ненадежными.

Для взлома такого вида шифров был придуман метод под названием «Частотный анализ» [2, с. 1], который сравнивает частоты встречаемости символов в открытом (исходном) и закрытом (зашифрованном) текстах. С годами учеными было выяснено, что в каждом алфавите у каждой буквы есть так называемая частота или вероятность встречаемости в тексте, так, например, в русском языке наиболее часто встречаемая буква «О». Ее частотность составляет приблизительно 11 %, и как раз на этом и основывается метод частотного анализа. Рассмотрим пример: пусть ключом шифрования является таблица 1.

	1	2	3	4	5	6
1	А	Б	В	Г	Д	Е\Ё
2	Ж	З	И	Й	К	Л
3	М	Н	О	П	Р	С
4	Т	У	Ф	Х	Ц	Ч
5	Ш	Щ	Ъ	Ы	Ь	Э
6	Ю	Я				

Таблица 1. Таблица шифрования

Каждая буква открытого текста шифруется по пересечению строки и столбца, т. е. букву А шифруем как 11, Б – 12. Зашифруем данным способом текст, который нам необходимо передать собеседнику. В примере мною был взят текст из произведения А. С. Пушкина «Сказка о рыбаке и рыбке» размером, превосходящим 1000 символов [1, стр. 1].

Посчитаем частоту встречаемости каждой буквы в тексте [рис. 1].

На гистограмме видно, что, как и говорилось ранее, наиболее часто встречаемая буква «О», за ней идет буква «Т» и так далее. Зашифруем наш текст и аналогично составим частоту встречаемости только уже не букв, а, в нашем случае, двухзначных цифр [рис. 2].

По построенным гистограммам частоты встречаемости символов в открытом и зашифрованном текстах сделаем следующие выводы:

Теоретически: чем ровнее гистограмма, тем сложнее вскрыть исходный текст. На гистограмме зашифрованного текста отчетливо просматривается самый высокий столбец, который соответствует комбинации «33». По предположению, это будет буква «О», что и действительно так. Далее столбец «41», это буква Т. Аналогично определяем все комбинации искомых букв первичного алфавита и легко расшифровываем текст.

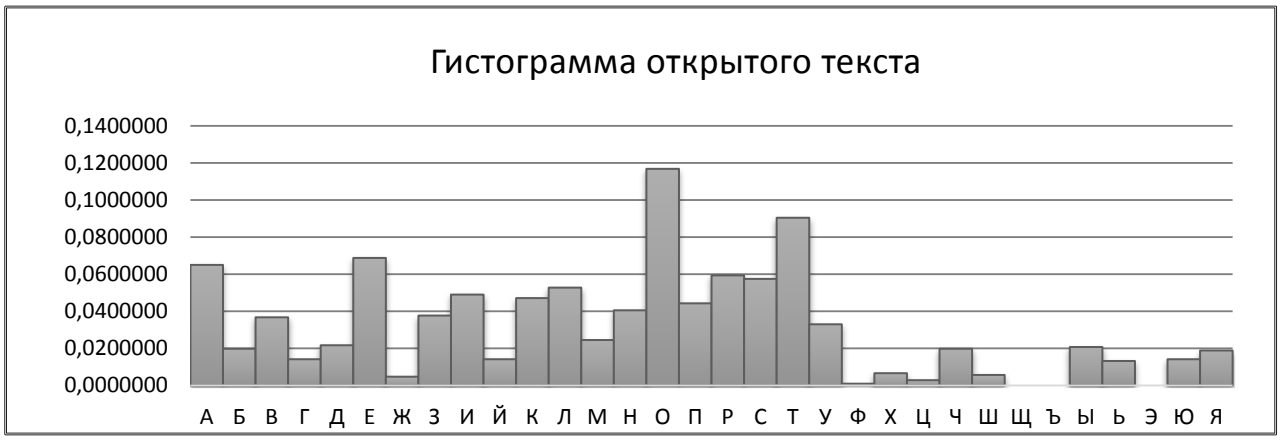

Рис. 1. Гистограмма открытого текста

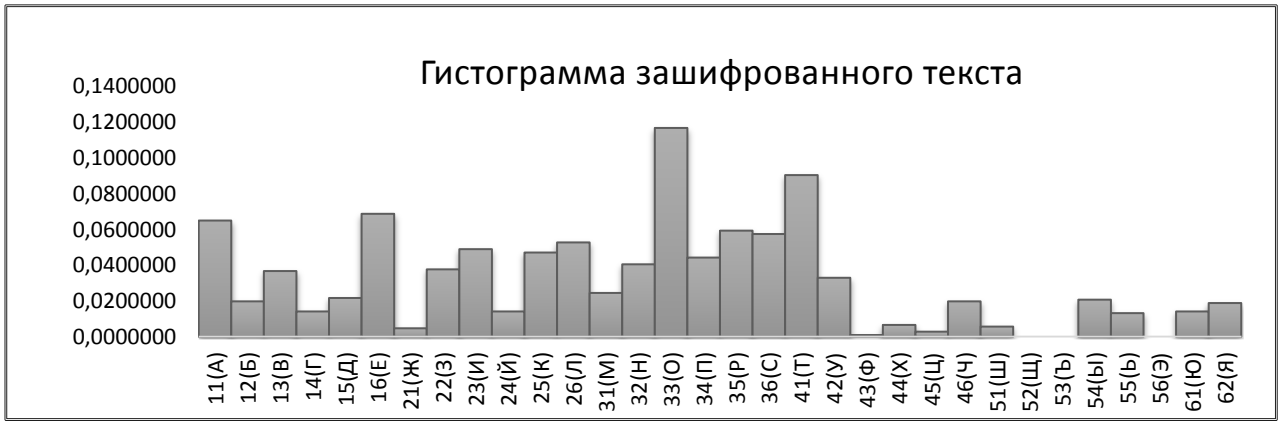

Рис. 2. Гистограмма зашифрованного текста

Подведем итог: как видно из полученного результата, метод частотного анализа является чуть ли не самым эффективным, когда дело касается вскрытия шифров, основанных на алгоритмах простой замены, гистограммы частот встречаемости символов текстов, чей объем превосходит 500 и более символов. Составленные гистограммы позволили полностью сопоставить символ исходного алфавита с соответствующим ему заменяемым символом, благодаря чему и получен столь исчерпывающий результат.

Литература

1. Текст для шифрования из книги А. С. Пушкина «Сказка о рыбаке и рыбке» [Электронный ресурс]. URL: http://rvb.ru/pushkin/01text/03fables/01fables/0799.htm.
2. Данные о частотном анализе [Электронный ресурс] URL: https://ru.wikipedia.org/wiki/%D0%A7%D0%B0%D1%81%D1%82%D0%BE%D1%82%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7.