Практика №8. Криптоанализ шифров замены.

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Шифры подстановки просты в реализации и понимании. К сожалению, их столь же легко взломать. Простой криптоанализ шифра Цезаря показывает, что если в шифре используется русский (или любой другой известный) алфавит, то все, что нужно знать для взлома, — это величина сдвига. Можно проверять каждую букву алфавита одну за другой, пока текст не будет расшифрован. Это значит, что для восстановления открытого текста в худшем случае потребуется 33 попытки (для русского языка).

Для аффинной подстановки такой метод взлома не сработает, так как разные буквы сдвигались на различную величину. Но при этом одна и таже буква сдвигалась на одно и тоже число, поэтому зашифрованный текст сохранил статистику. Используя знание о частоте появления символов, можно сделать предположения о реальном значении символа, а затем, в силу избыточности естественного языка восстановить исходный текст.

Буква	Вероят-ность	Буква	Вероят-ность	Буква	Вероят-Ность	Буква	Вероят-ность
Пробел	0.175	Р	0.040	Я	0.018	Х	0.009
О	0.090	В	0.038	Ы	0.016	Ж	0.007
Е	0.072	Л	0.035	З	0.016	Ю	0.006
А	0.062	К	0.028	Ъ	0.014	Ш	0.006
И	0.062	М	0.026	Б	0.014	Ц	0.004
Н	0.053	Д	0.025	Г	0.013	Щ	0.003
Т	0.053	П	0.023	Ч	0.012	Э	0.003
С	0.045	У	0.021	Й	0.010	Ф	0.002

ЗАДАНИЯ

Провести статистический анализ зашифрованного русского текста.

Используя данные о вероятностном распределении символов русского языка, восстановить исходный текст (или его часть).