ЗАДАНИ 7. Метод Симпсона для шифра Виженера (вычисление ключа при известной его длине):

Вход: 1) файл криптограммы $Y = y_1y_2\ldots y_N$ (шифра Виженера) над алфавитом открытых сообщений длины m.

2) Значение периода d (дина ключа).

3) Файл частот букв открытых сообщений.

Выход: список наиболее вероятных ключей или единственный ключ $k \in K$.

ЗАДАЧА:

Допустим длина ключа в шифре Виженера уже вычислена и равна d, т.е. ключ имеет вид $k = (k_1, k_2,\ldots,k_d)$. Символы алфавита отождествим с остатками по модулю m, и операции над символами будем осуществлять по модулю m.

Обозначим:

$$Y(1) = y_1, y_{1+d}, \ldots$$
$$Y(2) = y_2, y_{2+d}, \ldots$$
$$\ldots$$
$$Y(d) = y_d, y_{2d}, \ldots .$$

Последовательность $Y(2)$ зашифрована символом $k_2 = k_1 + \Delta_2$.

Для каждого $\Delta_2 \in \{0, 1, \ldots, m-1\}$ строим $Y(2-\Delta_2) = y_2 - \Delta_2, y_{2+d} - \Delta_2, \ldots$,
и вычисляем $I_{\text{ср}}(Y(1), Y(2-\Delta_2))$.

Из всех Δ_2 выбираем те, для который $I_{\text{ср}}(Y(1), Y(2-\Delta_2))$ имеет наибольшее значение. Их может оказаться несколько, обозначим их Δ_2.

Последовательность $Y(3)$ зашифрована символом $k_3 = k_1 + \Delta_3$.

Для каждого $\Delta_3 \in \{0, 1, \ldots, m-1\}$ строим $Y(3-\Delta_3) = y_3 - \Delta_3, y_{3+d} - \Delta_3, \ldots$,
и вычисляем $I_{\text{ср}}(Y(1), Y(3-\Delta_3))$.

Из всех Δ_3 выбираем те, для который $I_{\text{ср}}(Y(1), Y(3-\Delta_3))$ имеет наибольшее значение. Их может оказаться несколько, обозначим их Δ_3.

Повторяем вычисления для $Y(j)$: $j = 4, 5, \ldots, d$.

Последовательность

$$Y(1) = y_1y_{1+d},\ldots$$

зашифрована символом k_1. Подвергаем её обычному частотному анализу как в случае простого сдвига, и выделяем наиболее вероятные символы k_1 (возможно один).

Тогда $k = (k_1,\ k_1 + \Delta_2, \ldots, k_1 + \Delta_d)$ множество наиболее вероятных ключей (возможно будет состоять из одного ключа). Перебираем его и выделяем из него искомый ключ k.

ЗАДАНИЕ 8. Метод Томаса Якобсена (1995г.) (вычисление ключа при известной его длине):

1. Вычисление эталонной матрицы частот биграмм языка открытых сообщений.

Вход: файл большого текста на языке открытых сообщений.
Выход: файл матрицы $E = (E_{ij})$ частот биграмм, где E_{ij} частота биграммы ij во входном файле, $I = \{0, \ldots, m-1\}$ — алфавит, т.е. m — длина алфавита.

2. Вычисление ключа.

Вход: 1) файл криптограммы $Y = y_1 y_2 \ldots y_N$ (шифра Виженера) над алфавитом открытых сообщений.
2) Значение периода d (дина ключа).
3) Файл эталонной матрицы $E = (E_{ij})$.

ЗАДАЧА:

1. Обозначим: $k^0 = k_1 k_2 \ldots k_d$, экспериментальный ключ, полученный случайным выбором.
2. Полагаем $k = k^0$, вычисляем $k(Y) = z_1 z_2 \ldots z_N$ — текст, полученный из Y при расшифровании ключом k.
3. Вычисляем $D^k = (D_{ij}^k)$ — матрицу частот биграмм в тексте $k(Y)$.
4. Вычисляем значение целевой функции $W(k) = \sum_{ij} | D_{ij}^k - E_{ij} |$.
5. Для $k_1^* = \{0, \ldots, m-1\}$ вычисляем наименьшее $W(k^*)$, где $k^* = k_1^* k_2 \ldots k_d$, для которого $W(k^*) \le W(k)$. Пусть это значение соответствует k^*. Полагаем $k = k^*$. Для $k_2^* = \{0, \ldots, m-1\}$ вычисляем наименьшее $W(k^*)$, где $k^* = k_1 k_2^* \ldots k_d$, для которого $W(k^*) \le W(k)$. Пусть это значение соответствует k^*. Полагаем $k = k^*$. И т.д. до k_d^*. Получаем некоторый ключ k.
6. Если $k = k^0$, то стоп. Если $k \ne k^0$, то присваиваем $k^0 = k$, и повторяем пункт 5.

Выход: k^0.

https://drive.google.com/open?id=0B2AKc7ibPQ_WUDBtcHhMc0lGYlU