

Homework 1 - Viginere cipher

The code is written in Python and works as follows:

The functions for encoding and decoding a given message use a string of letters (only alphabetical) and a key. Both of these are compressed to spaceless strings and turned to uppercase.

The encryption is by the formula:

$$C_i = (M_i + K_{i \bmod |K|}) \bmod 26$$

Where:

M_i - The initial message letter

$K_{i \bmod |K|}$ - The letter of the key

C_i - The resulting ciphertext letter

The decryption uses - instead of +.

If we don't have the key, we can calculate it. The key calculation works like this:

1. We know that the Index of Coincidence - the probability to find two similar letters in an English text - is around 1.73. Using this information we can calculate the letter count for each letter in our ciphertext and calculate its IoC if a certain key length is used.
 1. The iteration is on lengths of 1-15, and for each key length (KL) we divide the ciphertext into substrings where the substring contains the letters that were encrypted by the same key letter.
That looks like stacking the cipher in layers of length KL and taking each column of letters at a time.
 2. For each substring we calculate its IoC.
 3. We take the average of all these IoC values and check how close it is to 1.73. The closest average indicates the best fit for the key length.
2. For the guessed key length we "stack" the ciphertext on itself again (the same way it is explained in 1-a) and for each of the columns - the substrings that were essentially encrypted by the same letter - we try to decode it using each letter in the English alphabet.
 1. We know for each letter how often it appears in English text.
 2. We try to calculate a chi value (X), which is the sum of [letter counts in the cipher multiplied by their respective known frequencies].
 3. A maximum value for chi achieved by the best fitting key letter.
 4. We collect all the letters and find the key - the bigger the ciphertext the more precise the key calculation will be.
3. Using the key we can now decrypt the ciphertext.

The end result is a poem named "The Crow" by Edgar Allan Poe (the key is THECROW).