Shift Cipher Report

Author: Jacob Craiglow

This cipher was easy to break as most shift ciphers are. The first step was to tokenize the cipher text by character in order to determine the most frequent unigrams, bigrams, and trigrams. After this cryptanalysis began. It quickly became annoying to flit from my program to a chart of most popular combinations of letters in the English language so I decided to add them to the program as a cheat sheet as well.

Upon performing unigram analysis we find that these are the top 15 unigrams in the cipher:

[('O', 10), ('K', 9), ('B', 7), ('C', 6), ('I', 6), ('M', 6), ('S', 6), ('V', 6), ('Z', 5), ('X', 4), ('Y', 4), ('D', 3), ('E', 3),

('R', 3), ('N', 2)]

Being that O is the most popular letter in the cipher and E is the most popular letter in the English language, it was worth adding to the partial solution. At this point I decided to look at bigrams including O in order to possibly crack other letters before trying out a key.

These are the bigrams occurring more than once in the cipher:

[('VI', 3), ('ZR', 2), ('MO', 2), ('OB', 2), ('BD', 2), ('KS', 2), ('SX', 2), ('CM', 2), ('IC', 2), ('CS', 2), ('EC', 2),

('ON', 2)]

I cross-referenced these with my cheat-sheet in order to determine if any would fit with O->E. O->E is a shift of 10. Going first MO I tried M-10 which gave me E. This is a feasible bigram, it's not in the top 15 so I decided to try another bigram. OB occurs twice as well. B->R and ER is a common bigram as well.

The partial solution at this point is:

C R - - - - - R - - - - C E R - - - - - - - - - C - - E - - - - - - - - - - C E - - - - - - C - E - - R R E - - R - E - - - - - E - - - - - - E - - - - - - - E C - - - E R

Knowing the nature of this class I decided to substitute the third letter of the cipher for Y giving:

C R Y - - - - R - - - Y C E R - - - - - Y - - - C - - E - - - - - - - Y - - - C E - - - - - - C - E - - R R E - - R - E - - Y - - E - - - - - - E - Y - - - - - E C - - - E R

At this point I was comfortable trying key = 10 and this gives:

CRYPTOGRAPHYCERTAINLYHASCOMEALONGWAYSINCEJULIUSCAESARREPORTEDLYUSEDANAIVELYSIMPLECIPHER