Kevin Bednar – CSCI4905

Cryptography Assignment 3 Report – Breaking the Vigenere Cipher

Our task for the third assignment was to break the Vigenere cipher given only the ciphertext to work with (ciphertext only attack). Written below are the procedures and methods I used in order to recover the plaintext.

**Vigenere Cipher**:

Similar to the hill cipher, the vigenere cipher is polyalphabetic – meaning that more than one character may be mapped to a single element given how the encryption is carried out. Due to this property, single letter cryptanalysis will prove useless from the get-go. The first method used to attack the vigenere cipher must include taking a stab at finding the key length. In my program I went about this using the Kasiski test and searching for groups of trigrams in the cipher. The following information was recovered:

(5, 'ZEE'), (5, 'QZE'), (5, 'HYA'), (5, 'EEK'), (5, 'DAQ'), (5, 'AQZ') – The most commonly occurring sets of characters.

ZEE found at 237

ZEE found at 317

ZEE found at 427

ZEE found at 522

ZEE found at 547

QZE found at 236

QZE found at 316

QZE found at 426

QZE found at 521

QZE found at 546

HYA found at 36

HYA found at 321

HYA found at 336

HYA found at 381

HYA found at 756

EEK found at 238

EEK found at 318

EEK found at 428

EEK found at 523

EEK found at 548

DAQ found at 234

DAQ found at 314

DAQ found at 424

DAQ found at 519

DAQ found at 544

AQZ found at 235

AQZ found at 315

AQZ found at 425

AQZ found at 520

AQZ found at 545

After finding the indexes, the space between the locations must be found. Once obtaining the differences, we can begin to work towards finding the gcd. Here are the differences:

[80, 110, 95, 25, 80, 110, 95, 25, 285, 15, 45, 375, 80, 110, 95, 25, 80, 110, 95, 25, 80, 110, 95, 25]

With a quick analysis, one is able to see the difference all end in 5’s and 10’s. From this we can gather that the key is some length times 5. The gcd of the above values ends up being 5 as well which allows us to proceed confidently guessing with a key of size 5. Once the keysize is decided upon, the cipher basically becomes a shift cipher. Since the keysize is 5, the ciphertext is put into 5 lists, each list containing characters in positions 1,2,3,4,5 in the ciphertext itself. These lists are each analyzed for singular letter frequency, similar to the process of solving the Caesar shift cipher. Here are the results gathered:

[(25, 'F'), (17, 'A'), (16, 'U'), (16, 'Q'), (12, 'M'), (10, 'X'), (10, 'K'), (10, 'E'), (8, 'Z'), (8, 'O'), (8, 'D'), (7, 'S'), (6, 'I'), (3, 'T'), (3, 'N'), (2, 'R'), (2, 'P'), (1, 'W'), (1, 'H'), (1, 'G')]

[(24, 'S'), (20, 'H'), (19, 'W'), (17, 'B'), (13, 'O'), (12, 'G'), (12, 'C'), (8, 'V'), (6, 'Z'), (6, 'Q'), (4, 'U'), (3, 'R'), (3, 'P'), (3, 'M'), (3, 'K'), (3, 'I'), (3, 'F'), (2, 'A'), (1, 'Y'), (1, 'X'), (1, 'T'), (1, 'J'), (1, 'D')]

[(18, 'K'), (18, 'F'), (17, 'Z'), (17, 'E'), (15, 'V'), (15, 'J'), (14, 'R'), (11, 'C'), (10, 'Y'), (6, 'L'), (5, 'X'), (5, 'U'), (4, 'P'), (2, 'T'), (2, 'N'), (2, 'I'), (1, 'W'), (1, 'S'), (1, 'G'), (1, 'D'), (1, 'B')]

[(27, 'T'), (20, 'E'), (14, 'S'), (14, 'N'), (13, 'A'), (11, 'L'), (10, 'O'), (9, 'I'), (7, 'Y'), (6, 'H'), (6, 'F'), (4, 'W'), (4, 'U'), (4, 'R'), (4, 'C'), (4, 'B'), (3, 'G'), (2, 'D'), (1, 'P'), (1, 'M'), (1, 'K'), (1, 'J')]

[(20, 'E'), (16, 'P'), (15, 'T'), (15, 'D'), (13, 'Y'), (11, 'Z'), (11, 'L'), (9, 'S'), (8, 'N'), (8, 'C'), (6, 'W'), (5, 'R'), (5, 'H'), (4, 'O'), (4, 'F'), (3, 'X'), (3, 'J'), (3, 'A'), (2, 'M'), (2, 'I'), (1, 'V'), (1, 'Q')]

Assuming the most frequent letter in each list is E gives the key: BOGPA – which contains only one correct letter. Left in the program was the method for making the ‘crappy guesses’ for obtaining the BOGPA key. Since the key is wrote (after verification) I developed a method which attempts mapping the most frequent list letters (F,S,K,T, and E) and gives each of the top 4 possible mappings of E, T, A, and O. The following information is now given:

Possible values for list:

B 1

M 12

F 5

R 17

Possible values for list:

O 14

Z 25

S 18

E 4

Possible values for list:

G 6

R 17

K 10

W 22

Possible values for list:

P 15

A 0

T 19

F 5

Possible values for list:

A 0

L 11

E 4

Q 16

After given these values, the user is allotted 10 guesses for which they may choose these values. I opted to avoid using the scoring mechanism and to utilize more user input to allow for user intuition to aid the decryption. For this particular cipher, the key is often a word, as opposed to seemingly random hill or affine values. After a few attempts at mixing and matching the above values, the key: M- O – R – A – L was able to successfully decrypt the cipher into the following plaintext:

Enter key 1: 12

Enter key 2: 14

Enter key 3: 17

Enter key 4: 0

Enter key 5: 11

WHEREHASHEBEENDURINGTHISWHOLECENTURYTHATISWHATTHISWHOLECENTURYHASBEENABOUTTHISSTRUGGLEBETWEENINTELLECTUALANDSOCIALPATTERNSTHATISTHETHEMESONGOFTHETWENTIETHCENTURYISSOCIETYGOINGTODOMINATEINTELLECTORISINTELLECTGOINGTODOMINATESOCIETYANDIFSOCIETYWINSWHATISGOINGTOBELEFTOFINTELLECTANDIFINTELLECTWINSWHATISGOINGTOBELEFTOFSOCIETYTHATWASTHETHINGTHATTHISEVOLUTIONARYMORALITYBROUGHTOUTCLEARERTHANANYTHINGELSEINTELLECTISNOTANEXTENSIONOFSOCIETYANYMORETHANSOCIETYISANEXTENSIONOFBIOLOGYINTELLECTISGOINGITSOWNWAYANDINDOINGSOISATWARWITHSOCIETYSEEKINGTOSUBJUGATESOCIETYTOPUTSOCIETYUNDERLOCKANDKEYANEVOLUTIONARYMORALITYSAYSITISMORALFORINTELLECTTODOSOBUTITALSOCONTAINSAWARNINGJUSTASASOCIETYTHATWEAKENSITSPEOPLESPHYSICALHEALTHENDANGERSITSOWNSTABILITYSODOESANINTELLECTUALPATTERNTHATWEAKENSANDDESTROYSTHEHEALTHOFITSSOCIALBASEALSOENDANGERITSOWNSTABILITY

Where has he been during this whole century? That is what this whole century has been about. This struggle between intellectual and social patterns. That is the theme song of the twentieth century… and if society wins what is going to be left of intellect and if intellect wins what is going to be left of society? That was the thing that this evolutionary morality brought out clearer than anything else: Intellect is not an extension of society any more than society is an extension of biology. Intellect is going its own way and in doing so is at war with society seeking to subjugate society - to put society under lock and key. An evolutionary morality says it is moral for intellect to do so but it also contains a warning just as society that weakens its people’s physical health then dangers its own stability. So does an intellectual pattern that weakens and destroys the health of its social base also endanger its own stability?

Answers to remaining hw (work attached)

2. a) 17-1 = 6 mod 101

b) 357-1 = 1075 mod 1234

c) 3125-1 = 1844 mod 9987

3. a) 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 – 2 | 4 | 6 | 1 | 8 | 3 | 5 | 7

b) Gentlemen do not read each other’s mail.

4. Involutory Keys

a) ek(ek(x)) = x

shift: (x+k) + (x + k) = x (mod 26)

affine: a(ax+b)+b = x (mod n)

b) shift:2 keys. 0 and 13

affine: a2=1 (mod n) and ab+b = 0 (mod n)