College of Computer and Information Science
Northeastern University
CSG252: Cryptography and Communication Security Fall 2004

Solution Set 1

Atilay Yilmaz ati@ccs.neu.edu / Guevara Noubir noubir@ccs.neu.edu

Problem 2 [20 points]:
Textbook problem 1.29, page 43.

a) Suppose we hypothesize that the keyword length is *m*. Define the following modified ciphertext:

$$y'_j = y_j - \left\lfloor \frac{j-1}{m} \right\rfloor, \text{ j = 1,2,…} \quad \text{(Shifting each block of } m \text{ letters)}$$

Then the string $y'_1, y'_2, y'_3 \ldots$ is the encryption of the same plaintext using the usual the Vigenere Cipher with the same keyword! Hence the methods used to cryptanalyze the Vigenere Cipher (Kasiski Test or Index of Coincidence) can be applied to this modified ciphertext string to determine the keyword length and the actual keyword. (page 31 in the book)

b) You need to write a program that finds:

1) Possible key length: guess the length starting from 2 and divide the ciphertext into blocks of size *m* and shift ciphertext by 1 for 2nd block, 2 for 3rd block, … so the ciphertext will be like encrypted with the same key and it will have the same probability distributions of letters in a standard English text. Then compute Index of coincidence and check if it is around 0.065. Do this for each length *m* and the one closest to 0.065 must be the key length. The others will be around 0.038, which is value for random text (not a language).
2) The key: Once you find the key length *m*, divide the ciphertext into blocks of size *m* and check for frequency of most used letters and match them to those in English like ‘e’, ‘a’,..

The key is PRIME and the plaintext is:

The most famous cryptologist in history owes his fame less to what he did than to what he said and to the sensational way in which he said it and this was most perfectly in character for Herbert Osborne Yardley was perhaps the most engaging articulate and technicolored personality in the business

Problem 3 [20 points]:
Textbook problem 1.22, page 41.

a) Suppose that $q'_j < q'_k$ for some $j < k$. Define

$$q''_j = \left\{ \begin{array}{ccc} q'_i & if & i \notin \{j,k\} \\ q'_k & if & i = j \\ q'_j & if & i = k \end{array} \right\} \qquad \text{(Swapping between } q_i \text{ values)}$$

Then we have

$$\sum_{i=1}^{n} p_i q''_i - \sum_{i=1}^{n} p_i q'_i = (p_j - p_k)(q'_k - q'_j) \geq 0$$

Therefore the desired sum is not decreased when q'_j and q'_k are exchanged. By a sequence of exchanges of this type, we see that the sum attains its maximum value when $q'_1 \geq q'_2 \geq ... \geq q'_n$.

b) Suppose that π is a permutation of {0,..,25} such that $p_{\pi(0)} \geq p_{\pi(1)} ... \geq p_{\pi(25)}$. Then it is 'likely' that $f_{\pi(0)} \geq f_{\pi(1)} ... \geq f_{\pi(25)}$. Assuming that this is the case, we proceed. When *g=0*, the following holds:

$$\sum_{i=0}^{25} \frac{p_i f_i}{n'} = \sum_{i=0}^{25} \frac{p_{\pi(i)} f_{\pi(i)}}{n'}$$

By the result proven in part (a), this sum is at least as great as any sum $\sum_{i=0}^{25} \frac{p_i f_{i+g}}{n'}$ where $g \neq 0$

Problem 4 [20 points]:
Textbook problem 1.25, page 42.
(By *Zina Saadi*)

Ciphertext:
LMQETXYEAGTXCTUIEWNCTXLZEWUAISPZYVAPEWLMGQWYAXFTCJMSQCADAGTXLMX
NXSNPJQSYVAPRIQSMHNOCVAXFV

The frequency of each string size 2:

TX 4
EW 3
LM 3
AG 2
AP 2
QS 2
AX 2
YV 2
AD 1
:

:

Finding the inverse Key Matrix:

We notice that the string “TX” occurs 4 times in the ciphertext, thus according to the list of the most frequent 2-letters strings in the English language, our first guess would be:
“TX” -> “th”
“EW” -> “in”

Then our decryption functions using the unknown matrix k is:
d(19,23)=(19,7)
d(4,22)=(8,12)

Note that we cannot have d(T) = t therefore, let us assume another string to decrypt to “th”:
“TX” -> “in”
“LM” -> “th”
“EW” -> “ou”

Thus the decryption function is as follow:
d(19,23)=(8,13)
d(11,12)=(19,7)
d(4,22)=(14,20)

Finding the inverse Key matrix:

Recall: $\mathsf{A} \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathsf{A}^{-1} = \frac{1}{|\mathsf{A}|} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$

$$K^{-1} * \begin{pmatrix} 19 & 23 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 8 & 13 \\ 19 & 7 \end{pmatrix}$$

$$Det\begin{pmatrix} 19 & 23 \\ 11 & 12 \end{pmatrix} = 19 * 12 - 11 * 23 = -25 \mod 26 = 1 \mod 26$$

$$Inv\begin{pmatrix} 19 & 23 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 12 & -23 \\ -11 & 19 \end{pmatrix} \mod 26 = \begin{pmatrix} 12 & 3 \\ 15 & 19 \end{pmatrix}$$

the inverse key matrix is:

$$K^{-1} = \begin{pmatrix} 12 & 3 \\ 15 & 19 \end{pmatrix} * \begin{pmatrix} 8 & 13 \\ 19 & 7 \end{pmatrix} = \begin{pmatrix} 23 & 21 \\ 13 & 16 \end{pmatrix}$$

$$Det\begin{pmatrix}23 & 21\\ 13 & 16\end{pmatrix} = 23*16-13*21 = 95 \mod 26 = 17 \mod 26$$

Then $$K = \begin{pmatrix}23 & 21\\ 13 & 16\end{pmatrix}^{-1} = \frac{1}{17}\begin{pmatrix}16 & -21\\ -13 & 23\end{pmatrix} = \frac{1}{17}\begin{pmatrix}16 & 5\\ 13 & 23\end{pmatrix} = \begin{pmatrix}4 & 11\\ 13 & 9\end{pmatrix}$$

checking for d(4,22)=?:

[4 22] inverse(k) = [14 20]

By applying the Key matrix to the cipher-text we obtain the following plaintext:

The plaintext:

THEKINGWASINHISCOUNTINGHOUSECOUNTINGOUTHISMONEYTHEQUEEN
WASINTHEPARLOUREATINGBREADANDHONEY