Homework #1

Gorkem Yar
27970

Due date: **27/10/2023**
Notes:

- Your work (code + written answers) must be submitted through SUcourse+.
- Winzip your programs and add a readme.txt document (**if necessary**) to explain the programs and how to use them.
- Name your winzip file as **"cs411_507_hw01_yourname.zip"**

1. (**15 pts**) Consider the shift cipher. Show that the ciphertext "NLPDLC" can be decrypted into a meaningful word. Find out this word and the corresponding encryption key.

For this question, I write a very simple python code to find all possible plaintexts. Since the question does not state any specific language, I used English Alphabet with 26 letters.

So, my approach is a Brute Force approach that tries all possible 26 different keys.

Answer:

CAESAR
Encryption Key is 11

2. (**20 pts**) Consider the ciphertext generated by Affine Cipher over Z_{26}. As a hint, you are told that the most frequent letter in the plaintext is 'T'. Find the plaintext, the encryption and decryption keys. Show your work.

"J gjg mxa czjq ayr arpa. J ulpa cxlmg ayerr ylmgerg rqrwrm hzdp ax gx ja hexmn."

To solve this question, I made an attack based on the frequencies of the alphabetic letters. The frequency of the letters are: {'j': 5, 'g': 6, 'm': 5, 'x': 5, 'a': 8, 'c': 2, 'z': 2, 'q': 2, 'y': 3, 'r': 8, 'p': 3, 'u': 1, 'l': 3, 'e': 3, 'w': 1, 'h': 2, 'd': 1, 'n': 1}. As it can be seen most frequent letters are "a" and "r". Since the most common letter in English is "e", I assume that "e" should correspond to one of these letters.

Later, I find all possible alpha values in Z26 with their gamma equivalents which are {1: 1, 3: 9, 5: 21, 7: 15, 9: 3, 11: 19, 15: 7, 17: 23, 19: 11, 21: 5, 23: 17, 25: 25}. There are 12 pairs for alpha and gamma values. There is one unique (for each frequent letter) Beta value for each alpha to map “e” to “a” or “r”.

After these steps, there are 24 possible valid encryptions in which 12 encryptions when “e” is map to “a” and 12 encryptions when “e” is map to “r”. I examined all these encryptions and find the plaintext as the following.

I did not fail the test. I just found three hundred eleven ways to do it wrong.

The encryption and decryption keys are as the following:
Alpha: 11
Beta: 25
Gamma: 19
Theta: 19

3. (15 pts) Assume that you design a new affine cipher where you encrypt two letters at a time, where your alphabet is

{'A':0, 'B':1, 'C':2, 'D':3, 'E':4, 'F':5, 'G':6, 'H':7, 'I':8, 'J':9, 'K':10, 'L':11, 'M':12, 'N':13, 'O':14, 'P':15, 'Q':16, 'R':17, 'S':18, 'T':19, 'U':20, 'V':21, 'W':22, 'X':23, 'Y':24, 'Z':25, '.':26, ' ':27, ',':28, '!':29}.

In other words, you group your plaintext message in bigrams (i.e., two-character words) and encrypt each bigram of the plaintext separately using this affine cipher. For example, if the first two letters of a plaintext is “TH” then it will be encoded as follows

TH ⇒ 19 × 30 + 7 = 577.

If the number of letters in the plaintext is not a multiple of two, you pad it with the letter “X” at the end. Determine the modulus and the size of the key space.

There are 30*30 = 900 unique encodings like (AA, AB, AC, ….). As a result, the modulus should be 900. To find the size of key space we need to find the find the count of possible alpha values. Alpha cannot exceed 900 as well as Beta. The possible alpha values should satisfy the following condition: gcd(alpha, 900) = 1. We can calculate the number of alphas by using phi function.

Phi(900) = 240
The number of Beta’s are equal to 900 since there is no such constraint on Beta.

So, the size of the key space is 240*900 = 216000.

Modulus: 900
Key Space: 216000

The encryption algorithm as the following:

alpha * encoded_bigram + beta
encoded_bigram varies between 0 to 900

4. (**15 pts**) Is the affine cipher defined in question (3) secure against the letter frequency analysis?

No, it does not secure against letter frequency analysis. In the English there are phrases that commonly used and some of the letters generally come after each other. There is also a frequency list for bigrams like the normal letter frequency. I found the following table on the internet which shows the bigram frequencies in the English.

th 3.56%	of 1.17%	io 0.83%
he 3.07%	ed 1.17%	le 0.83%
in 2.43%	is 1.13%	ve 0.83%
er 2.05%	it 1.12%	co 0.79%
an 1.99%	al 1.09%	me 0.79%
re 1.85%	ar 1.07%	de 0.76%
on 1.76%	st 1.05%	hi 0.76%
at 1.49%	to 1.05%	ri 0.73%
en 1.45%	nt 1.04%	ro 0.73%
nd 1.35%	ng 0.95%	ic 0.70%
ti 1.34%	se 0.93%	ne 0.69%
es 1.34%	ha 0.93%	ea 0.69%
or 1.28%	as 0.87%	ra 0.69%
te 1.20%	ou 0.87%	ce 0.65%

As it can be seen some of the bigrams are common. We can select the most common bigrams in our ciphertext, and they would probably correspond to the bigram “th” or “he”.

The source of the table is:

Bigram. (2023, October 26). Wikipedia. https://en.wikipedia.org/wiki/Bigram

5. **(20 pts**) Consider the following ciphertext that is encrypted with the affine cipher defined in question (3):

"ZHOFC.BNZCLRZ WNJ.XGI.WMBDV.MEJ!GGYKGDZ ERGMWNJ.KDGD RSW"

Find the key and decrypt the ciphertext.

(**Hint 1:** The plaintext is a sentence that ends with a dot.)
(**Hint 2:** The length of the plaintext (plen) is not a multiple of 2; plen = 2k+1 for an integer k.)

Using the hint given by the question, the last two bits in the encrypted text should be ".X". These two letters are ciphered into "SW".

Since we can calculate the encoding of both ".X" and "SW", we can easily reduce the number of possible alphas and betas. As it stated in the question 3, there are 240 different alpha variables in mod 900. For each alpha there is only one beta corresponding to satisfy the following equation:

alpha*encode(".X") + beta = encode("SW") mod 900

Since encode(".X") and encode("SW") are known constants and beta should take a specific value for each alpha. As a result, we have 240 different alpha, beta pairs.
Using these 240 pairs, I generated their equivalent gamma and theta keys for decryption purposes.

After using all of them, I was able to decrypt the ciphertext into the original plaintext from.

Answer:
SING, GODDESS, OF THE ANGER OF ACHILLES, SON OF PELEUS.

Keys:
Alpha: 91
Beta: 389
Gamma: 811
Theta: 421

6. (15 pts) If we select a different shift amount for every letter in the plaintext uniformly randomly, the shift cipher becomes a one-time-pad with perfect security. Suppose p_α is the probability of the plaintext letter α from the Turkish alphabet, where $\alpha \in \{A, B, C, Ç, \ldots, Z\}$. Suppose also that p_β is the probability of the ciphertext letter β, where $\beta \in \{A, B, C, Ç, \ldots, Z\}$. Demonstrate that $p_\beta = 1/29$ for every $\beta \in \{A, B, C, Ç, \ldots, Z\}$ independent of the values of p_α.

Since the shift amount for every letter in the plaintext is uniformly random and since there are 29 (the letter amount of the alphabet) possible shifts, the probability of any different shift for any letter is 1/29.

Let's say the probability of a letter in the plaintext is p_α where $\alpha \in \{A, B, C, Ç, \ldots, Z\}$. (The plaintext letters are not balanced.)

$$\sum_{\alpha \in \{A, B, C, Ç, \ldots, Z\}} p_\alpha = 1$$

To get any letter $\beta \in \{A, B, C, Ç, \ldots, Z\}$ in the ciphertext, we need to encrypt a letter in the plaintext. For each letter in the plaintext there is one unique shift to get the encrypted letter β. The probability of having this shift count is 1/29 because the shift count is random.

We can calculate the probability of letter β as the following:

$$P_\beta = \sum_{\alpha \in \{A, B, C, Ç, \ldots, Z\}} 1/29 * p_\alpha \quad \text{// 1/29 is the probability of the shift count for getting letter } \beta \text{ from } \alpha$$

$$= 1/29 * \sum_{\alpha \in \{A, B, C, Ç, \ldots, Z\}} p_\alpha$$

$$= 1/29 * 1$$

$$= 1/29$$

7. BONUS (20 pts) The following was encrypted using the Vigènere cipher:

"FNZ FFZZMLQQZVO GAXXH PZ UPU QXGIHU UY NWJXR AHBDLPOMK YOUPZM, VOZAYCD. J TGQH B XUIJJZM ARS XOAH, BZJ D JP AT GLWUTB LO EVDWF AL GRHUI. OKPGMC L NME IRU NKGLFHK DQ UTK JUEQX JI UTK PQJHKMVF, KKO L MABZ WIQ YOLDWE GLUFRZ OFMBZV BE ZCHZ AVZQ JZ YKUJZM. D OPHK OKF NRPH TWE, D OPHK NRNQ VZRQXK, RKPY UIH MABZV ZAA FQPI YJPFFOHHT IOOKPGZ FQPIOIJ XTE. D OPHK NRNQ MMHBF JZHEE JJQF NE HHO, FNJXHT O'QH MATB FFMYZG QQXCDQE ZJ KBHK ADJFN DQ UTKH, BFF LMRN ARY KBNOO ROQ'Y CHBDZ KUJLKN WIQS. CHSQ ZCHZ TGQH CDUPJIF ZCH TAAK IPD EJX, FMZ DW, JF CDOM PU TRV SUJG. JF'Y ALSEZ-MDUQ YJXQ, FNZB LZUR KPI ZJ PBWK DW IQXZ. L XMTO WP FXVYFX OI HVDUKH, BXEJVIM, O NKBXR NHU ALA ISAS CHSQ. GIG ZQZ D NOAC OKBF O VP PZRT JPUTB WP M MMDWQEVUE, NAO LU'E G HRTF VMHDUUPV HDGQHZMXY, WIMZ'N ZIMZ DW JE. VMHDUUPV BDK OKF PKVG UTGO OJQ ZCHSQ, KQHSK YOROQ UQHS FNZP TBKVNT AL NXDT HPUOUTB OJRK DQ UTK KDTF, UA VVON KDTEOJQBFK ADJFN DQ UTKDU XAXF, WIQOM WSGZC, WIQOM VUDABJMQ GIG UTKDU TOOZQDQ, ZCDU U QIRX U YCDMX LVOM AT OKF SXJXOP GIG LUYN WIAYZ VUATZV BZJ RHFB UQHS FNZP; UTUPJI U'S XROHOIFFP OI PZ TKVUU FNVW JF'Y GROS HZHO ZUOKJZM WXU M MMDWQEVUE. MTY L TTGGO OAZ RHFB LMRN

PKNSBUX, WXU EOHSMK HZFBGYZ L TTGGO CQ NVSQK OI PZ FKVUT, U YCDMX YOHFB ST VPGR DQ FYUOLPZ. O GRWQ ZCH TFOXNZ XKVYFE OI VQDOIJ, UTK WOVQ YFB - UTGO'V BXR DW JE. OO'V OAZ V PBFZZU PR OIWFXRZFU AX GRHUI, DW'T XUQLOS CDWI ATZ'V JZYDGF, IOOK PZK'N VUASVFI."

Attack it and find the key length and the key. Note that only the letter characters are encrypted.

As it stated in the lecture, I firstly tried to find the key length of the Vigenère Cipher. To find it, I created a loop that shifts the original text by one letter in each iteration and count the intersected letters. In this approach, I found the key length as 5.

After, I divided the ciphertext into 5 parts to solve each part separately. Each part is a shift cipher, and we can solve shift ciphers by frequency attacks. Since the letters "e", "t", and "a" are the most common letters in English I attacked the ciphers with these letters. In each cipher, I find the most frequent letter and assumed that one of the most frequent letters in English should mapped to this letter. Since I choose the most common three letters in English, I get 3 different solutions for each cipher. For 5 different ciphers in which I have 3 solutions for each of them. I get 3^5 = 243 different solutions.

One of the solutions among this 243 is the true one. Which is:
The key is **"MGVDB"**

THE CENTRIPETAL FORCE ON OUR PLANET IS STILL FEARFULLY STRONG, ALYOSHA. I HAVE A LONGING FOR LIFE, AND I GO ON LIVING IN SPITE OF LOGIC. THOUGH I MAY NOT BELIEVE IN THE ORDER OF THE UNIVERSE, YET I LOVE THE STICKY LITTLE LEAVES AS THEY OPEN IN SPRING. I LOVE THE BLUE SKY, I LOVE SOME PEOPLE, WHOM ONE LOVES YOU KNOW SOMETIMES WITHOUT KNOWING WHY. I LOVE SOME GREAT DEEDS DONE BY MEN, THOUGH I'VE LONG CEASED PERHAPS TO HAVE FAITH IN THEM, YET FROM OLD HABIT ONE'S HEART PRIZES THEM. HERE THEY HAVE BROUGHT THE SOUP FOR YOU, EAT IT, IT WILL DO YOU GOOD. IT'S FIRST-RATE SOUP, THEY KNOW HOW TO MAKE IT HERE. I WANT TO TRAVEL IN EUROPE, ALYOSHA, I SHALL SET OFF FROM HERE. AND YET I KNOW THAT I AM ONLY GOING TO A GRAVEYARD, BUT IT'S A MOST PRECIOUS GRAVEYARD, THAT'S WHAT IT IS. PRECIOUS ARE THE DEAD THAT LIE THERE, EVERY STONE OVER THEM SPEAKS OF SUCH BURNING LIFE IN THE PAST, OF SUCH PASSIONATE FAITH IN THEIR WORK, THEIR TRUTH, THEIR STRUGGLE AND THEIR SCIENCE, THAT I KNOW I SHALL FALL ON THE GROUND AND KISS THOSE STONES AND WEEP OVER THEM; THOUGH I'M CONVINCED IN MY HEART THAT IT'S LONG BEEN NOTHING BUT A GRAVEYARD. AND I SHALL NOT WEEP FROM DESPAIR, BUT SIMPLY BECAUSE I SHALL BE HAPPY IN MY TEARS, I SHALL STEEP MY SOUL IN EMOTION. I LOVE

THE STICKY LEAVES IN SPRING, THE BLUE SKY - THAT'S ALL IT IS. IT'S NOT A MATTER OF INTELLECT OR LOGIC, IT'S LOVING WITH ONE'S INSIDE, WITH ONE'S STOMACH.