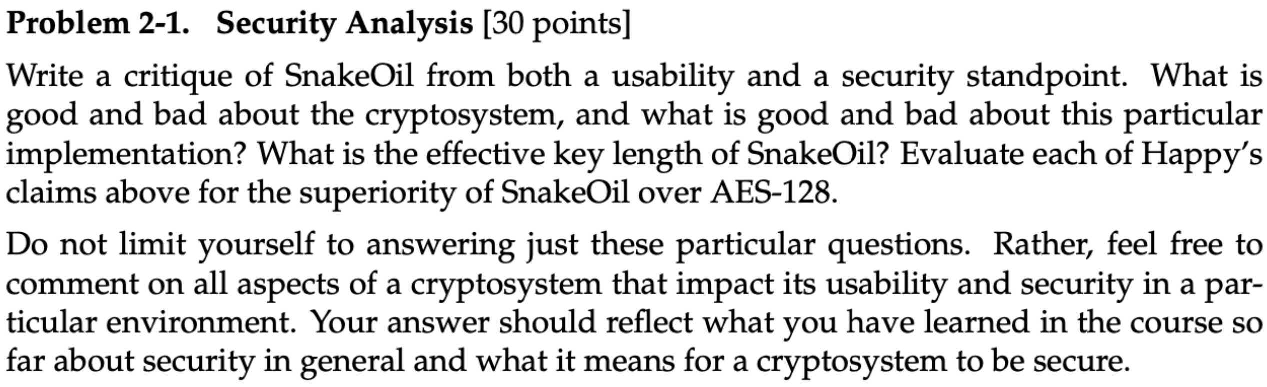
Aryan Nath November 3, 2024

CS-1306 / MAT-1201 Problem Set 2

Collaborators: *none*

Problem 1



Problem 2-1. Security Analysis [30 points]

Write a critique of SnakeOil from both a usability and a security standpoint. What is good and bad about the cryptosystem, and what is good and bad about this particular implementation? What is the effective key length of SnakeOil? Evaluate each of Happy's claims above for the superiority of SnakeOil over AES-128.

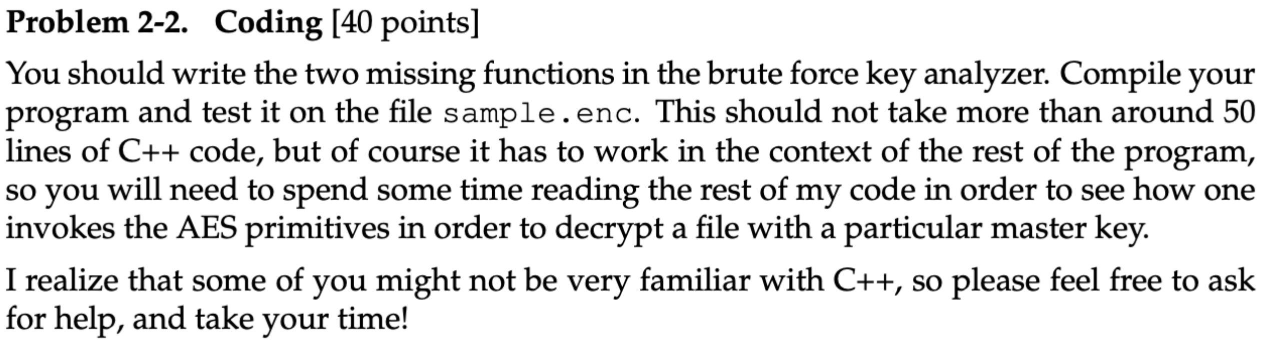
Do not limit yourself to answering just these particular questions. Rather, feel free to comment on all aspects of a cryptosystem that impact its usability and security in a particular environment. Your answer should reflect what you have learned in the course so far about security in general and what it means for a cryptosystem to be secure.

(a) A good system is one which makes it very hard for an adversary to guess a message plaintext using just ciphertext and priors about the cryptosystem that have been published. There are more than one factors that need to be satisfied to make this problem hard for the adversary. Firstly, and most basically, the letters of the ciphertext should atleast be a shuffling of the plaintext - the adversary should not be able to just look at the ciphertext and guess the plaintext. The lesser information that the adversary is able to retrieve about the plaintext from just the ciphertext and the public parameters of the cryptosystem, the better is the cryptosystem, or more precisely, the probability of guessing the plaintext should be increased by having the ciphertext. A good cryptosystem will produce a ciphertext that will look completely random to the adversary, this should even be so for repeated letter plaintexts - which was not the case for the breakable cryptosystem from assignment 1. Besides randomness, the adversary should not be able to combine or perform operations on ciphertexts and get some meaningful information that they can use to break the cryptosystem. For example, if a cryptosystem decrypts the two messages 'Good morning, the weather is great today', and 'Good morning, I'm going for jog', now if the encryption is letter by letter, the adversary should not be able to for example subtract the ASCII values of the sets of letters and some representation of difference between the key values used to decrypt the two ciphertexts. Lastly, the cryptosystem should be resistant to brute force attacks. The adversary should not be able to efficiently check decryptions for all possible keys(brute force attack) in the key space and get the correct plaintext through some additional processing such as letter frequency analysis. A cryptosystem that does not satisfy any one of these factors is not a good cryptosystem and can be 'broken' by an adversary without knowing the encryption key. In case of usability, the encryption function should be computationally feasible otherwise the cryptosystem will be redundant.

(b) Now let's consider the SnakeOil modification of AES-128/CBC. Firstly, if zeroes are interpreted as NUL/NULL, then this will not create a problem for decoding the texts where the ciphertext receiver can assume no NUL/NULL values in the text. For any other file type which can contain NUL/NULL values, there will be decoding errors as simply padding with 0's will not help the receiver identify the number of the 0's to remove as a 0 byte (NUL/NULL) can be a part of the plaintext. But just for the case of text files, this zero padding cryptosystem is good in the context of padding as the adversary cannot use the zero padding to perform operations on the ciphertext to retrieve information about the plaintext. The other modification made is the key management system. In terms of key generation, the XOR'ed master key will still be random because the XOR of two random values is also random. So just from this perspective as well, the cryptosystem is good as there are not a certain set of master keys that are more likely to occur in the key space that the adversary can use to guess the encryption key. Besides this, since Happy is still following AES for encryption, the randomness of the ciphertexts and inability to guess the plaintext with just the ciphertext will be preserved.

(c) However, if the key shares are public, Happy has effectively reduced the key space size of AES-128 from 2^{128} to ${}^{100}C_2$. The adversary now has to just brute force check all possible pairs of key shares to generate a master key instead of trying for all possible 2^{128} encryption keys with undisclosed keys(making a subset of encryption keys public was already a big mistake). The key space is now fixed in terms of the public key shares and not 128 unkown bits per key. This made it *feasible* for the adversary to brute force the effective key space and guess the plaintext after some processing like letter frequency analysis.

(d) Beyond the three characteristics I have mentioned above, a cryptosystem is good if its ciphertexts are random, which can be measured using good confusion and diffusion values - how does a change in the jey and plaintext have in the ciphertext output, without either of the two being traceable from the ciphertext. Moreover, the security of the cryptosystem should rely completely on the key being kept secret so that the cryptosystem itself can be made public to other users for communication. Lastly, encryption and decryption with any key should be computationally feasible.



Problem 2-2. Coding [40 points]

You should write the two missing functions in the brute force key analyzer. Compile your program and test it on the file `sample.enc`. This should not take more than around 50 lines of C++ code, but of course it has to work in the context of the rest of the program, so you will need to spend some time reading the rest of my code in order to see how one invokes the AES primitives in order to decrypt a file with a particular master key.

I realize that some of you might not be very familiar with C++, so please feel free to ask for help, and take your time!

The code for this question has been submitted on google classroom.

The plaintext that I retrieved is:

```
This is a sample text that is encrypted using keyshares with key
indices 4 and 31.
```

Files used:

- Frequency table: data/space.dat
- keyshares hex codes: data/keyshares
- file encrypted using snakeoil: data/sample.enc
- output file: project_root/output.txt

Problem 2-3. Experiments [30 points]

You will find several files in the `data` folder. Files ending in `.dat` are frequency tables. Files ending in `.enc` are ciphertexts. The file `keyshares` is the key share file that was used in all of the encryptions. All ciphertexts are valid encryptions of English-language text files, but not all are easy to decipher.

You should run your brute force key finder with every frequency table and every ciphertext file. For each, report the key indices that the program found, the decryption produced, and whether or not the program succeeded in finding the correct key.

Next, you should analyze your results and try to draw conclusions about the effects of the frequency table and the length of the ciphertext on the effectiveness of the attack. Based on the insights gained, construct a 40 character "message" that cannot be cracked using any of the furnished frequency tables. (Your message does not have to be real English text, but it must consist of printable ASCII characters.)

1. mystery1.enc

   Length of ciphertext: 1488 bytes

   (a) frequency table used: merrywives.dat
   Key Indices: 14 and 93

      i. Decryption:
      Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.
      Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.
      But, in a larger sense, we can not dedicate – we can not consecrate – we can not hallow – this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us – that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion – that we here highly resolve that these dead shall not have died in vain – that this nation, under God, shall have a new birth of freedom – and that government of the people, by the people, for the people, shall not perish from the earth.

   ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the merriwives.dat frequency table.

(b) frequency table used: space.dat
Key Indices: 14 and 93

   i. Decryption:
   Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.
   Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.
   But, in a larger sense, we can not dedicate – we can not consecrate – we can not hallow – this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us – that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion – that we here highly resolve that these dead shall not have died in vain – that this nation, under God, shall have a new birth of freedom – and that government of the people, by the people, for the people, shall not perish from the earth.

   ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the space.dat frequency table.

(c) frequency table used: ulysses.dat
Key Indices: 14 and 93

   i. Decryption:
   Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.
   Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.
   But, in a larger sense, we can not dedicate – we can not consecrate – we can not hallow – this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The

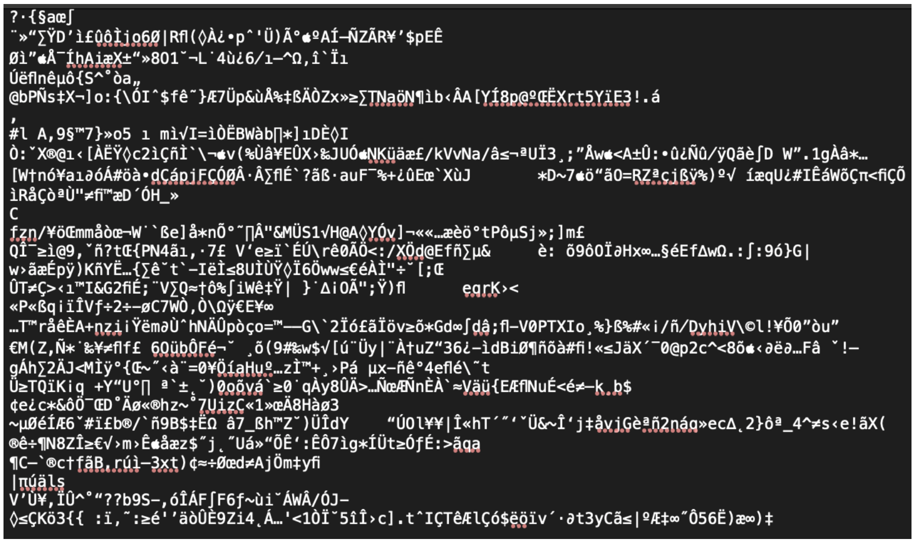
world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us – that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion – that we here highly resolve that these dead shall not have died in vain – that this nation, under God, shall have a new birth of freedom – and that government of the people, by the people, for the people, shall not perish from the earth.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the ulysses.dat frequency table.

(d) frequency table used: uniform.dat

Key Indices: 43 and 65

i. Decryption:



```
?·{§aœ∫
¨»“∑ŸD’ì£ûôÌjo6Ø|Rfl(◊À¿•p^'Ü)Ã°ºAÍ–ÑZÃR¥’$pEÊ
Øì”Å¯ÍhAiæX±“»801˘¬L˙4ù¿6/ı–^Ω,î`Ïı
Úëflnêµô{S^˚òa„
@bPÑs‡X¬]o:{\ÓI^$fê˜}Æ7Üp&ùÅ%‡ßÄÒZx»≥∑TNaöN¶ìb‹ÂA[YÍ8p@ºŒËXrt5YïE3!.á
,
#l A,9§™7}»o5 ı mì√I=ìÒËBWàb∏*]ıDÈ◊I
Ò:˘X®@ı‹[ÀËŸ◊c2ìÇñÌ`\¬v(%Ùâ¥EÛX›‰JUÓNKüäæ£/kVvNa/â≤¬ªUÏ3¸;”Åw<A±Û:•û¿Ñû/ÿQãè∫D W”.1gÀâ*…
[W†nó¥aı∂óÁ#öà•dÇápjFÇÓØÂ·Â∑flÉ`?ãß·auF¯%+¿ûEœ`XùJ       *D~7ö“ãO=RZªçjßÿ%)º√ íæqU¿#IÊáWõÇπ<fiÇÕ
ìRåÇòªÙ"≠fi™æD´ÓH_»
C
fzn/¥öŒmmåòœ¬W˙`ße]å*nÕ°˜∏Â"&MÜS1√H@A◊YÓv]¬««…æèö°tPôµSj»;]m£
QÎ¯≥ì@9,˘ñ?tŒ{PN4ãı,·7£ V‘e≥ï`ÉÚ\rê0ÃÖ<:/XÖd@Efñ∑µ&      è: õ9ôOÏ∂Hx∞…§éEfΔwΩ.:∫:9ó}G|
w›ãæÉpÿ)KñYË…{∑ê˘t`–IëÌ≤8UÌÙŸ◊Ï6Öww≤€éÀÌ"÷˘[;Œ
ÛT≠Ç>‹ı™I&G2fiÉ;¨V∑Q≈†ô%∫iWê‡Ÿ| }˙ΔiOÃ";Ÿ)fl       eqrK›<
«P«ßqiïÎVƒ÷2÷–øC7WÒ,Ò\Ωÿ€E¥∞
…T™råêÈA+nziiŸëm∂Ù^hNÄÛpòço=™——G\`2Ïó£ãÏöv≥õ*Gd∞∫dâ;fl–V0PTXIo¸%}ß%#«i/ñ/DyhiV\©l!¥Õ0”òu”
€M(Z,Ñ*˙‰¥≠flf£ 6OübÔFé¬˘ ¸õ(9#‰w$√[ú¨Üy|¨À†uZ“36¿–ìdBiØ¶ñõà#fi!«≤JäX´¯0@p2c^<8õ‹∂ë∂…Fâ ˘!–
gÁh∑2ÄJ<MÌÿ°{Œ~˝‹à¨=0¥ÖíaHuº…zÌ™+¸›Pá µx–ñê°4eflé\˜t
Ü≥TQïKiq +Y“U°∏ ª`±¸˘)0oõvá`≥0˙qÀy8ÛÄ>…ÑœÆÑnÈÀ`≈Väü{EÆflNuÉ<é≠–k¸b$
¢e¿c*&ôÖ¯ŒD˚Äø«®hz~˚7UizC«1»œÄ8Hàø3
~µØéÍÆ6˘#ï£b®/`ñ9B$‡ËΩ â7_ßh™Z˘)ÜÎdY    “ÚOl¥¥|Î«hT´˝‘˘Ü&~Î‘j‡åvjGèªñ2náq»ecΔ¸2}ôª_4^≠s‹e!ãX(
®ê÷¶N8ZÎ≥€√›m›Êåæz$˝j¸˝Uá»“ÕÊ‘:ÊÔ7ìg*ÍÜt≥ÓƒÉ:>ãqa
¶C–`®c†fãB¸rúì–3xt)¢≈÷Øœd≠AjÖm‡yfi
|πúäls
V’Ù¥,ÏÛ^°“??b9S–,óÎÁF∫F6ƒ~ùi˘ÁWÂ/ÓJ–
◊≤ÇKö3{{ :ï,˜:≥é'’äòÛÈ9Zi4¸Á…'<1ÒÏ˘5îÎ›c].t^IÇTêÆlÇó$ëöïv´·∂t3yCã≤|ºÆ‡∞˝Ô56Ë)æ∞)‡
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the uniform.dat frequency table.

2. mystery2.enc

Length of ciphertext: 736 bytes

(a) frequency table used: merrywives.dat

Key Indices: 55 and 82

i. Decryption:

I have a dream that one day this nation will rise up and live out the true meaning of its creed: "We hold these truths to be self-evident: that all men are created equal." I have a dream that one day on the red hills of Georgia the sons of former slaves and the sons of former slaveowners will be able to sit down together at a table of brotherhood. I have a dream that one day even the state of Mississippi, a desert state, sweltering with the heat of injustice and oppression, will be transformed into an oasis of freedom and justice. I have a dream that my four children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the merriwives.dat frequency table.

(b) frequency table used: space.dat

Key Indices: 55 and 82

i. Decryption:

I have a dream that one day this nation will rise up and live out the true meaning of its creed: "We hold these truths to be self-evident: that all men are created equal." I have a dream that one day on the red hills of Georgia the sons of former slaves and the sons of former slaveowners will be able to sit down together at a table of brotherhood. I have a dream that one day even the state of Mississippi, a desert state, sweltering with the heat of injustice and oppression, will be transformed into an oasis of freedom and justice. I have a dream that my four children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the space.dat frequency table.

(c) frequency table used: ulysses.dat

Key Indices: 55 and 82

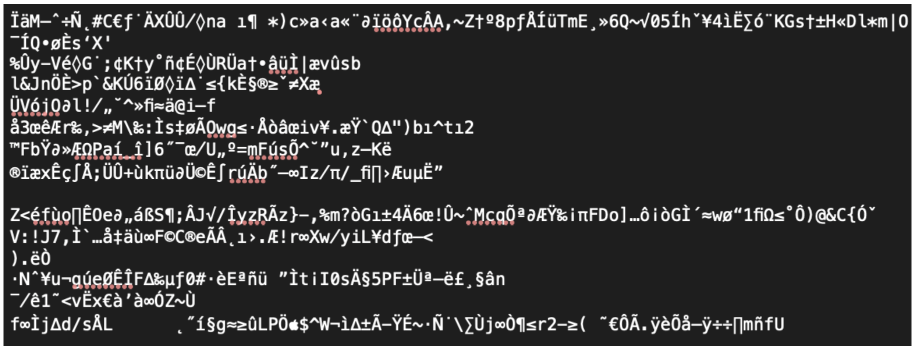
i. Decryption:

I have a dream that one day this nation will rise up and live out the true meaning of its creed: "We hold these truths to be self-evident: that all men are created equal." I have a dream that one day on the red hills of Georgia the sons of former slaves and the sons of former slaveowners will be able to sit down together at a table of brotherhood. I have a dream that one day even the state of Mississippi, a desert state, sweltering with the heat of injustice and oppression, will be transformed into an oasis of freedom and justice. I have a dream that my four children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the ulysses.dat frequency table.

(d) frequency table used: uniform.dat
Key Indices: 6 and 3

i. Decryption:



```
ÏäM—ˆ÷Ñ˛#C€ƒ˙ÄXÛÛ/◊na ı¶ *)c»a‹a«¨∂ïöôYcÂA,~Z†º8pƒÅÍüTmE¸»6Q~√05Íhˇ¥4ìË∑ó¨KGs†±H«Dl*m|O
¯ÍQ•øÈs‘X'
%Ûy—Vé◊G˙;¢K†y˚ñ¢É◊ÙRÜa†•âüÌ|ævûsb
l&JnÖÈ>p`&KÚ6ïØ◊ïΔ˙≤{kÈ§®≥ˇ≠Xæ
ÜVójO∂l!/„˘^»ﬁ≈ä@i—f
å3œêÆr‰,>≠M\‰:Ìs‡øÃOwg≤·Åòâœiv¥.æŸ`QΔ")bı^tı2
™FbŸ∂»ÆΩPaí_î]6˝¯œ/U„º=mFúsÕ^˘”u,z—Kë
®ïæxÊçƒÅ;ÜÛ+ùkπü∂Ü©Ê∫rúÄb˝—∞Iz/π/_ﬁ∏›ÆuµË”

Z<éfùo∏Ê0e∂„áßS¶;ÂJ√/ÎyzRÃz}–,%m?òGı±4Ä6œ!Û~ˆMcqÕª∂ÆŸ‰iπFDo]…ô¡òGÌ´≈wø“1ﬁΩ≤˚Ô)@&C{Óˇ
V:!J7,Ì`…å‡äù∞F©C®eÃÂ˛ı›.Æ!r∞Xw/yiL¥dƒœ—<
).ëÒ
·Nˆ¥u¬gúeØÊÎFΔ‰µƒ0#·èEªñü ”Ìt¡I0sÄ§5PF±Üª—ë£¸§ân
¯/ê1˜<vËx€à’à∞ÓZ~Ù
f∞ÌjΔd/sÅL       ˛˝í§g≈≥ûLPÖ$^W¬ìΔ±Ã–ŸÉ~·Ñ˙\∑Ùj∞Ò¶≤r2–≥( ˜€ÔÃ.ÿèÕå–ÿ÷÷∏mñfU
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the uniform.dat frequency table.

3. mystery3.enc

Length of ciphertext: 64 bytes

(a) frequency table used: merrywives.dat
Key Indices: 2 and 56

i. Decryption:
That's one small step for man; one giant leap for mankind.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the merriwives.dat frequency table.

(b) frequency table used: space.dat
Key Indices: 2 and 56

i. Decryption:
That's one small step for man; one giant leap for mankind.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the space.dat frequency table.

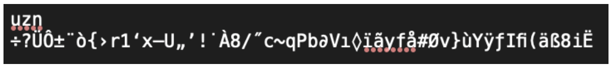
(c) frequency table used: ulysses.dat
Key Indices: 2 and 56

i. Decryption:
That's one small step for man; one giant leap for mankind.

ii. Success in decryption: Yes, the program was able to decrypt the ciphertext using the ulysses.dat frequency table.

(d) frequency table used: uniform.dat
Key Indices: 3 and 18

i. Decryption:



```
uzn
÷?ÜÔ±¨ò{›r1‘x—U„’!˙À8/˝c~qPb∂Vı◊ïãyfå#Øv}ùYÿƒIﬁ(äß8iË
```

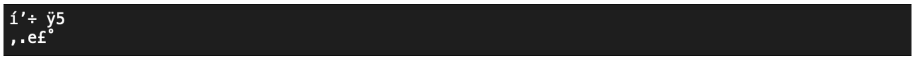
ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the uniform.dat frequency table.

4. mystery4.enc

Length of ciphertext: 16 bytes

(a) frequency table used: merrywives.dat
Key Indices: 13 and 82

i. Decryption:

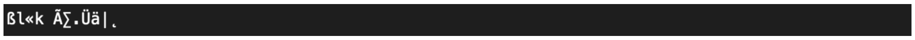


```
í’÷ ÿ5
,.e£˚
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the merriwives.dat frequency table.

(b) frequency table used: space.dat
Key Indices: 1 and 97

i. Decryption:

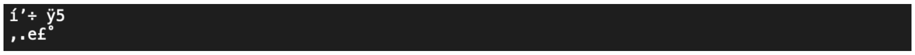


```
ßl«k Ã∑.Üä|˛
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the merriwives.dat frequency table.

(c) frequency table used: ulysses.dat
Key Indices: 13 and 82

i. Decryption:

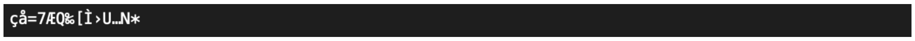


```
í’÷ ÿ5
,.e£˚
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the merriwives.dat frequency table.

(d) frequency table used: uniform.dat
Key Indices: 0 and 1

i. Decryption:



```
çå=7ÆQ‰[Ì›U…N*
```

ii. Success in decryption: No, the program was not able to decrypt the ciphertext using the uniform.dat frequency table.

For the second part of the question, I need to find a plaintext such that after it has been encrypted with the selected keyshares pair, I wont be able to retrieve the correct plaintext using a frequency analysis attack without knowing the correct keyshares pair.

I noted that the *uniform.dat* has assigned frequency of 1 to each of the ascii characters, and using it we cannot decrypt any of the given encrypted messages using a frequency analysis attack.

space.dat assigns 0 to all of the ascii characters except 1.

Let's first check if these two are able decrypt the ciphertext of a plaintext consisting of just a single letter.

I used 40 0's consecutively as my plaintext and I couldn't perform a successful bruteforce attack with any of the frequency tables.

Selected plaintext:

```
0000000000000000000000000000000000000000
```

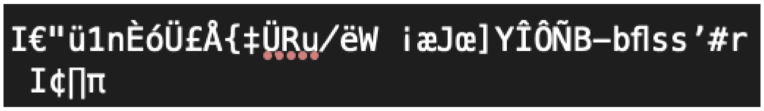
Key indices used for encryption: 4 and 62

1. merrywives.dat

   Key Indices: 59 and 71

   Decryption:

   (a) Decryption:



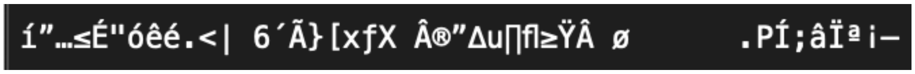
```
I€"ü1nÈóÜ£Å{‡ÜRu/ëW iæJœ]YÎÔÑB–bflss’#r
 I¢∏π
```

2. space.dat

   Key Indices: 22 and 59

   Decryption:

   (a) Decryption:



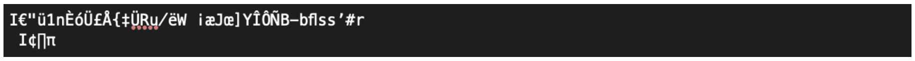
```
í”…≤É"óêé.<| 6´Ã}[xƒX Â®”Δu∏fl≥ŸÂ ø      .PÍ;âÏª¡–
```

3. space.dat

   Key Indices: 59 and 71

   Decryption:

   (a) Decryption:



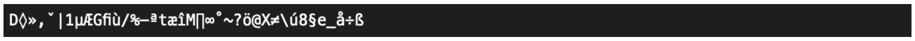
```
I€"ü1nÈóÜ£Å{‡ÜRu/ëW iæJœ]YÎÔÑB–bflss’#r
 I¢∏π
```

4. uniform.dat

   Key Indices: 3 and 86

   Decryption:

   (a) Decryption:



```
D◊»,ˇ|1µÆGfiù/%–ªtæîM∏∞˚~?ö@X≠\ú8§e_å÷ß
```

Now let's analyze why I could not perform a successful frequency analysis attack using the frequency tables.

Firstly, I cannot directly infer that the reason for this was the ciphertext length since in the given encryptions, only *mystery4.enc* had a small enough size - 16 bytes - that one of reasons that none of the frequencies could be used to find the plaintext was because of the small size of the plaintext and hence the frequency distributions of the decryptions for not comparable to that expected from normal english sentences.

Secondly, I noticed that for none of the ciphertexts, the *uniform.dat* frequency table could give a feasible distribution. The frequency table of uniform.dat is just 256 consecutive 1's. This means that based on *unform.dat*, each ascii character is equally likely in the plaintext. So we were not able to use it to identify whether the frequency distribution of an obtained plaintext has a unique divergence with respect to uniform.dat. A similar case follows for *space.dat*. In this frequency distribution, just a single ascii letter is expected to be present in the plaintext, and rest all of the ascii characters are expected to be absent. For the other two frequency tables, I considered having the frequency distribution of the plaintext as different as possible from the given frequency table. Having a plaintext consisting of the same character seems to work, even for *space.dat* as long as the character is different from the which is assigned 1 in the frequency table. So any plaintext here consisting of all the same ascii character, other than the ascii character assigned 1 in the frequency table *space.dat*, would work as the frequency distribution of such a plaintext would be quite different from all of the other frequency tables.