Secret Key Encryption Project Report

Task 1: Frequency Analysis

Description: In this task, we applied frequency analysis to decipher a ciphertext that was encrypted using a monoalphabetic substitution cipher. The goal was to identify the original plaintext based on the frequency of individual letters and common letter pairs (bigrams and trigrams) in English.

Observations:

- By analyzing the frequency of letters in the ciphertext, we matched high-frequency letters to common letters in the English language.
- Bigram and trigram frequencies were analyzed, helping us identify frequent letter combinations and map them to common English word patterns.

Approach:

- Used the freq.py script to generate 1-gram, 2-gram, and 3-gram statistics from the ciphertext file. This script reads the file and identifies the most common patterns to help decipher the text.
- Mapped frequent ciphertext letters to common English letters such as 'E', 'T', 'A', etc
- Replaced ciphertext characters gradually to reveal the original message.

```
#!/usr/bin/env python3
from collections import Counter
import re
TOP_K = 20
N_{GRAM} = 3
# Generate all the n-grams for value n
def ngrams(n, text):
   for i in range(len(text) -n + 1):
       # Ignore n-grams containing white space
       if not re.search(r'\s', text[i:i+n]):
          yield text[i:i+n]
# Read the data from the ciphertext
with open('ciphertext.txt') as f:
   text = f.read()
# Count, sort, and print out the n-grams
for N in range(N_GRAM):
  print("-----
   print("{}-gram (top {}):".format(N+1, TOP_K))
   counts = Counter(ngrams(N+1, text)) # Count
   sorted_counts = counts.most_common(TOP_K) # Sort
   for ngram, count in sorted_counts:
      print("{}: {}".format(ngram, count))
                                            # Print
```

Task1 (freq.py)

```
[10/01/24]seed@VM:-/../Files$ python3
Python 3.8.5 (default, Jul 28 2020, 12:59:40)
[GCC 9.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import random
>>> s = "abcdefghijklmnopqrstuvwxyz"
>>> list = random.sample(s, len(s))
>>> key = ''.join(list)
>>> print(key)
leyzfmcitjbxdnhswkvaqougpr
```

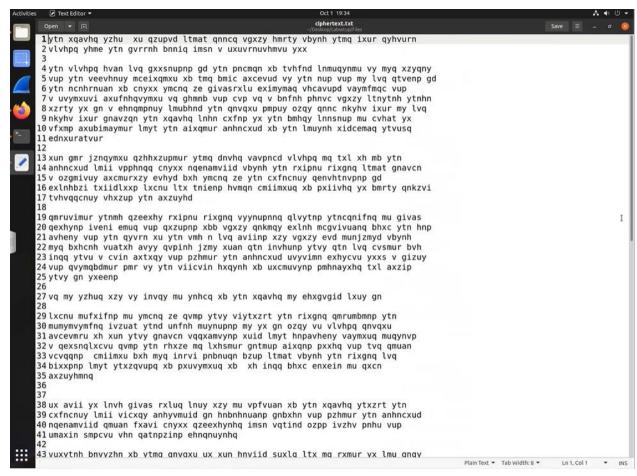
Task1 (imported random in order to generate key)

```
Seed-Ubuntu20.04 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities  Terminal 

                                                                                                                                - ø ×
                                                                     Oct 1 19:32
                                                                  seed@VM: ~/.../Files
[10/01/24]seed@VM:~/.../Files$ ./freq.py
```

Task1 (This is the n-gram that was produced)

Task1 (This is the n-gram that was produced)

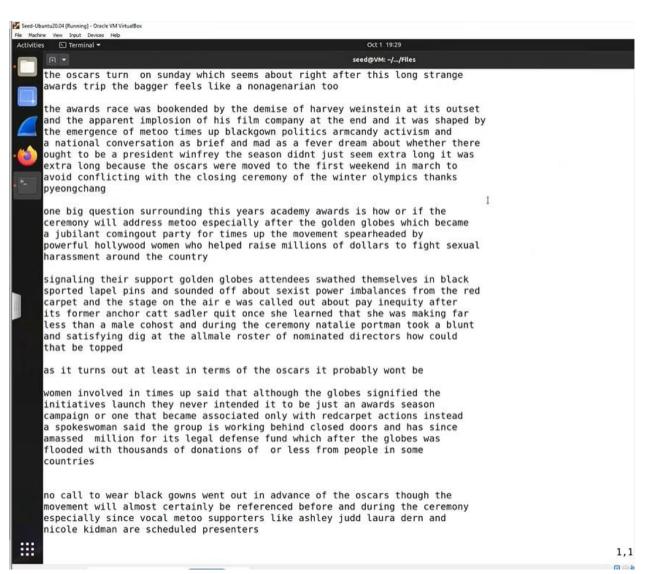


Task1 (ciphertext)

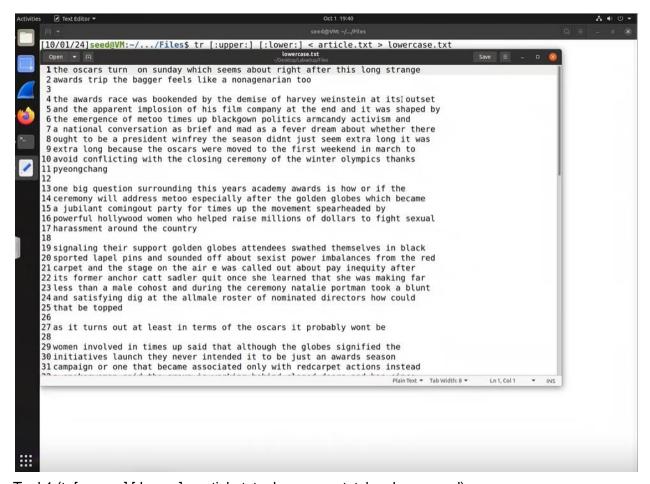
Based on the n-gram, we can analyze the ciphertext.txt file. The encryption key could be as follows:

```
\begin{array}{l} ytn \rightarrow THE \\ y \rightarrow T \\ t \rightarrow H \\ n \rightarrow E \\ tn \rightarrow HE \\ . \\ . \\ . \end{array}
```

so on...



Task1 (decrypted ciphertext)



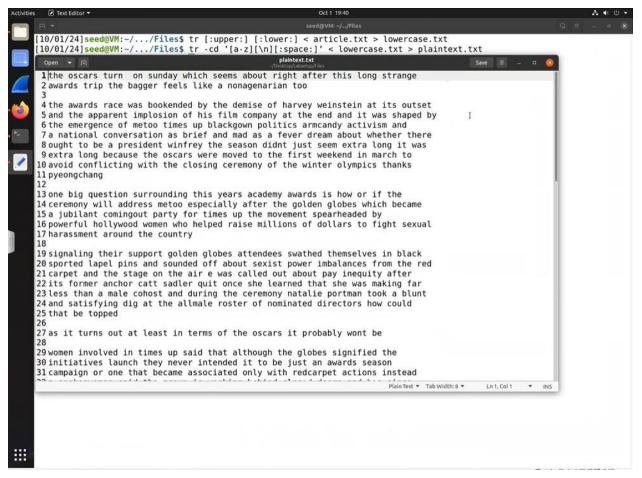
Task1 (tr [:upper:] [:lower:] < article.txt > lowercase.txt has been used)

The actual alphabet letter goes first, followed by the cipher key it corresponds to:

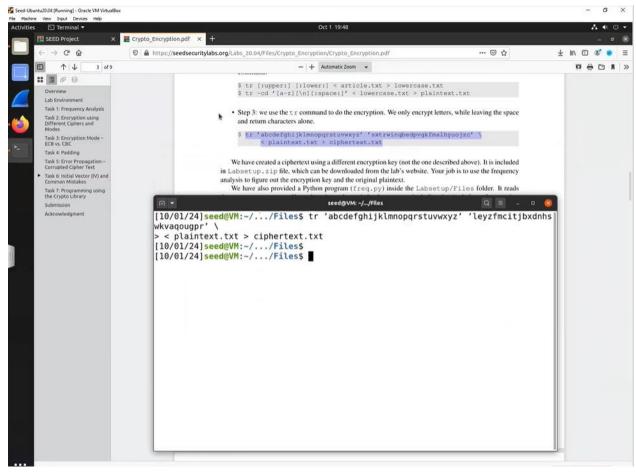
```
\bullet A=V \bulletB=G \bulletC=A \bulletD=P \bulletE=N \bulletF=B \bulletG=R \bulletH=T \bulletI=M \bulletJ=W
```

$$\bullet$$
 K=S \bullet L=I \bullet M=C \bullet N=U \bullet O=X \bullet P=E \bullet Q=O \bullet R=H \bullet S=Q \bullet T=Y

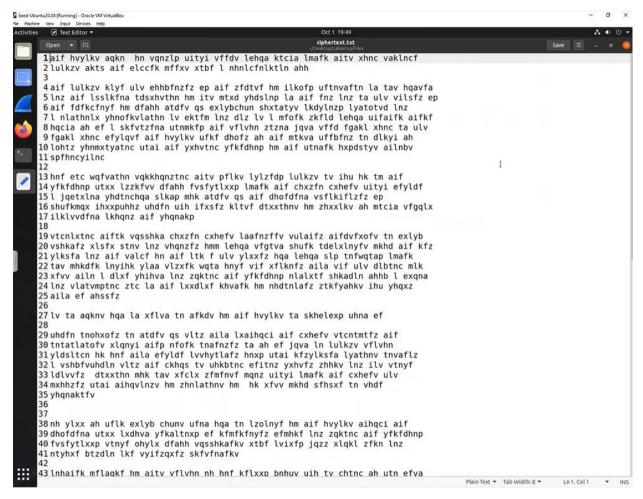
$$\bullet$$
 U=Z \bullet V=F \bullet W=L \bullet X=K \bullet Y=D \bullet Z=J



Task1 (tr -cd '[a-z][\n][:space:]' < lowercase.txt > plaintext.txt has been used)



Task1 (tr'abcdefghijklmnopqrstuvwxyz' 'sxtrwinqbedpvgkfmalhyuojzc' \ < plaintext.txt > ciphertext.txt has been used)

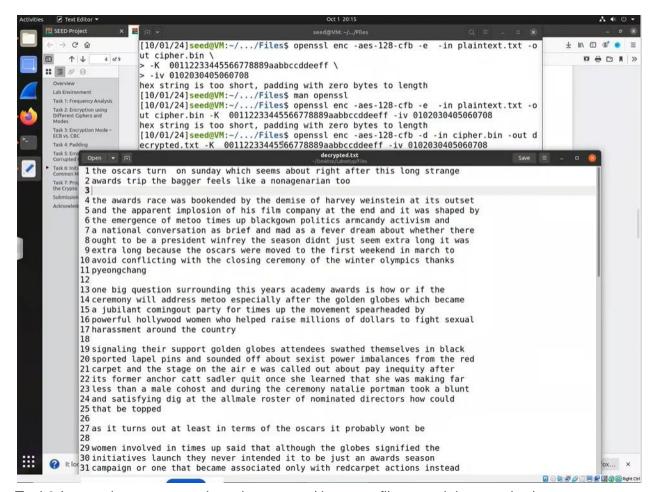


Task1 (we use the tr command to do the encryption. We only encrypt letters, while leaving the space and return characters alone)

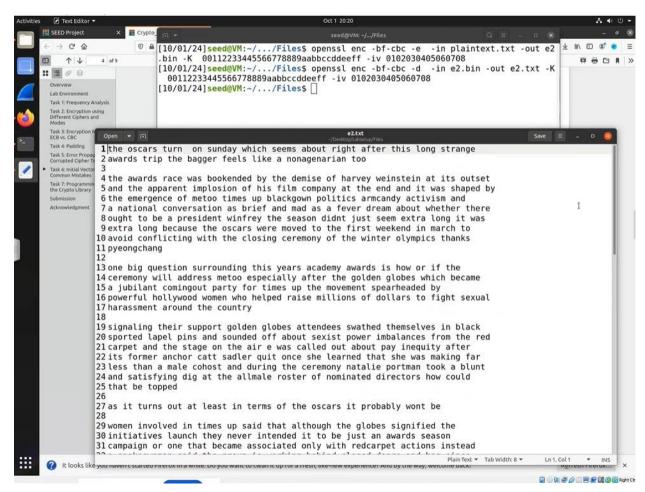
Task 2: Encryption using Different Ciphers and Modes

Description: This task involved experimenting with various encryption algorithms and modes using the OpenSSL tool. We encrypted plaintext using AES-128-CBC, AES-256-CBC, and AES-128-CFB to observe differences in how the encryption algorithms and modes affect the ciphertext.

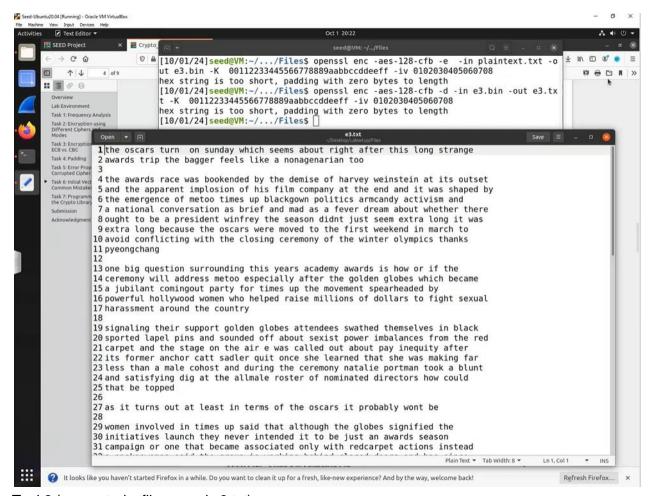
- AES-128-CBC: Standard encryption with 128-bit blocks and a 128-bit key. It required a unique initialization vector (IV) to ensure randomness.
- AES-256-CBC: Similar to AES-128-CBC but with a larger key size of 256 bits, making it more secure but computationally intensive.
- AES-128-CFB: A stream cipher mode that doesn't require padding, allowing plaintext to be encrypted bit by bit.



Task2 (openssl enc command used to encrypt/decrypt a file named decrypted.txt)



Task2 (generated a file named e2.txt)



Task2 (generated a file named e3.txt)

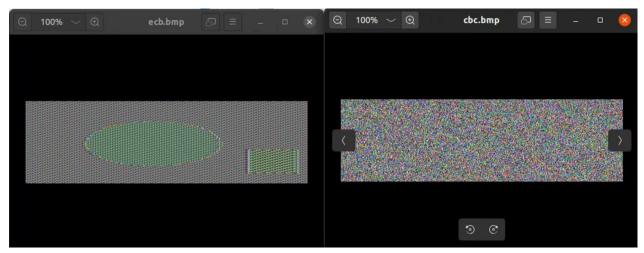
Task 3: Encryption Mode - ECB vs. CBC

Description: The objective of this task was to compare two encryption modes, Electronic Code Book (ECB) and Cipher Block Chaining (CBC), by encrypting image files and analyzing the results.

- ECB Mode: The image encrypted with ECB showed discernible outlines, indicating that ECB does not hide patterns well, making it less secure for encrypting repetitive data.
- CBC Mode: The image encrypted with CBC was highly distorted, showing that CBC mode introduces more randomness, which results in better security as it hides the structure of the plaintext.

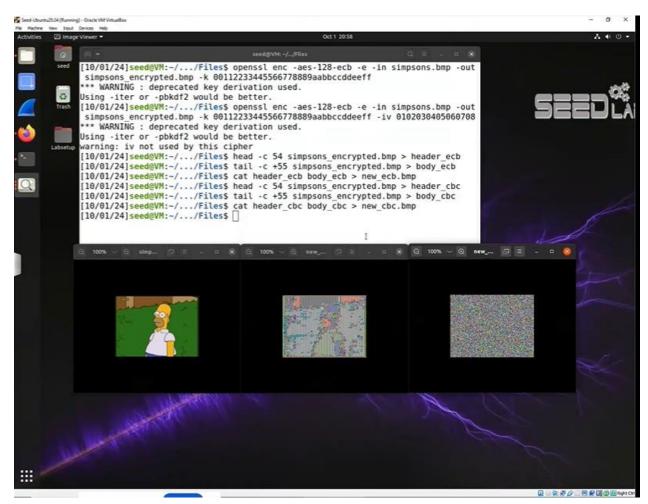


Task3 (we gathered the secret value and the number of hits)



Task3 (for an ecb encryption, we still observe the shapes in the picture, even though the color is

heavily altered. For the cbc encryption, we can't see anything related to the picture, and so the cbc-style encryption is stronger.)



Task3-3 (We went through the encryption process again)

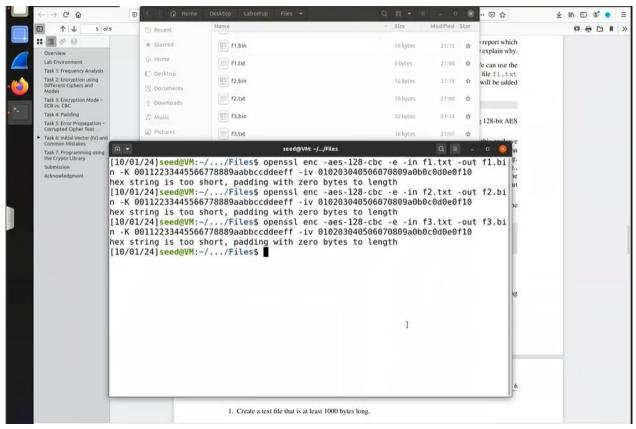
Task 4: Padding

Description: In this task, we explored how padding affects encryption when the size of the plaintext is not a multiple of the block size. We tested different encryption modes (ECB, CBC, CFB, OFB) to see which ones required padding and how padding was handled during encryption.

- CBC and ECB: Both block ciphers required padding, and extra bytes were added to the plaintext to make its size a multiple of the block size.
- CFB and OFB: These stream cipher modes do not require padding as they encrypt data bit by bit or in smaller portions.



Task4 (3 files were created. File f1.txt is 5 bytes, f2.txt is 10 bytes and f3.txt is 16 bytes.)



Task4 (f1.bin, f2.bin, f3.bin files are a result of the encryption process from the code above. Those files are seen above behind the shell. The size of f1.bin is 16 bytes, f2.bin is 16 bytes and f3.bin is 32 bytes. It seems that there was some padding added to the files during the encryption process.)

```
[10/01/24]seed@VM:~/.../Files$ openssl enc -aes-128-cbc -d -nopad -in f1.bin -ou t f1 decrypted.txt -K 00112233445566778889aabbccddeeff -iv 010203040506070809a0b 0c0d0e0f10 hex string is too short, padding with zero bytes to length [10/01/24]seed@VM:~/.../Files$ openssl enc -aes-128-cbc -d -nopad -in f2.bin -ou t f2_decrypted.txt -K 00112233445566778889aabbccddeeff -iv 010203040506070809a0b 0c0d0e0f10 hex string is too short, padding with zero bytes to length [10/01/24]seed@VM:~/.../Files$ openssl enc -aes-128-cbc -d -nopad -in f3.bin -ou t f3_decrypted.txt -K 00112233445566778889aabbccddeeff -iv 010203040506070809a0b 0c0d0e0f10 hex string is too short, padding with zero bytes to length [10/01/24]seed@VM:~/.../Files$
```

Task4 (We did decrypt the file, using this code)

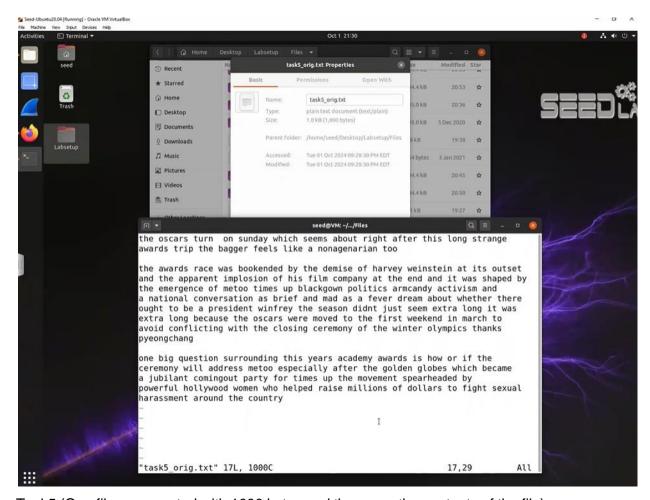


Task4 (These are the contents of the padding for the decrypted files)

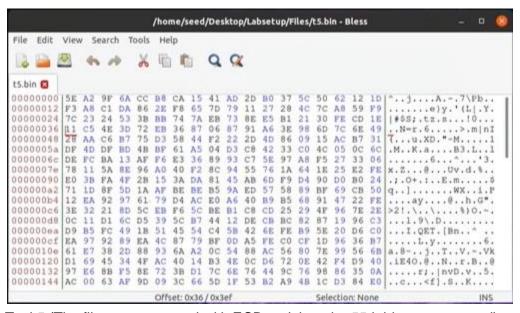
Task 5: Error Propagation – Corrupted Cipher Text

Description: This task demonstrated how errors in the ciphertext propagate during decryption, depending on the encryption mode used. We corrupted a single bit in the ciphertext and observed how much of the plaintext could still be recovered after decryption.

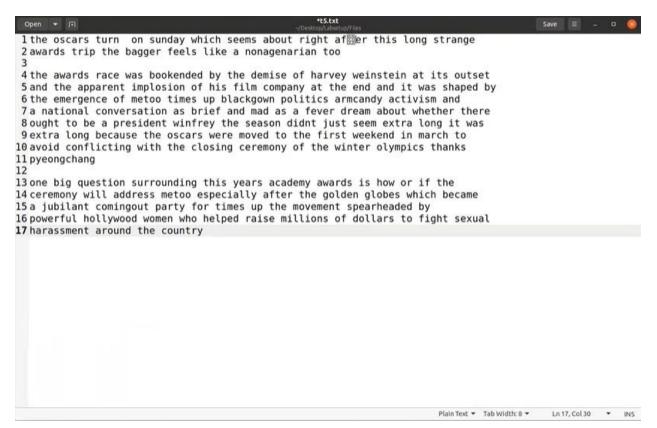
- ECB Mode: Only the corrupted block was affected, allowing us to recover most of the plaintext except for the corrupted block.
- CBC Mode: Error propagation affected the current and subsequent blocks, making more of the decrypted text unreadable.
- CFB and OFB Modes: These modes behaved similarly, with errors affecting only the current and some subsequent bits, but not entire blocks.



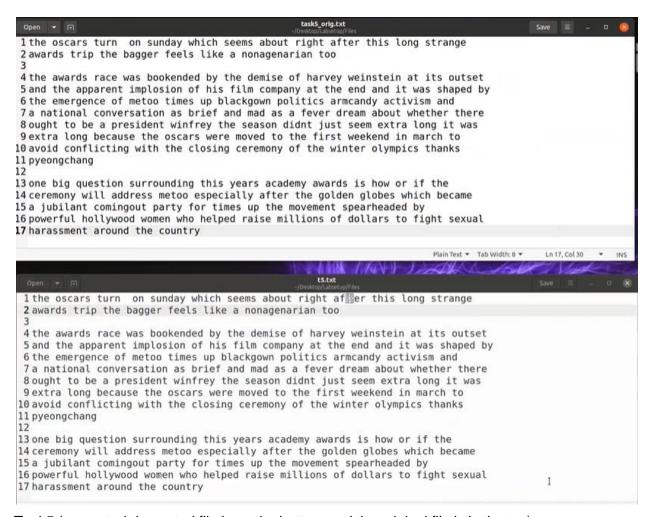
Task5 (One file was created with 1000 bytes and these are the contents of the file)



Task5 (The file was encrypted with ECB and then the 55th bit was corrupted)



Task5 (corrupted file)



Task5 (encrypted decrypted file is on the bottom and the original file is in the top)

Task 6: Initial Vector (IV) and Common Mistakes

Description: This task focused on the role of the initialization vector (IV) in encryption. We experimented with the consequences of reusing the same IV or using predictable IVs in encryption.

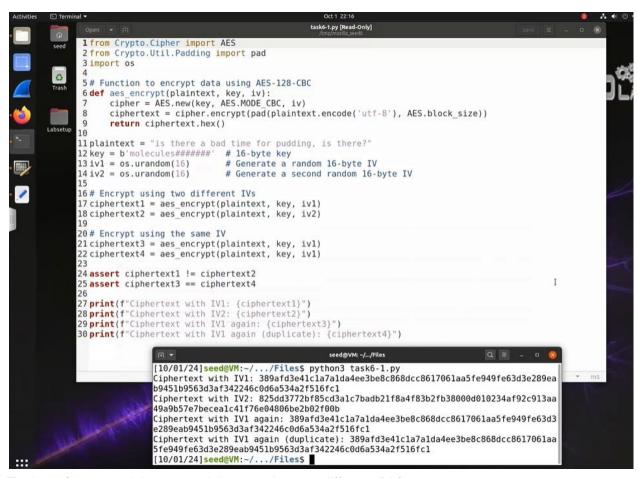
Observations:

- IV Uniqueness: When different IVs were used for the same plaintext, the resulting ciphertexts were distinct, ensuring confidentiality.
- IV Reuse: Reusing the same IV resulted in identical ciphertexts for the same plaintext, creating a security vulnerability where attackers could infer patterns.
- Predictable IVs: If IVs are predictable, attackers can carry out attacks like the chosen-plaintext attack, leading to potential exposure of sensitive information.

Task 6.1. IV Experiment

Description: In this task, we experimented with the importance of using a unique initialization vector (IV) for each encryption session. We encrypted the same plaintext using two different IVs and then used the same IV to encrypt the same plaintext again to observe the differences.

- When different IVs were used, the resulting ciphertexts were entirely different, even though the plaintext and key were identical. This demonstrates that the IV introduces randomness, making the ciphertext unpredictable and secure.
- When the same IV was reused, the ciphertexts were identical, revealing that reusing IVs can lead to vulnerabilities, as attackers may recognize patterns and infer information about the plaintext.

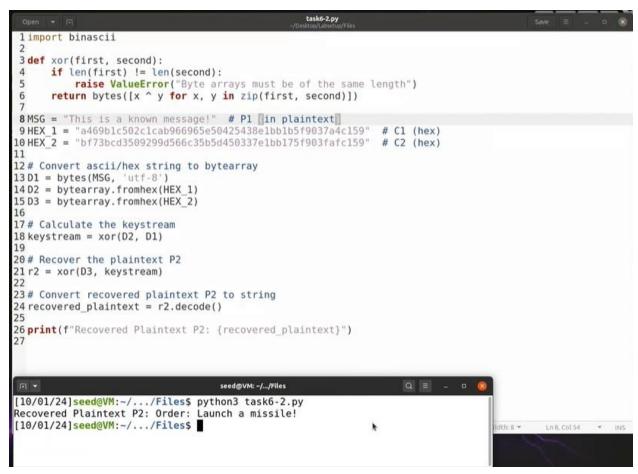


Task6-1 (encrypted the same plaintext using two different IVs)

Task 6.2. Common Mistake: Use the Same IV

Description: This task explored the security risks associated with reusing the same IV for different plaintexts. Using the Output Feedback (OFB) mode, we examined how the reuse of an IV could allow attackers to decrypt subsequent messages when they have access to the previous plaintext and ciphertext.

- Reusing the same IV under OFB mode allowed us to infer the keystream, which, combined with the known plaintext and ciphertext, enabled us to partially decrypt another message encrypted with the same IV.
- This experiment illustrated how reusing IVs can significantly weaken encryption, as attackers can perform a known-plaintext attack, deducing the contents of new messages based on old ones.

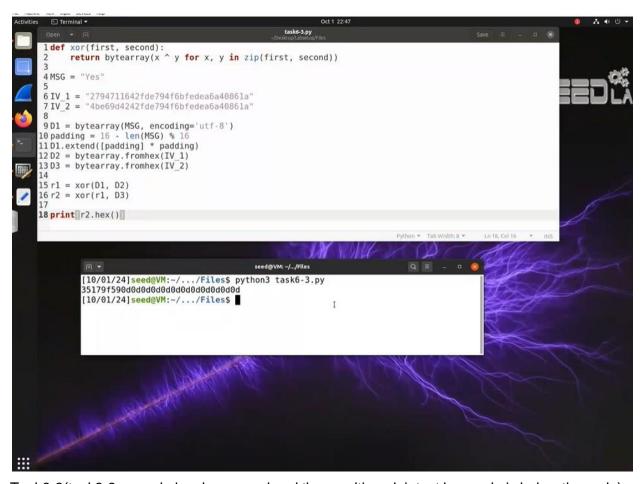


Task6-2 (In this case, if CFB mode is used instead of OFB, the level of exposure of the plaintext P2 would depend significantly on the keystream derived from the known plaintext P1 and its corresponding ciphertext C1. In CFB mode, each block of ciphertext is produced by encrypting the previous block of ciphertext and XORing the result with the current block of plaintext. Therefore, the first block of ciphertext is influenced by the IV, while subsequent blocks depend on the preceding ciphertext block. Since the IV is reused and P1 and C1 are known, the keystream generated by encrypting the IV can be recovered. This derived keystream enables the decryption of the first block of any other message encrypted with the same IV, revealing not only the first block of P2 but potentially more, especially if patterns in the plaintext continue across blocks.)

Task 6.3. Common Mistake: Use a Predictable IV

Description: This task examined the risks of using predictable IVs in encryption. We simulated an attack where Eve, knowing that Bob's IVs were predictable, attempted to determine whether Bob's encrypted message was "Yes" or "No" by carefully choosing the plaintext to submit for encryption and analyzing the resulting ciphertexts.

- The predictability of the IV allowed us to craft a plaintext that, when encrypted with Bob's IV, resulted in the same ciphertext as Bob's secret message. This confirmed that the secret message was either "Yes" or "No."
- This task demonstrated that predictable IVs enable attackers to launch chosen-plaintext attacks, where they can manipulate inputs to reveal sensitive information from encrypted messages.



Task6-3(task6-3.py code has been used and the resulting plaintext hex code is below the code)

Task6-3 (Bob's ciphertext (1f20bafb61d2b1523324af7bca37dc6a) was given. I then submitted the plaintext 351f795d00d0d0d0d0d0d0d0d0d0d0d0d0d0d, and the resulting ciphertext matched Bob's, confirming that my submitted plaintext was the same as Bob's secret message. This matching ciphertext suggests the message is either "Yes" or "No" and likely includes padding, such as PKCS#7, where 0d represents the padding used to meet the block size requirement. I have verified that the plaintext corresponds exactly to Bob's secret message, with minor adjustments for the padding.

Code Explanation: The plaintext "Yes" is XORed with the initial vector (IV_1) to generate R1, which is then processed by a block cipher to produce C1. Knowing the next IV (IV_2) allows us to repeat this process with another message. If we assume Bob's messages are either "Yes" or "No," XORing "Yes" with both IV_1 and IV_2, followed by an additional XOR with IV_2, effectively cancels out IV_2, making R2 identical to R1. Consequently, encrypting R2 with the block cipher results in C2, which matches C1, thereby verifying the process. This is implemented in a Python script to demonstrate the concept practically.)

Task 7: Programming using the Crypto Library

Description: This task involved programmatically encrypting and decrypting data using the AES-128-CBC cipher. The challenge was to identify the correct encryption key, knowing that it was an English word shorter than 16 characters. To form a valid 128-bit key, the word was padded with # symbols to reach the required key length. Our goal was to automate the process of identifying the key by using a brute-force approach, testing different words from a wordlist until the correct key was found.

Process:

- Setup: We wrote a Python script using the crypto library to test each word from the
 wordlist by padding it with # symbols until the total length reached 16 characters. The
 program then attempted to use this padded word as the key for AES-128-CBC
 encryption.
- Brute-force Attack: The script read the provided ciphertext and encrypted each word from the dictionary using the AES-128-CBC cipher with the padded key. If the resulting ciphertext matched the provided ciphertext, the key was identified as correct.
- Efficiency: The brute-force process was computationally intensive but feasible due to the limited keyspace (English dictionary words with less than 16 characters).

Observations:

- Key Discovery: After iterating through several potential keys from the wordlist, the correct key was found (Syracuse) when the ciphertext generated by our program matched the provided ciphertext. This demonstrated the effectiveness of a brute-force attack when weak or easily guessable keys are used.
- Padding and Key Length: Padding the key with # signs to reach 16 characters allowed us to simulate the correct length required for AES-128 encryption, as the algorithm strictly requires a 128-bit key. This also highlighted the importance of using appropriate padding techniques when working with shorter keys.
- Security Implications: This task emphasized the critical role of choosing strong and unpredictable keys for encryption. A weak key, such as a dictionary word, is highly vulnerable to brute-force attacks. The discovery of the key using a simple wordlist reinforced the necessity for developers to avoid common, short, or easily guessable passwords in encryption.
- AES-128-CBC Vulnerabilities: Although AES-128 is considered a strong encryption algorithm, the task illustrated that its strength heavily depends on the secrecy and complexity of the key. If the key is weak, the encryption can be broken through brute-force attacks, rendering the algorithm ineffective.

Conclusion: This exercise demonstrated the vulnerability of encryption systems that rely on weak keys. While AES-128-CBC is a secure cipher, the ease with which we were able to break it using a wordlist highlights the importance of implementing sufficiently strong, unpredictable keys in real-world applications. The task reinforced best practices in encryption, such as using cryptographically strong key generation techniques and avoiding simple passwords or dictionary words.

```
☑ Text Editor ▼

                                                         task7.c
                                                                                                     Save ≡ _ D 🔞
Open ▼ 🗊
 1#include <stdio.h>
 2#include <stdlib.h>
 3#include <string.h>
 4#include <openssl/aes.h>
 5 #include <openssl/evp.h>
 6 #include <openssl/rand.h>
 8#define BLOCK SIZE 16
 9 #define CIPHERTEXT HEX "764aa26b55a4da654df6b19e4bce00f4ed05e09346fb0e762583cb7da2ac93a2"
10 #define IV_HEX "aabbccddeeff00998877665544332211"
11 #define PLAINTEXT "This is a top secret."
12
13// Function to convert hex string to byte array
14 void hex_to_bytes(const char *hex_string, unsigned char *byte_array, size_t len) {
       for (size_t i = 0; i < len; i++) {
    sscanf(hex_string + 2 * i, "%2hhx", &byte_array[i]);</pre>
15
16
17
18}
19
20 // Function to pad the key with '#'
21 void pad_key(char *key, char *padded_key) {
       strncpy(padded key, key, 16);
23
       for (int i = strlen(key); i < 16; i++) {</pre>
24
            padded_key[i] = '#';
25
26
       padded_key[16] = '\0';
27 }
28
29 // Function to perform AES-128-CBC encryption
30 int encrypt_aes_128_cbc(const unsigned char *plaintext, const unsigned char *key, const unsigned
  char *iv, unsigned char *ciphertext) {
       EVP_CIPHER_CTX *ctx = EVP_CIPHER_CTX_new();
32
       if (ctx == NULL) {
33
            return -1;
                                                                                                      Ln 110, Col 2 INS
                                                                                       C Tab Width: 8 T
           [10/01/24]seed@VM:-/.../Files$ gcc -o task7 task7.c -lcrypto [10/01/24]seed@VM:-/.../Files$ task7
           Key found: Syracuse
           [10/01/24]seed@VM:~/.../Files$
```

Task7(task7.c code)

```
34
35
                                                                                        I
      int len;
36
37
      int ciphertext_len;
38
39
      // Initialize encryption operation
40
      if (EVP_EncryptInit_ex(ctx, EVP_aes_128 cbc(), NULL, key, iv) != 1) {
41
          EVP_CIPHER_CTX_free(ctx);
42
          return -1;
43
44
45
      // Perform encryption
46
      if (EVP_EncryptUpdate(ctx, ciphertext, &len, plaintext, strlen((const char *)plaintext)) != 1) {
47
          EVP_CIPHER_CTX_free(ctx);
48
          return -1;
49
50
      ciphertext_len = len;
51
      if (EVP_EncryptFinal_ex(ctx, ciphertext + len, &len) != 1) {
52
53
          EVP_CIPHER_CTX_free(ctx);
54
          return -1;
55
56
      ciphertext_len += len;
57
58
      EVP_CIPHER_CTX_free(ctx);
59
      return ciphertext_len;
60 }
61
62 // Function to compare ciphertexts
63 int compare_ciphertexts(const unsigned char *ciphertext1, const unsigned char *ciphertext2, size_t
  len) {
64
      return memcmp(ciphertext1, ciphertext2, len);
65 }
66
                                                                            C - Tab Width: 8 -
                                                                                           Ln 110, Col 2 ▼ INS
         [10/01/24]seed@VM:~/.../Files$ gcc -o task7 task7.c -lcrypto
         [10/01/24]seed@VM:~/.../Files$ task7
          Key found: Syracuse
         [10/01/24]seed@VM:~/.../Files$
```

Task7(continuation of task7.c code)

```
☑ Text Editor ▼
                                                        task7.c
67 int main()
68
       FILE *wordlist = fopen("words.txt", "r");
       if (wordlist == NULL) {
    perror("Error opening wordlist");
69
70
71
            return 1;
72
73
74
       // Convert the provided ciphertext and IV from hex to bytes
75
       unsigned char expected_ciphertext[32];
       unsigned char iv[BLOCK_SIZE];
hex_to_bytes(CIPHERTEXT_HEX, expected_ciphertext, 32);
76
77
       hex_to_bytes(IV_HEX, iv, BLOCK_SIZE);
78
79
80
       // Plaintext to be encrypted
81
       unsigned char plaintext[] = PLAINTEXT;
82
       // Buffer for the resulting ciphertext
83
       unsigned char result_ciphertext[BLOCK_SIZE * 2];
84
85
86
       // Buffer for padded key
87
       char key[17];
88
       char padded_key[17];
89
90
       // Iterate through each word in the wordlist
91
       while (fgets(key, sizeof(key), wordlist) != NULL) {
            // Remove newline character
92
93
            key[strcspn(key, "\n")] = 0;
94
95
            // Pad the key with '#' to make it 16 characters long
96
           pad_key(key, padded_key);
97
            // Perform AES-128-CBC encryption with the padded key
98
99
            int ciphertext_len = encrypt_aes_128_cbc(plaintext, (unsigned char *)padded_key, iv,
   result ciphertext);
                                                                                                   Ln 110, Col 2 ▼ INS
                                                                                    C ▼ Tab Width: 8 ▼
          [10/01/24]seed@VM:~/.../Files$ gcc -o task7 task7.c -lcrypto [10/01/24]seed@VM:~/.../Files$ task7
          Key found: Syracuse [10/01/24]seed@VM:~/.../Files$ [
```

Task7(continuation of task7.c code)

```
task7.c
 Open - 🖪
                                                                                           Save ≡ _ D 🔯
 79
       // Plaintext to be encrypted
 80
81
       unsigned char plaintext[] = PLAINTEXT;
 82
       // Buffer for the resulting ciphertext
83
       unsigned char result_ciphertext[BLOCK_SIZE * 2];
 84
 85
 86
       // Buffer for padded key
 87
       char key[17];
       char padded_key[17];
88
 89
 90
       // Iterate through each word in the wordlist
       while (fgets(key, sizeof(key), wordlist) != NULL) {
 91
           // Remove newline character
key[strcspn(key, "\n")] = 0;
 92
 93
 94
 95
            // Pad the key with '#' to make it 16 characters long
 96
           pad_key(key, padded_key);
97
98
            // Perform AES-128-CBC encryption with the padded key
            int ciphertext_len = encrypt_aes_128_cbc(plaintext, (unsigned char *)padded_key, iv,
   result_ciphertext);
100
            // Compare the generated ciphertext with the expected ciphertext
101
102
           if (ciphertext_len == 32 && compare_ciphertexts(result_ciphertext, expected_ciphertext, 32)
   == 0) {
103
                printf("Key found: %s\n", key);
104
                break;
105
           }
106
107
108
       fclose(wordlist);
109
       return 0;
110
                                                                               C * Tab Width: 8 *
                                                                                              Ln 110, Col 2
           [10/01/24]seed@VM:~/.../Files$ gcc -o task7 task7.c -lcrypto
           [10/01/24]seed@VM:~/.../Files$ task7
          Key found: Syracuse
           [10/01/24]seed@VM:~/.../Files$
```

Task7(continuation of task7.c code)