Data Security & Privacy
Project 3
Group 20

Raghu Pusapati

Venkata Koyyalamudi

Solomon Sukumar Pulavarhi

Most of the code has been taken from the AES project. For keygen, we used SHA256. We generated a random 256-bit string and performed SHA256 on it. For encryption, we used AES in CBC mode. We encrypted the files using AES and stored them in the ciphertextfiles folder. For PRF, we used SHA256. The keywords have been hashed using SHA256. These hashed keywords are used in forming the inverted index file. Before building an inverted index, we build an index to invert it. For token generation, we used SHA256. The input keyword is hashed and is stored as a token. Finally, for searching, we used the inverted index to find out the files associated and then decrypted them using AES.

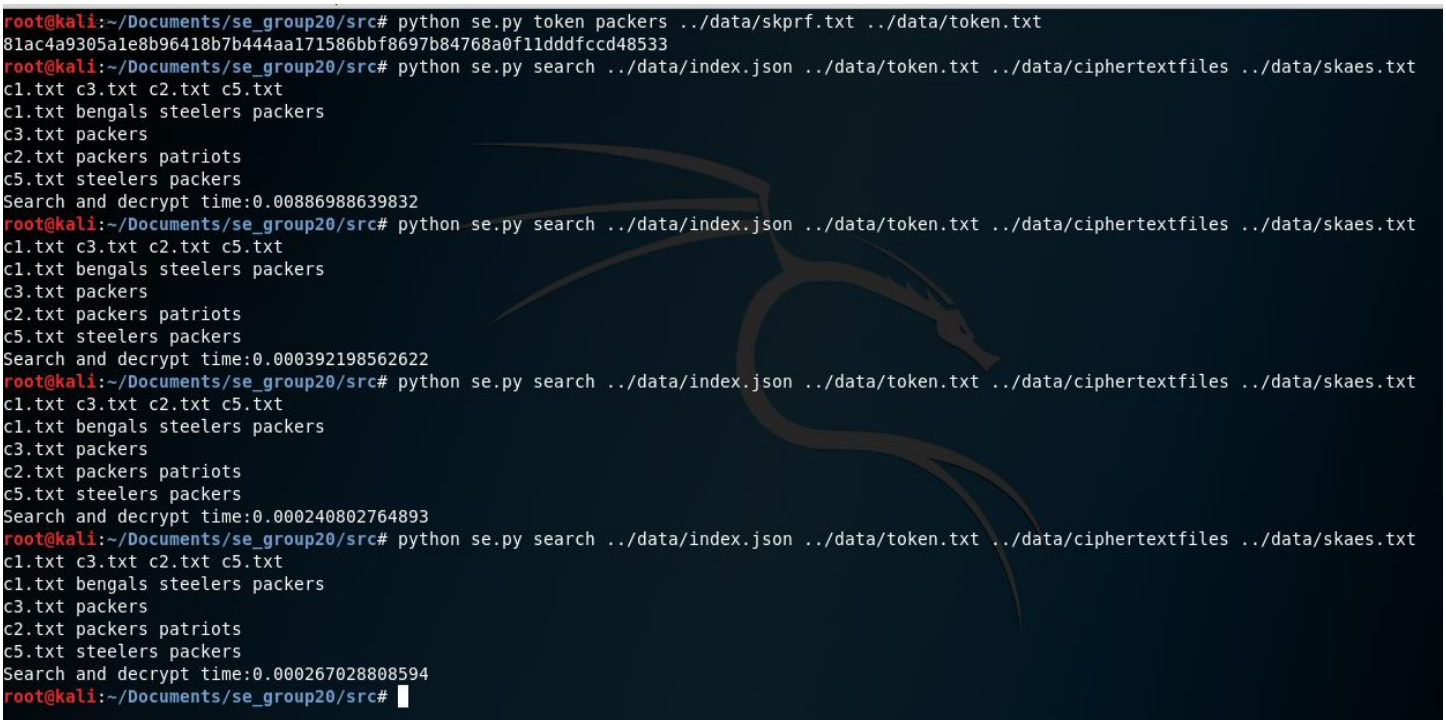
The following is a sample run of the above four functions

```
root@kali:~/Documents/se_group20/src# python se.py keygen ../data/skprf.txt ../data/skaes.txt
e577fd951fab8849dcd32940f79e93fc09a4b3072b10e0c3a0cd5f35675670f8
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src# python se.py token packers ../data/skprf.txt ../data/token.txt
81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533
root@kali:~/Documents/se_group20/src# python se.py search ../data/index.json ../data/token.txt ../data/ciphertextfiles ../data/skaes.txt
c1.txt c3.txt c2.txt c5.txt
c1.txt bengals steelers packers
c3.txt packers
c2.txt packers patriots
c5.txt steelers packers
root@kali:~/Documents/se_group20/src#
```

Furthermore, we have calculated the encryption and search times for building the 6 files and for ‘packers’ search.

```
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
Encryption and build time:0.000756978988647
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
Encryption and build time:0.000747919082642
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
Encryption and build time:0.000639915466309
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
Encryption and build time:0.00101399421692
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src# python se.py enc ../data/skprf.txt ../data/skaes.txt ../data/index.json ../data/files ../data/ciphertextfiles
Encryption and build time:0.00088906288147
{'462aec2ca8883242483416b27326b2258b9ae4d116aef0c55c38b1820d07355e': ['c2.txt'], '0bfe298af0976c9460d159cd802af953354c3c15344ef3240ceb16a15d23819a': [
'c1.txt', 'c6.txt', 'c4.txt'], '81ac4a9305a1e8b96418b7b444aa171586bbf8697b84768a0f11dddfccd48533': ['c1.txt', 'c3.txt', 'c2.txt', 'c5.txt'], 'aae0a09f
f5c9b924621f84fecff5989846af92b22624bbfca50eadb94d7e5aeb': ['c1.txt', 'c5.txt', 'c4.txt']}
root@kali:~/Documents/se_group20/src#
```

The average encryption and build time is 0.00080952 seconds



The average search time is 0.002442 seconds. The first run seems to be slow, but the rest of the runs are quicker than encryption. We would say that search is quicker than encrypting and building.

Platform details:

OS: Kali Linux

Language: Python

Libraries installed: PyCrypto (linux only)