ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC - KỸ THUẬT MÁY TÍNH

MẬT MÃ VÀ ANH NINH MẠNG

Đề tài

Ứng dụng mã hóa dữ liệu

GVHD: Nguyễn Hữu Hiếu
SV: Nguyễn Trần Lê Minh - 1511003
Lê Duy Thanh - 1512990
Nguyễn Xuân Nam - 1512098

TP. HỒ CHÍ MINH, THÁNG 03/2019

Mục lục

Mã hóa là phương pháp bảo vệ dữ liệu cá nhân nhạy cảm trên máy tính của bạn. Việc mã hóa còn ngăn chặn bất cứ ai đọc dữ liệu của bạn khi bạn gửi thông tin qua mạng hay đồng bộ lên máy chủ, cloud,...

Trong bài tập lớn này, nhóm sẽ thực hiện một số giải thuật mã hóa dễ cho các tập tin trong máy được an toàn.

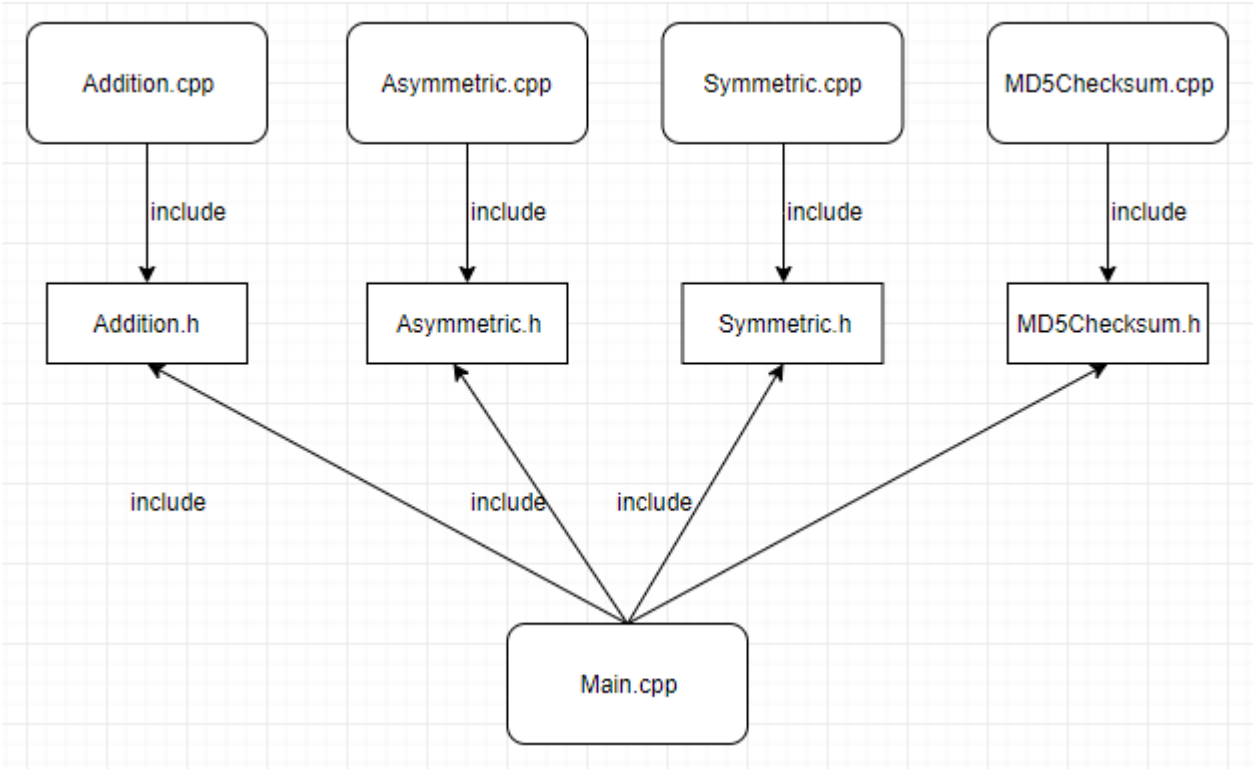
1 Giới thiệu đề tài

Nhóm đã hiện thực ba giải thuật mã hóa phổ biến đó là RSA, DES và Steganography.

Với RSA, nhóm đọc dữ liệu cần mã hóa (plaint text) từ tệp văn bản (*.txt file - text file).
Với Steganography, nhóm đọc dữ liệu từ text file trộn vào một ảnh mang.
Với DES, nhóm có thể mã hóa các tệp cơ bản như text, mp4, pdf, png,... bằng key có sẵn từ tệp.
Để chứng minh plaint text trùng với dữ liệu được giải mã, nhóm sử dụng hàm băm để đối chiếu hai tệp.

2 Cấu trúc của ứng dụng

Cấu trúc của ứng dụng được miêu tả trong hình 1: Chương trình gồm 5 mô đun chính là:



Hình 1: Các mô đun của chương trình

- Symmetric: Chứa các hàm mã hóa đối xứng tập tin sử dụng giải thuật DES.
- Asymmetric: Chứa các hàm mã hóa bất đối xứng tập tin sử dụng giải thuật RSA.
- Addition: Chứa các hàm mã hóa tập tin sử dụng kỹ thuật giấu tin (Steganography).
- MD5: Hàm MD5 sử dụng cho việc so sánh hai tệp.
- Main: Đọc các đối số và gọi các hàm phù hợp trong bốn thư viện trên.

2.1 Symmetric

Giải thuật mã hóa đối xứng nhóm sử dụng là DES với ECB mode. Trong tệp "Symmetric.h":

```
char *stringFromFile(char filename[], const char type[]);
```

Trả về chuỗi giá trị trong tệp filename.
type: chế độ đọc tệp ("r", "rb", ...).
Lưu ý: Cần giải phóng vùng nhớ cho con trỏ trả về do có sử dụng malloc trong hàm này.

```
unsigned char *DES_encrypt(EVP_CIPHER_CTX *en, unsigned char *plaintext, int *plain_l
```

Mã hó plaintext với độ dài plain_len theo giải thuật được quy định trong (*en). Trả về giá trị của chuỗi mã hóa. Lưu ý: Cần giải phóng vùng nhớ cho con trỏ trả về do có sử dụng malloc trong hàm này.

```
char *DES_decrypt(EVP_CIPHER_CTX *de, unsigned char *ciphertext, int *cipher_len);
```

Ngược với hàm trên, hàm này giải mã ciphertext với độ dài cipher_len.
Giá trị của (*de) quy định giải thuật sử dụng để giải mã.
Giá trị trả về là plaintext.
Lưu ý: Cần giải phóng vùng nhớ cho con trỏ trả về do có sử dụng malloc trong hàm này.

2.2 Asymmetric

Giải thuật mã hóa bất đối xứng nhóm chọn là RSA.Trong tệp "Asymmetric.h":

```
RSA * create_RSA(RSA *keypair, int pem_type, char *file_name);
```

Được sử dụng để tạo ra cặp key (private và public key) cho việc mã hóa và giải mã.
Key được lưu trong tệp file_name.

```
int public_encrypt(int flen, unsigned char* from, unsigned char *to, RSA* key, int pa
```

Sử dụng public key để mã hóa chuỗi (*from) với độ dài flen và lưu chuỗi mã hóa vào (*to). Trả về độ dài của chuỗi mã hóa.

```
int private_decrypt(int flen, unsigned char* from, unsigned char *to, RSA* key, int p
```

Sử dụng private key để giải mã chuỗi (*from) với độ dài flen và lưu vào chuỗi (*to).
Trả về giá trị độ dài của plaintext.

2.3 Addition

Giải thuật mã hóa ngoài bài giảng nhóm sử dụng là kỹ thuật che giấu tập tin. Trong tệp "Addition.h":

```
void steganography_encode(char* inputfile, Mat image, char*outputfile);
```

Dọc tệp inputfile dưới dạng bit.
Sử dụng phép tính "and" bit cho mỗi bit trong inputfile ứng với giá trị bit cuối trong mỗi kênh màu, pixels của ảnh image. Từ đó thu được một ảnh có sai khác rất nhỏ so với ảnh gốc.
Ghi ảnh kết quả vào tệp outputfile.

```
void steganography_decode(Mat image, char*outputfile);
```

Từ ảnh image, rút ra thông điệp và lưu nó vào outputfile.

2.4 MD5

Trong tệp "MD5Checksum.h":

```
unsigned char *getMd5Hash(unsigned char *data, size_t dataLen, int *mdLen);
```

Trả về giá trị Md5Hash của data với độ dài dataLen và độ dài của Md5 được lưu trong mdLen.
Lưu ý: Cần giải phóng vùng nhớ của con trỏ trả về.

2.5 Main

Hàm main chịu trách nhiệm đọc khác đối số khi chạy chương trình. Từ đó quyết định thực hiện giải thuật mã hóa nào. Nhóm có một số định nghĩa cho đối số đầu tiên như sau:

```
#define SYMMETRIC_ENCODE        10
#define SYMMETRIC_DECODE        11
#define RSA_CREATE                      20
#define RSA_ENCODE                      21
#define RSA_DECODE                      22
#define STREGAN_ENCODE          30
#define STREGAN_DECODE          31
#define MD5_CHECKSUM            40
```

Từ đây, với mỗi tùy chọn của đối số đầu tiên mà người dùng có thể lựa chọn tạo key, giải mã, mã hóa hay so sánh hai tệp.

Ngoài ra, trong hàm main, nhóm có gọi hàm clock() trước và sau khi thực hiện chương trình để đo thời gian mã hóa và giải mã.

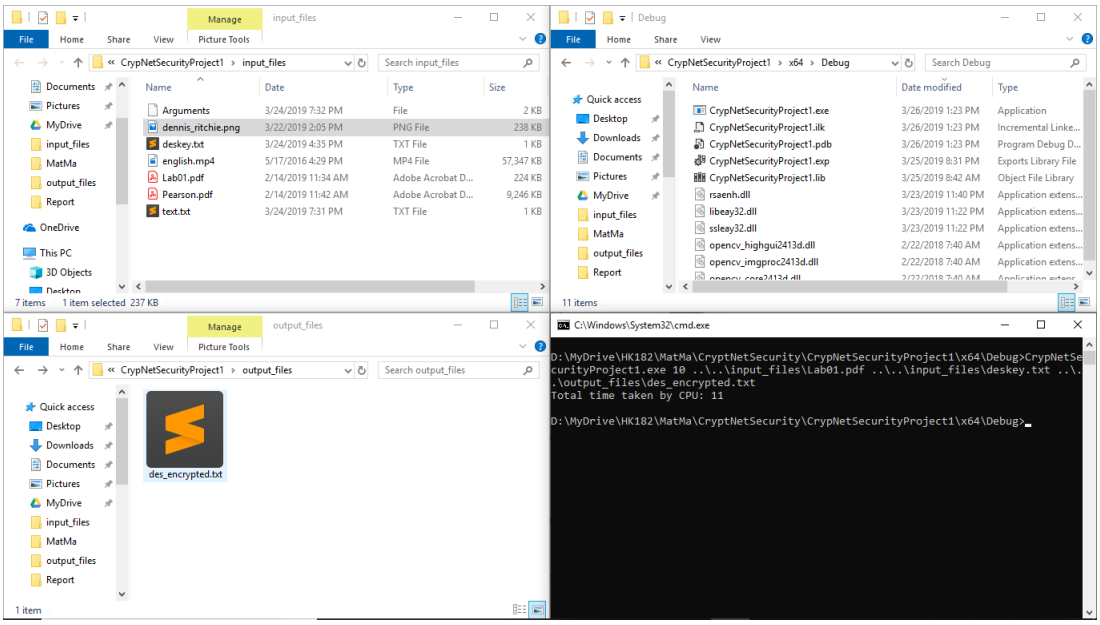
3 Chạy thử chương trình

3.1 Mã hóa DES

Chạy chương trình với command:
$File.exe 10 inputfile inputkeyfile outputfile
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.
- 10 : để lựa chọn thực thi mã hóa DES.
- inputfile: đường dẫn đến tệp cần mã hóa.
- inputkeyfile: đường dẫn đến tệp chứa key mã hóa.
- outputfile: đường dẫn đến tệp kết quả.

Ví dụ hình 2



Hình 2: Mã hóa DES

Từ hình 2, ta thấy mất 11ms để mã hóa tệp Lab01.pdf và tệp mã hóa (des_encrypted.txt) được tạo.

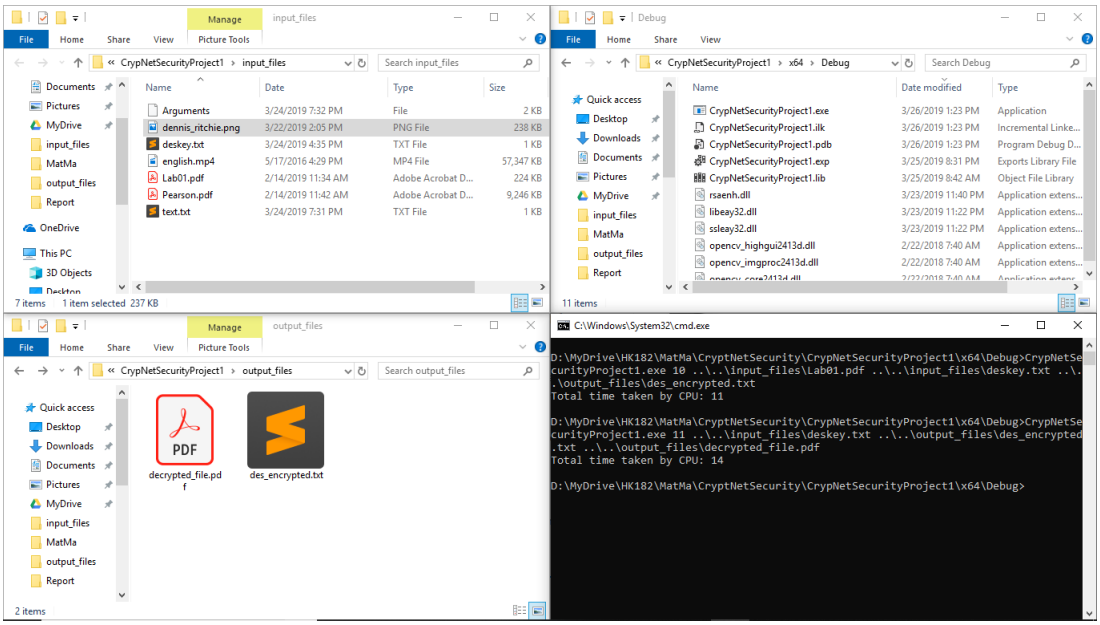
3.2 Giải mã DES

Chạy chương trình với command:
$File.exe 11 inputkeyfile inputfile outputfile
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.

- 11 : để lựa chọn thực thi giải mã DES.
- inputfile: đường dẫn đến tệp cần giải mã.
- inputkeyfile: đường dẫn đến tệp chứa key mã hóa.
- outputfile: đường dẫn đến tệp kết quả.

Ví dụ hình 3 Từ hình 3, ta thấy tệp des_encrypted.txt được giải mã trong vòng 14ms và tạo



Hình 3: Giải mã DES

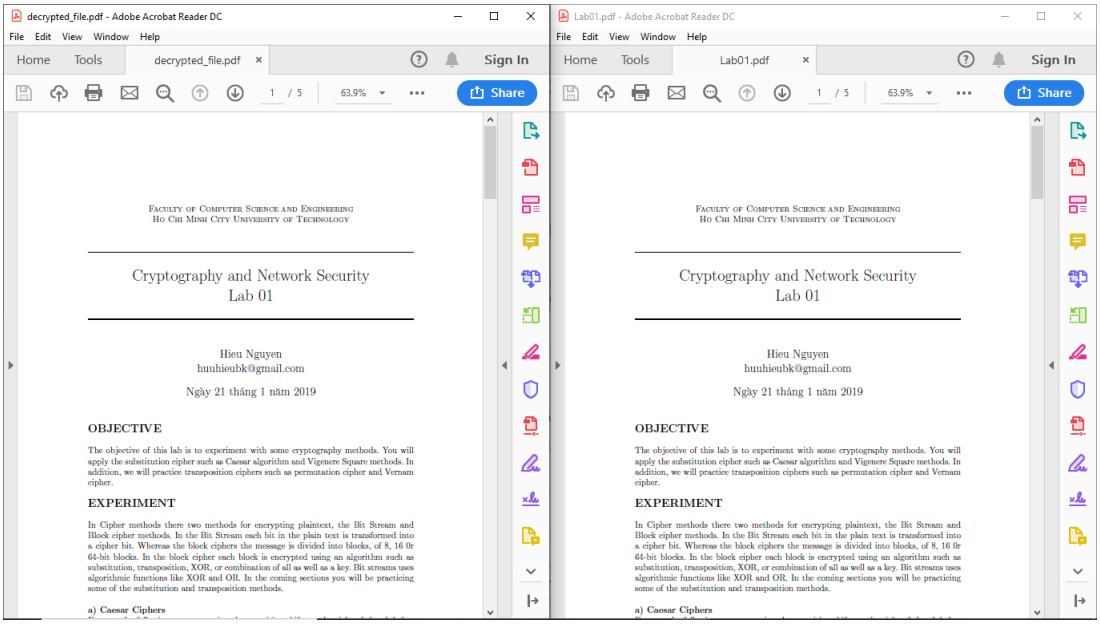
ra tệp giải mã decrypted_file.pdf. Kết quả của giải thuật DES đã tạo ra tệp giải mã và được miêu tả ngắn gọn trong hình 4.

3.3 Tạo khóa RSA

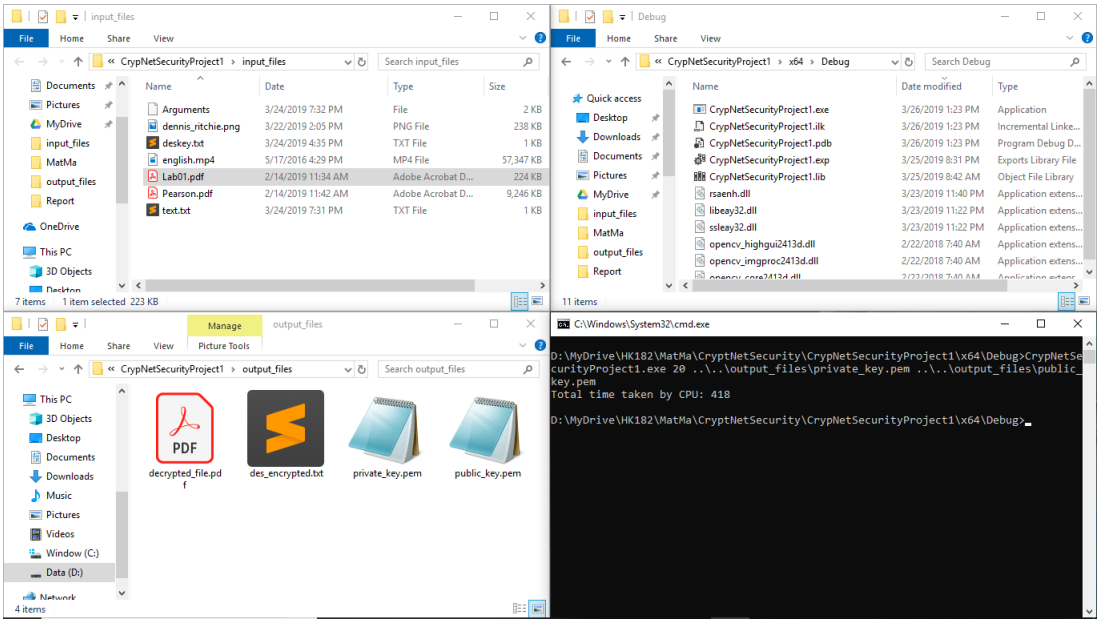
Chạy chương trình với command:
$File.exe 20 privateouput publicoutput
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.
- 20 : để lựa chọn thực thi tạo khóa bí mật và công khai.
- privateouput: đường dẫn đến tệp chứa khóa bí mật.
- publicoutput: đường dẫn đến tệp chứa khóa công khai.
- outputfile: đường dẫn đến tệp kết quả.

Ví dụ hình 5. Mất 418ms để tạo ra cặp khóa bí mật và công khai. Hai khóa trên ghi vào hai tệp: private_key.pem và public_key.pem.



Hình 4: Kết quả của giải thuật DES



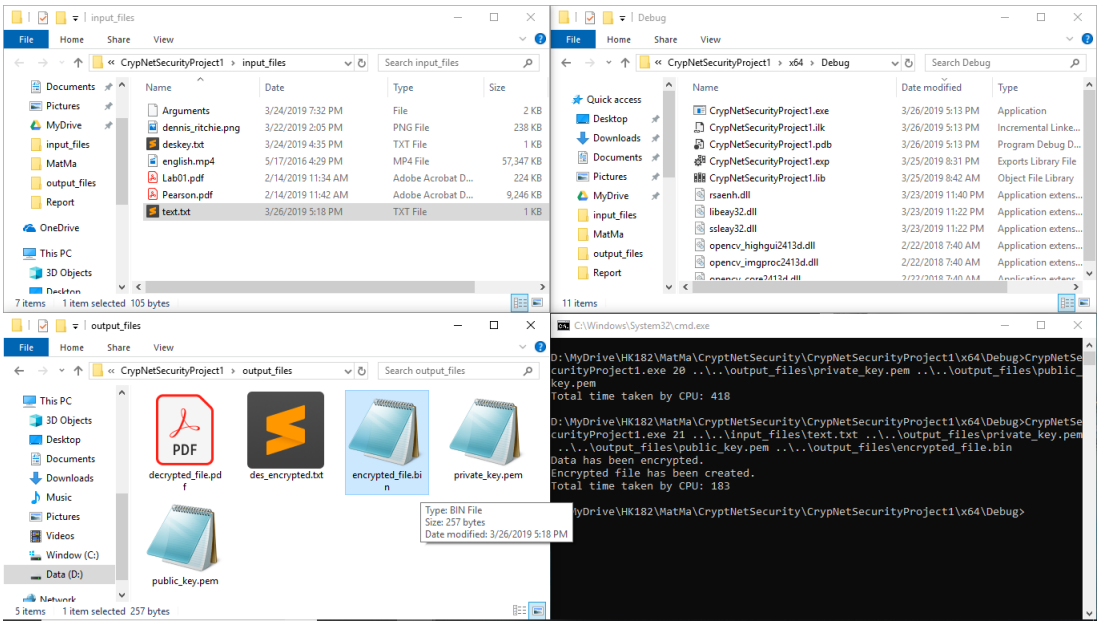
Hình 5: Tạo cặp khóa công khai và bí mật

3.4 Mã hóa RSA

Chạy chương trình với command:
$File.exe 21 inputfile privateouput publicoutput outputfile
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.
- 21 : để lựa chọn thực thi tạo khóa bí mật và công khai.
- inputfile: đường dẫn đến tệp cần mã hóa.
- privateouput: đường dẫn đến tệp chứa khóa bí mật.
- publicoutput: đường dẫn đến tệp chứa khóa công khai.

Ví dụ hình 6. Mất 183ms để mã hóa tệp text.txt. Kết quả mã hóa được ghi xuống tệp encrypted_file.bin



Hình 6: Tạo cặp khóa công khai và bí mật

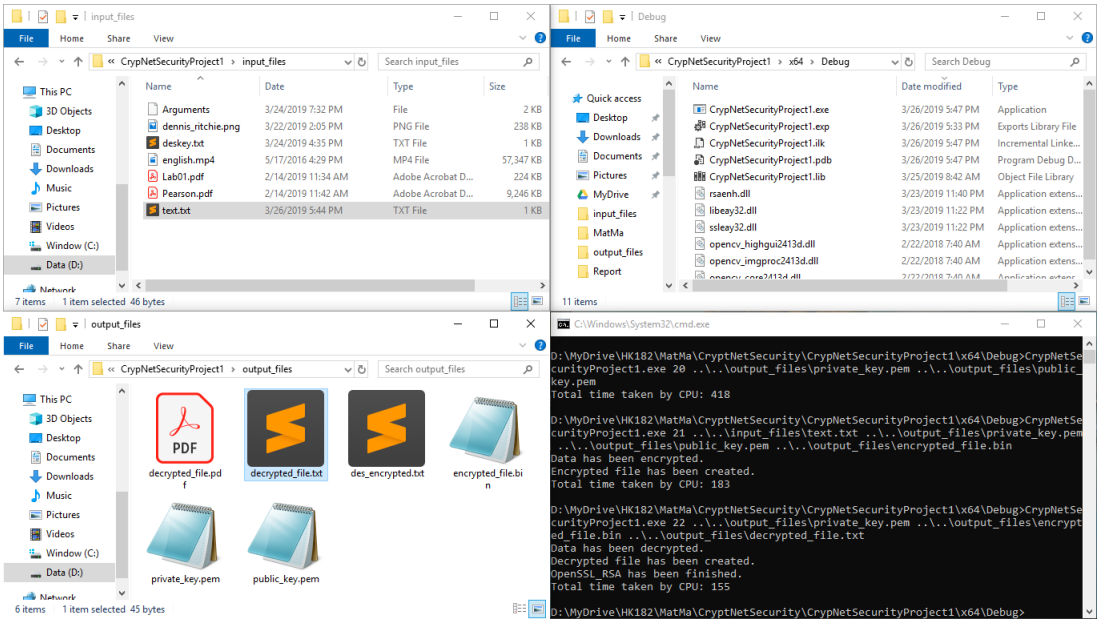
3.5 Giải mã RSA

Chạy chương trình với command:
$File.exe 22 privateinput inputfile outputfile
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.
- 22 : để lựa chọn thực thi mã hóa tệp theo giải thuật RSA.
- privateinput: đường dẫn đến tệp chứa khóa bí mật.

- inputfile: đường dẫn đến tệp cần giải mã.
- outputfile: đường dẫn đến tệp kết quả của phép giải mã.

Ví dụ hình 7. Mất 155ms để giải mã tệp encrypted_file.bin. Kết quả mã hóa được ghi xuống tệp decrypted_file.txt.
Hình 8 thể hiện trực quan so sánh giữa tệp ban đầu và tệp sau mã hóa.



Hình 7: Tạo cặp khóa công khai và bí mật

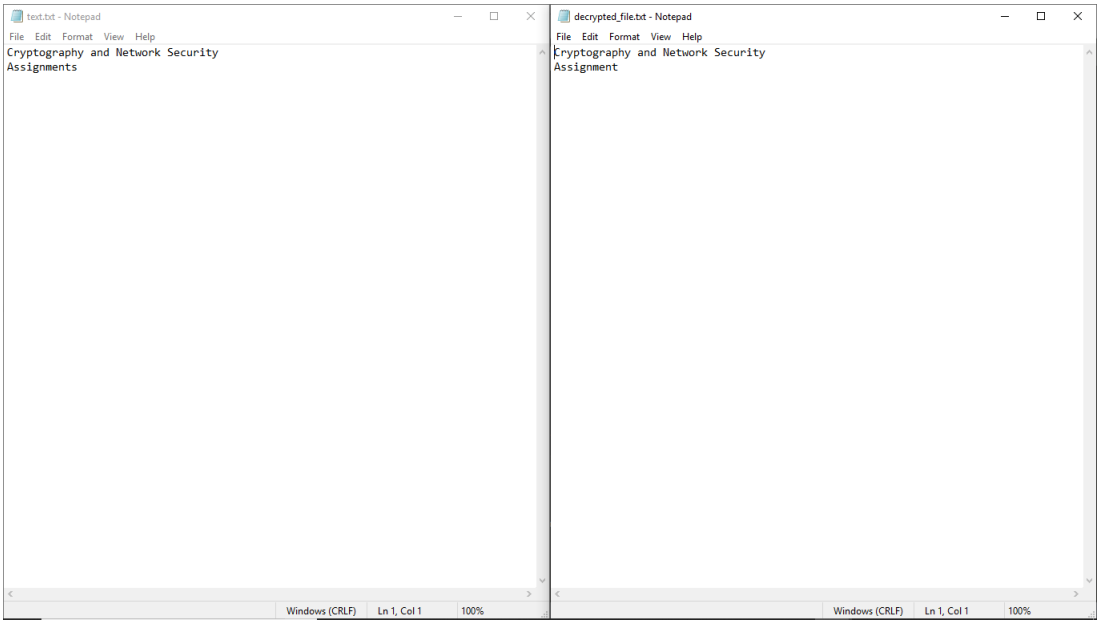
3.6 Kỹ thuật giấu tin

Chạy chương trình với command sau để gắn thông điệp vào ảnh mang:
$File.exe 30 inputfile1 inputfile2 outputfile
Trong đó:

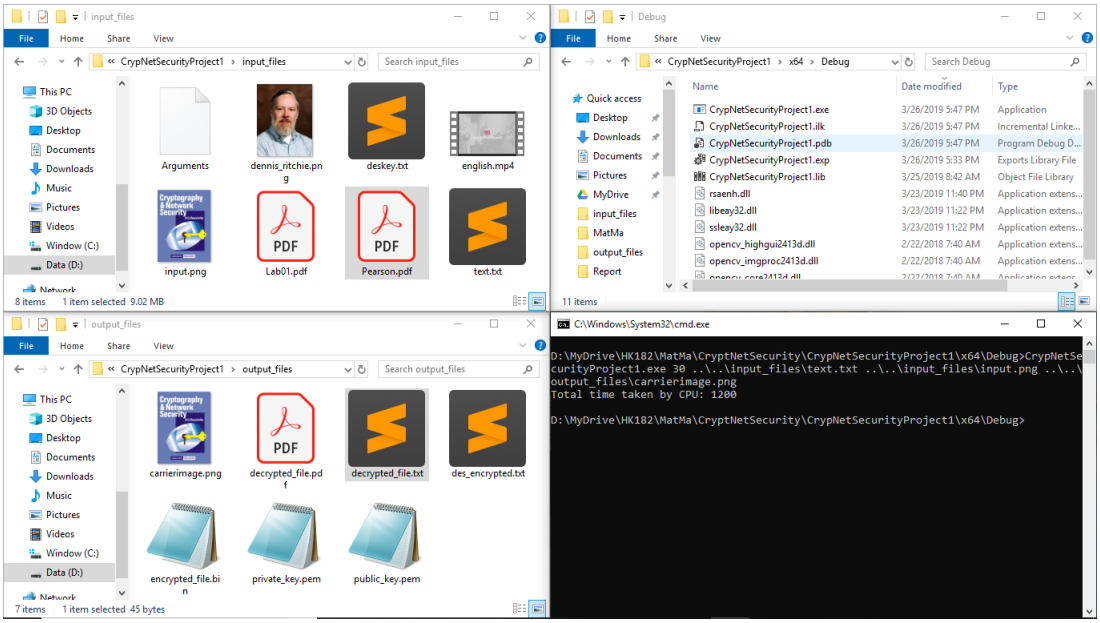
- File.exe : tệp thực thi được tạo từ MSVC.
- inputfile1: đường dẫn đến tệp chứa thông điệp.
- inputfile2: đường dẫn đến tệp chứa ảnh mang.
- outputfile: đường dẫn đến tệp kết quả của phép giấu tin.

Ví dụ hình 9
Chạy chương trình với command sau để gắn thông điệp vào ảnh mang:
$File.exe 30 inputfile outputfile
Trong đó:

- File.exe : tệp thực thi được tạo từ MSVC.
- inputfile: đường dẫn đến tệp chứa ảnh mang thông điệp.



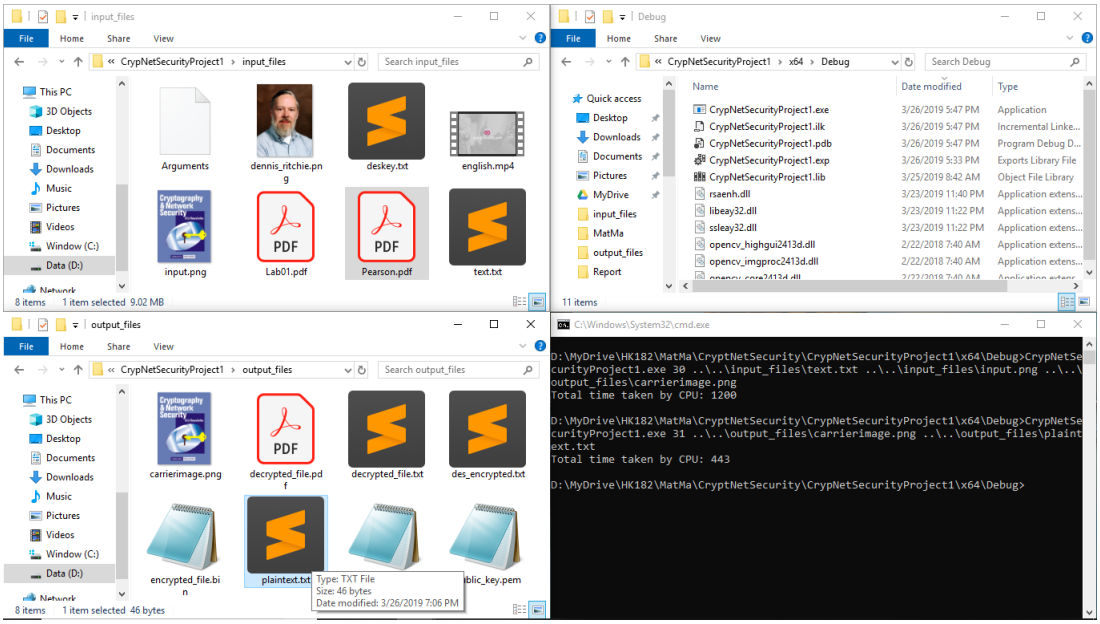
Hình 8: Tạo cặp khóa công khai và bí mật



Hình 9: Gắn thông điệp vào ảnh mang

- outputfile: đường dẫn đến tệp giải mã.

Ví dụ hình 10



Hình 10: Lấy thông điệp từ ảnh mang

Kết quả của quá trình giải mã và mã hóa được thể hiện qua hình 11

4 Hướng dẫn vận hành trên MSVC

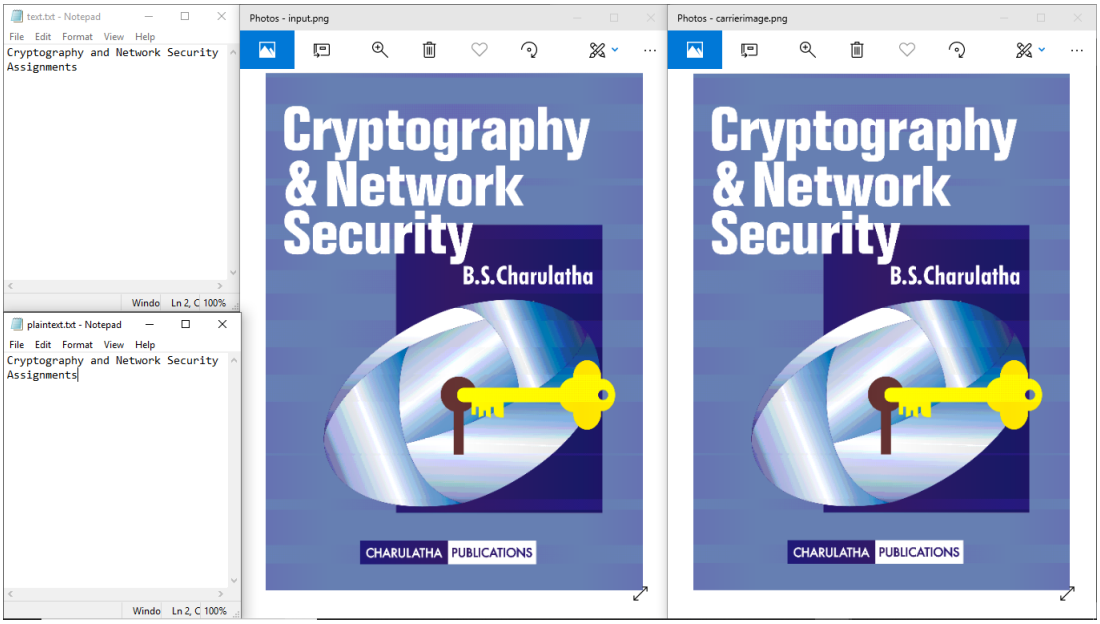
Bước 1: Cài đặt MSVC.
Bước 2: Cài đặt opencv.
Bước 3: Cài đặt opensll.
Bước 4: Tạo project trên MSVC.
Bước 5: Thiết lập đường dẫn đến thư viện opencv và openssl như hình 12 13 và 14
Bước 6: Tải mã nguồn tại urlhttps://github.com/nxnam714/CryptNetSecurity.git và gom vào project.

5 Phân tích và kết luận

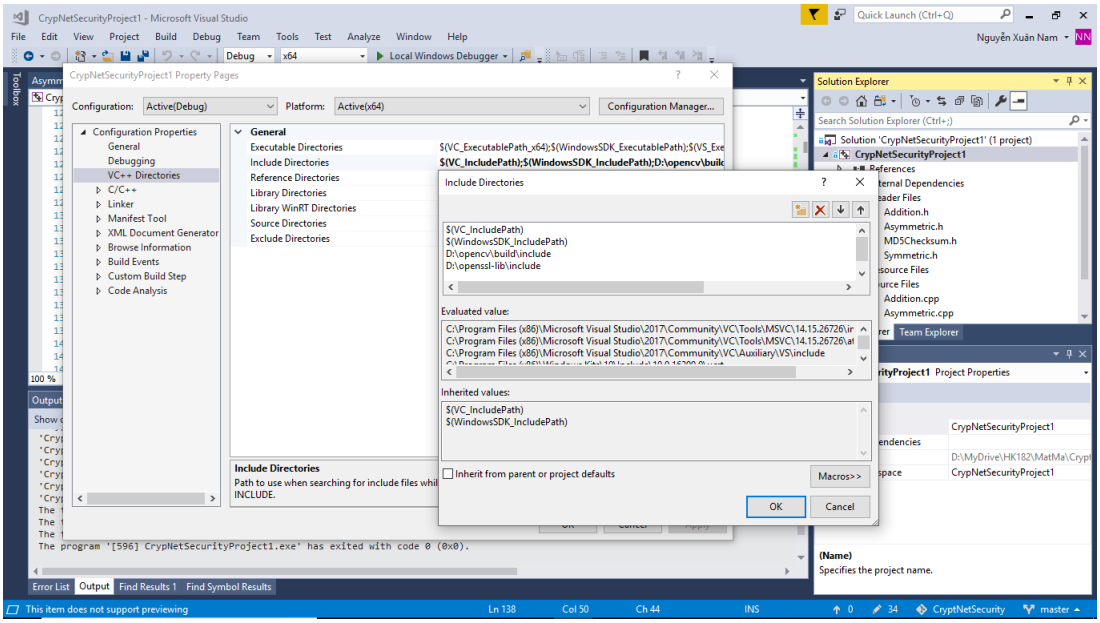
Ứng dụng có thể chạy tốt ba giải thuật mã hóa.

- Kết quả giải mã là đồng nhất với dữ liệu đưa vào dựa vào hàm kiểm tra MD5.
- Có thể mã hóa các tệp cơ bản như pdf, png, txt,...
- Thời gian thực thi nhanh.

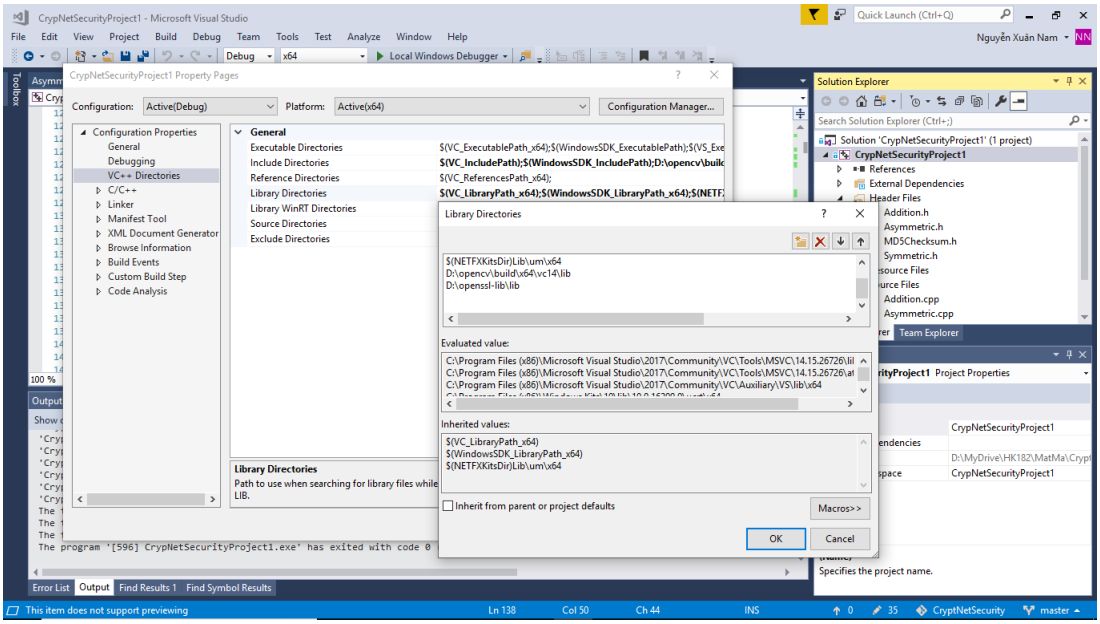
Tuy nhiên còn một số khuyết điểm:



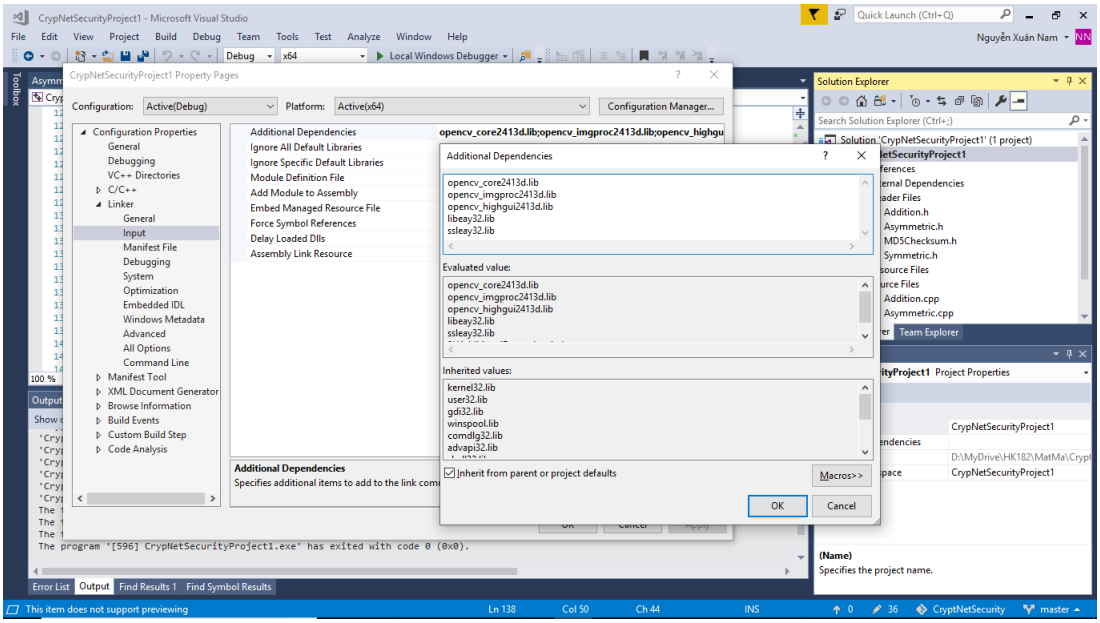
Hình 11: So sánh input và output của kỹ thuật giấu tin



Hình 12: Lấy thông điệp từ ảnh mang



Hình 13: Lấy thông điệp từ ảnh mang



Hình 14: Lấy thông điệp từ ảnh mang

- Tính thực tế của ứng dụng chưa cao, chỉ có thể sử dụng cho việc mã hóa các tệp trong máy. Đồng thời việc bảo khóa bí mật chưa được quan tâm đến.
- Ứng dụng không có giao diện nên tạo khó khăn cho việc sử dụng.
- Không có thanh trạng thái báo cho người dùng biết quá trình mã hóa đang diễn ra.

Hướng phát triển ứng dụng:

- Khắc phục các nhược điểm được nêu ra.
- Kết hợp các giải thuật mã hóa đã hiện thực để tao ứng dụng nhắn tin an toàn.

6 Phân công công việc

Khối lượng công việc được chia đều cho các thành viên trong nhóm như sau:

- Nguyễn Trần Lê Minh: Hiện thực khối DES.
- Lê Duy Thanh: Hiện thực khối RSA.
- Nguyễn Xuân Nam: Hiện thực khối MD5 và Steganography.

Tài liệu

[1] Thư viện openssl, https://www.openssl.org/docs/man1.0.2/

[2] Sách giáo khoa môn học: Stallings, William - Cryptography and network security principles and practice (2017, Pearson).