**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

****

|  |
| --- |
| **BÀI TẬP LỚN**  **AN TOÀN VÀ BẢO MẬT THÔNG TIN** |
|  |
| **ĐỀ TÀI:**  **CHIA SẺ BÍ MẬT THÔNG TIN DẠNG ẢNH VỚI HỆ MÃ HÓA HILL VÀ LƯỢC ĐỒ NGƯỠNG SHAMIR**   |  |  | | --- | --- | | **Giảng viên hướng dẫn :** | **Th.S Trần Phương Nhung** | | **Nhóm :** | **9** | | **Sinh viên thực hiện :** | **Nguyễn Ngọc Khánh 2019605636**  **Đỗ Tuấn Long 2019607069**  **Nguyễn Minh Long 2019607842**  **Ngọ Trọng ong 2019606620** | |

**Hà Nội – Năm 2022**

**LỜI CẢM ƠN**

Em xin chân thành cảm ơn các thầy, các cô khoa Công nghệ thông tin - trường Đại học Công Nghiệp Hà Nội đã tận tình dạy dỗ, truyền đạt cho chúng em nhiều kiến thức bổ ích và quý báu trong suốt những năm học đã qua.

Em xin tỏ lòng biết ơn sâu sắc đến cô Trần Phương Nhung, người đã trực tiếp hướng dẫn, giúp đỡ và truyền đạt cho em những kiến thức và kinh nghiệm để đề tài này có thể thực hiện được và hoàn thành.

Vì thời gian có hạn, trình độ hiểu biết của bản thân còn nhiều hạn chế. Cho nên trong bài tập lớn không tránh khỏi những thiếu sót, em rất mong nhận được sự đóng góp ý kiến của tất cả các thầy cô giáo cũng như các bạn bè để bài tập lớn của em được hoàn thiện hơn trong tương lai.

Em xin chân thành cảm ơn.

**MỤC LỤC**

[CHƯƠNG 1. TỔNG QUAN VỀ CHƯƠNG TRÌNH CHIA SẺ THÔNG TIN DẠNG ẢNH 4](#_Toc107400304)

[1.1. Lý do chọn đề tài 4](#_Toc107400305)

[1.2. Cở sở lý thuyết 5](#_Toc107400306)

[1.2.1. Tập các số nguyên (Set of Integers) 5](#_Toc107400307)

[1.2.2. Thuật toán Euclid 5](#_Toc107400308)

[1.2.3. Nghịch đảo nhân (Multiplicative Inverse) 6](#_Toc107400309)

[1.3. Mật mã Hill 6](#_Toc107400310)

[1.3.1. Quá trình mã hóa Hill (Encoded Hill Cipher) 6](#_Toc107400311)

[1.3.2. Quá trình giải mã Hill (Decoded Hill Cipher) 7](#_Toc107400312)

[1.4. Lược đồ ngưỡng Shamir 7](#_Toc107400313)

[1.5. Giải pháp đề xuất 8](#_Toc107400314)

[1.5.1. Tiến trình mã hóa chia sẻ ảnh thông tin mật 9](#_Toc107400315)

[1.5.2. Giai đoạn mã hóa ảnh thông tin mật 9](#_Toc107400316)

[1.5.3. Giai đoạn chia sẻ 10](#_Toc107400317)

[1.5.4. Tiến trình khôi phục ảnh thông tin 12](#_Toc107400318)

[1.5.5. Giai đoạn khôi phục ảnh 12](#_Toc107400319)

[1.5.6. Giai đoạn giải mã ảnh 13](#_Toc107400320)

[CHƯƠNG 2. KẾT QUẢ NGHIÊN CỨU 15](#_Toc107400321)

[2.1. Giới thiệu 15](#_Toc107400322)

[2.2. Nội dung: 15](#_Toc107400323)

[2.3. Yêu cầu 15](#_Toc107400324)

[2.4. Các bước triển khai đề tài bao gồm: 15](#_Toc107400325)

[2.5. Nội dung thuật toán 15](#_Toc107400326)

[2.5.1. Tập các số nguyên (Set of Integers) 15](#_Toc107400327)

[2.5.2. Thuật toán Euclid 16](#_Toc107400328)

[2.5.3. Nghịch đảo nhân (Multiplicative Inverse) 17](#_Toc107400329)

[2.5.4. Mật mã Hill 17](#_Toc107400330)

[2.5.5. Quá trình mã hóa Hill (Encoded Hill Cipher) 17](#_Toc107400331)

[2.5.6. Quá trình giải mã Hill (Decoded Hill Cipher) 17](#_Toc107400332)

[2.5.7. Lược đồ ngưỡng Shamir 17](#_Toc107400333)

[2.6. Thiết kế, cài đặt chương trình demo thuật toán 19](#_Toc107400334)

[2.6.1. Môi trường cài đặt 19](#_Toc107400335)

[2.6.2. Giao diện chương trình demo 19](#_Toc107400336)

[2.7. Thực hiện bài toán 22](#_Toc107400337)

[2.7.1. Các chức năng chính 22](#_Toc107400338)

[2.7.2. Phân công công việc 23](#_Toc107400339)

[CHƯƠNG 3. KẾT LUẬN 24](#_Toc107400340)

[3.1. Nội dung đã thực hiện 24](#_Toc107400341)

[3.1.1. Kiến thức cần nắm được 24](#_Toc107400342)

[3.2. Bài học kinh nghiệm 25](#_Toc107400343)

[3.3. Hướng phát triển. 26](#_Toc107400344)

[Tài liệu tham khảo 27](#_Toc107400345)

# TỔNG QUAN VỀ CHƯƠNG TRÌNH CHIA SẺ THÔNG TIN DẠNG ẢNH

## Lý do chọn đề tài

Tốc độ phát triển mạnh mẽ của hạ tầng mạng Internet cùng giá thành các thiết bị máy tính và mạng máy tính ngày càng giảm đã làm một cuộc cách mạng trong giao thức xử lý thông tin và truyền tải thông tin. Việc lưu trữ, xử lý thông tin bằng các thiết bị tính toán (PC, laptop, smartphone, tablet, camera…) và chia sẻ chúng qua môi trường mạng ngày càng trở nên phổ dụng. Sự phát triển của mạng Internet ở Việt Nam được xem là nền tảng, động lực cho sự phát triển thương mại điện tử, giáo dục, y tế… Tuy nhiên, như chúng ta biết, Internet là một hệ thống mạng mở, toàn cầu, trong đó môi trường truyền thông thông tin được xem là không an toàn. Trong khi yêu cầu của người dùng mỗi khi tiếp nhận cũng như truyền tin là phải an toàn và chính xác. Vì vậy, việc lưu trữ và truyền tải dữ liệu phải được giữ bí mật đối với người không được phép là một yêu cầu bắt buộc của người thiết kế hệ thống thông tin.

Việc bảo mật thông tin được đặt ra từ rất sớm, chẳng hạn trong (Liu C.L 1968) đưa ra vấn đề: Có 11 nhà khoa học cùng làm chung một dự án bí mật, họ mong muốn rằng toàn bộ hồ sơ của dự án được lưu giữ trong một cái hộp. Hộp này chỉ có thể mở ra nếu như có sự hợp tác đủ ít nhất 6/11 nhà khoa học. Vậy hỏi, số ổ khóa và chìa khóa ít nhất mà mỗi nhà khoa học đề cập trên cần phải có là bao nhiêu? Không khó để có lời giải cho bài toán này là mỗi nhà khoa học phải mang theo 462 ổ khóa và 252 chìa khóa. Tuy nhiên, ngày nay, đây là một lời giải không thực tế, và nó sẽ tăng theo lũy thừa khi số lượng nhà khoa học tham gia cùng dự án tăng lên.

Năm 1979, (A. Shamir) đề xuất một phương pháp để giải quyết bài toán nêu trên gọi là chia sẻ bí mật (secret sharing). Secret sharing là phương pháp chia sẻ thông tin bí mật thành hai hay nhiều phần (share/shadow), mà mỗi phần sẽ do một người nắm giữ. Và thông tin bí mật ấy chỉ có thể khôi phục khi có sự hợp tác của tối thiểu một số lượng người đã định trước. Một cách logic, mỗi phần chia là một phần của thông tin, nhưng phần thông tin đó là vô nghĩa nếu như chúng đứng riêng lẻ.

Chia sẻ bí mật là một giải pháp quan trọng trong lĩnh vực bảo mật thông tin, là kỹ thuật cho phép tạo ra nhiều bóng tin (shadow data) hay còn gọi là thông tin chia sẻ (shared data) của thông tin gốc mà mỗi bóng tin là hiển thị thông tin không có giá trị. Tuy nhiên, khi tập hợp đủ số lượng thông tin chia sẻ cần thiết thì thông tin gốc ban đầu sẽ được phục hồi. Bài báo này sẽ tập trung nghiên cứu giải pháp chia sẻ bí mật thông tin mà mỗi thông tin gốc ở đây được xem là một tập tin hình ảnh (image). Ảnh thông tin bí mật (secret image) sẽ được mã hóa trước khi phân mảnh. Điều này giúp tăng cường tính bảo mật cho ảnh thông tin cần giữ bí mật.

Xuất phát từ nhu cầu thực tế đó, nhóm sinh viên chúng em đã lựa chọn đề tài :

***“Chia sẻ bí mật thông tin dạng ảnh với hệ mã hóa Hill và lược đồ ngưỡng Shamir ”***.

## Cở sở lý thuyết

### Tập các số nguyên (Set of Integers)

Tập hợp các số nguyên, được ký hiệu là **Z**, là tập các số nguyên (không chứa số phân số) có miền xác định từ -∞ đến +∞ **Z** = {…, -2, -1, 0, 2, …}

Tập các số dư nguyên, được ký hiệu là **Zn** , là kết quả của phép toán modulo với n.

Khi đó **Zn** = {0, 1, …, n-1}

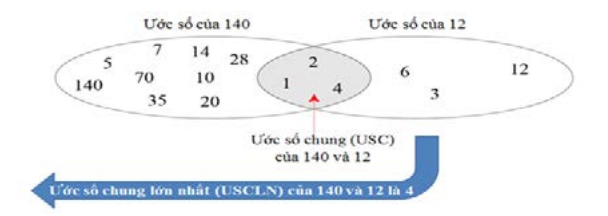
Ví dụ:

Z2 = {0, 1}; Z6 = {0, 1, …, 5} ; Z11 = {0, 1, …, 10}

### Thuật toán Euclid

Hai số nguyên dương a và b có thể có nhiều ước số, nhưng chỉ có một số chung là lớn nhất.

Ví dụ:



Như vậy, tìm USCLN của 2 số a và b là liệt kê tất cả ước số của mỗi số a và b, sau đó tìm tập giao USC của 2 số và rút ra USCLN cùa 2 số a và b. Rõ ràng với cách này là không thực tế khi 2 số a, b lớn. Thật may mắn, cách đây hơn 2000 năm, nhà toán học tên là Euclid đã phát triển một thuật toán có thể giúp ta tìm USCLN của 2 số nguyên dương. Thuật toán Euclid để tìm USCLN(a,b) có thể được mô tả như sau:

USCLN(a,b)

{ r1=a; r2=b;

While (r2>0)

{ q=a/b;

r=r1-q\*r2;

r1=r2;

r2=r;

}

return r1;

}

### Nghịch đảo nhân (Multiplicative Inverse)

Trong tập Zn , hai số a và b được gọi là nghịch đảo nhân của nhau nếu:

a ×b ≡ 1 (mod n)

Ví dụ:

Nếu giá trị modulus n = 10, thì nghịch đảo nhân của 3 trong Z10 là 7, vì:

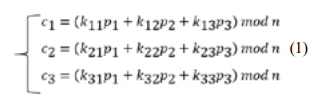
(3 × 7) mod 10 = 1

## Mật mã Hill

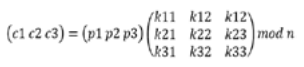
Mật mã Hill do nhà toán học người Mỹ tên là Lester S. Hill đề xuất năm 1929. Ý tưởng của giải thuật mã hóa Hill là lần lượt lấy m ký tự liên tiếp trong bản rõ (Plaintext) để thay thế bởi m ký tự mã hóa. Tiến trình thay thể được xác định bởi m phương trình tuyến tính.

### Quá trình mã hóa Hill (Encoded Hill Cipher)

Giả sử m=3, ta có hệ 3 phương trình tuyến tính như sau:



hay



hay C = E(P,K) = PK (mod n); trong đó E là hàm mã hóa, P là ký tự rõ, K là ma trận khóa, C là mật mã.

### Quá trình giải mã Hill (Decoded Hill Cipher)

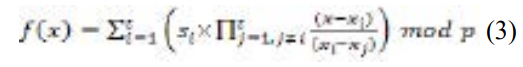
P’ = D(K, C) = CK-1 mod n = PKK-1 = PI = P, Trong đó: K-1là ma trận nghịch đảo của ma trận K, I là ma trận đơn vị.

## Lược đồ ngưỡng Shamir

Khái niệm chia sẻ bí mật đầu tiên được giới thiệu vào năm 1979 bởi Shamir [1] và được gọi là lược đồ ngưỡng (t, n); trong đó n là số thành viên tham gia vào hệ thống, t là số thành viên tối thiểu (ngưỡng) cùng tham gia để phục hồi thông tin mật. Lược đồ Shamir dựa trên hàm đa thức bậc t−1 được định nghĩa như sau:



Để mã hóa chia sẻ bí mật s thành n mảnh/phần chia (shadow/share), ký hiệu là: s1 , s2 , …, sn , với p là số nguyên tố và các hệ số a1 , a2 , …, at-1 được lựa chọn ngẫu nhiên sao cho ai ∈[0, p−1]. Chọn n số nguyên x1 , x2 , …, xn khác nhau từng đôi một, tượng trưng cho n thành phần tham gia vào hệ thống. Các phần chia s1 , s2 , …, sn được tính si = f(xi), với i = 1..n . Khi đó mỗi thành phần tham gia vào hệ thống được chia sẻ cặp (xi, si). s chỉ có thể được khôi phục khi tập hợp đủ ít nhất t phần chia (share) (có ít nhất t cặp (xi , si )) và áp dụng vào đa thức nội suy Lagrange:



Ví dụ:

Lược đồ ngưỡng (t, n)=(2,3) với giá trị bí mật s=3, chọn p=11, từ (2), ta có:

f(x) = 3+2x mod 11.

Nếu ta chọn x1 =1, x2 =2, x3 =3, f(1)=1, f(2)=2, f(3)=3, theo đó 3 phần chia tương ứng là: (1,5), (2,7), (3,9).

Thật ra chúng là những điểm trên đồ thị của phương trình f(x) = 3+2x mod 11.

Hình 1.1: Đồ thị f(x) = 3+2x

Để khôi phục lại giá trị bí mật s, ta phải tập đủ ít nhất 2 phần trong 3 phần chia. Chẳng hạn, ta có 2 phần chia: (1,5) và (2,7).

Từ công thức (3):

Ta được:

suy ra, s=3.

## Giải pháp đề xuất

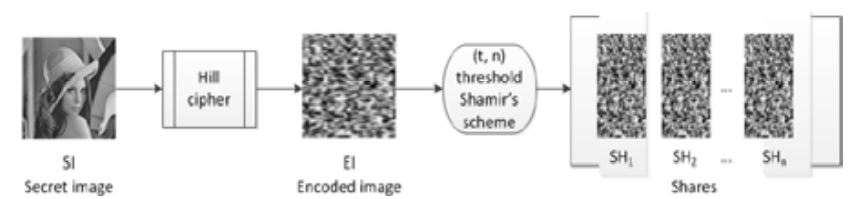
Một hệ mật mã luôn bao gồm hai tiến trình: mã hóa và giải mã. Giải pháp chia sẻ bảo mật thông tin với phần thông tin được xem là ảnh số đề xuất trong nghiên cứu này bao gồm tương ứng hai tiến trình như vậy:

(1) tiến trình mã hóa chia sẻ ảnh thông tin mật (ảnh bí mật – Secret image)

(2) tiến trình khôi phục ảnh thông tin (Recovered image).

### Tiến trình mã hóa chia sẻ ảnh thông tin mật

Trong tiến trình này, để tăng tính bảo mật cho ảnh bí mật (SI – secret image) cần chia sẻ, ta mã hóa SI dựa trên phương pháp mã hóa Hill, với ma trận vuông khóa K cấp (2×2) được lựa chọn ngẫu nhiên sao cho có tồn tại ma trận nghịch đảo K-1 theo modulo cường độ điểm ảnh. Như sơ đồ được mô tả trong Hình 2, ảnh thu được sau khi được mã hóa là EI sẽ áp dụng lược đồ ngưỡng (t, n) của Shamir. Theo cách này, ảnh EI sẽ phân chia thành n mảnh khác nhau ký hiệu là SH1 , SH2 , …, SHn.



Hình 1.2: Tiến trình mã hóa chia sẻ ảnh thông tin mật

Chi tiết của tiến trình này có thể được mô tả như sau:

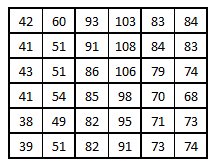
### Giai đoạn mã hóa ảnh thông tin mật

Bước 1: Chia ảnh đa cấp xám gồm có 256 cấp xám (ảnh xám) SI với kích thước (m×n) thành những khối (blocks) gồm 2 điểm ảnh (pixel) theo nguyên tắc trái-phải-trên-dưới. Nếu SI là ảnh màu true-color, ta tách lớp ảnh SI thành 3 lớp (layer) SIR, SIG, SIB và thực hiện từng lớp ảnh như một đa cấp ảnh xám.

Bước 2: Lấy lần lượt từng khối 2 điểm ảnh (p1 , p2 ) áp dụng vào công thức (1) với modulo 256 (giá trị mỗi điểm ảnh xám là 8 bits),1 tạo thành công thức mã hóa ảnh như sau :

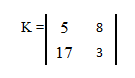


Với P1’ và P2’ là những điểm ảnh bị mã hóa (encoded pixel); k11, k12, k21, k22, là những khóa dùng để mã hóa theo phương pháp Hill. Ví dụ: Giả sử ta có ma trận điểm ảnh SI như sau:



Hình 1.3: Ma trận ảnh SI

Khóa được chọn là :

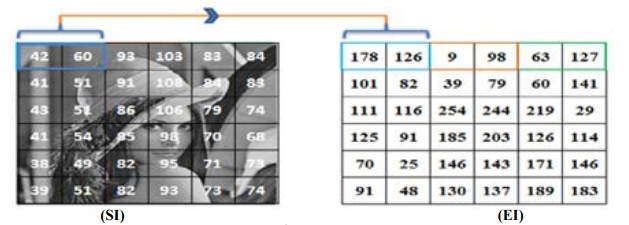


Khối ảnh đầu tiên để mã hóa là (p1 , p2 ) = (42,60)

*P1’ = (5x42 + 8x60) mod 256*

*P2’ = (17x42 + 3x60) mod 256*

Áp dụng vào (4): Ta có: và . Tiếp tục lấy những khối ảnh kế tiếp để mã hóa, cuối cùng ta nhận được ảnh mã hóa EI như hình 3.

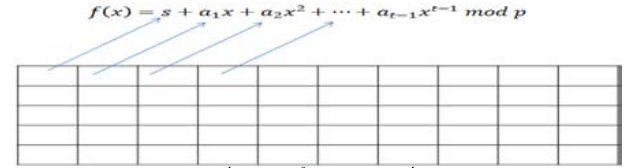


Hình 1.4: Ảnh mã hóa EO từ SI

### Giai đoạn chia sẻ

Sau khi mã hóa ảnh SI, ta thu được ảnh mã hóa EI. Tiếp theo của tiến trình này là chia sẻ ảnh mã hóa EI thành những phần chia (shares) dựa trên lược đồ ngưỡng

(t, n) của Shamir. Trong giai đoạn này, chúng tôi xem giá trị các pixels như các hệ số {s, a1 , a2 , …, at-1} của đa thức (2) như Hình 4 và số nguyên tố p = 251 là số nguyên tố lớn nhất nhỏ hơn 255 (giá trị lớn nhất của điểm ảnh xám).



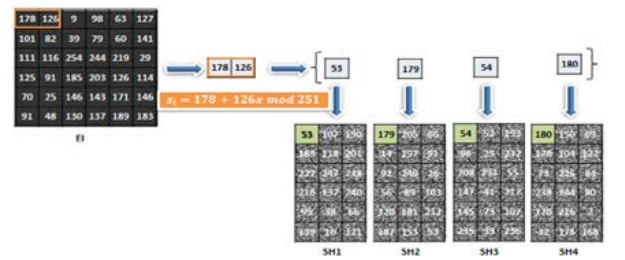
Hình 1.5: Bố trí giá trị điểm ảnh như hệ số của đa thức (2)

Theo đó, ảnh EI kích thước (m×n) sẽ bị chia thành (m×n)/t khối, mỗi khối có t điểm ảnh tương ứng với các hệ số {s, a1 , a2 , …, at-1} của công thức (2). Và kết quả đa thức này ta thu được n giá trị s1 =f(x1 ), s2 =f(x2 ), …, sn =f(xn ), trong đó x1 , x2 , …, xn được định ngẫu nhiên (xem như mã số của từng thành phần tham gia vào hệ thống) và khác nhau từng đôi một. Các giá trị s1 , s2 , ..sn phân bố một cách tương ứng vào n ảnh chia sẻ (shadow image) ký hiệu là: SH1 , SH2 , …, SHn . Và kích thước mỗi ảnh chia sẻ bằng 1/t kích thước ảnh EI.

Để mô tả giai đoạn chia sẻ ảnh, ta tiếp ví dụ ở phần trên (Hình 3), với ngưỡng (2, 4), có nghĩa là t=2, n=4. Như vậy hàm đa thức được thiết lập như sau đối với cặp điểm ảnh đầu tiên (178, 126):

f(x) = (178 + 126x) mod 251 (5)

Chọn ngẫu nhiên 4 giá tri x khác nhau, tượng trưng mã số của bốn (n=4) thành viên tham gia vào hệ thống và thay vào (5), chẳng hạn x={1, 2, 3, 4}. Kết quả là ta nhận được 4 điểm chia sẻ (1, 53); (2, 179); (3, 54); (4, 180). Bốn điểm chia sẻ này trở thành điểm ảnh đầu tiên của mỗi 4 ảnh chia sẻ SH1 , SH2 , SH3 , SH4 . Điểm ảnh thứ 2 của mỗi ảnh chia sẻ SHi (i=1..4) được tính tương tự như trên. Hình 5 mô tả chi tiết tiến trình chia sẻ ảnh dựa trên lược đồ ngưỡng Shamir.

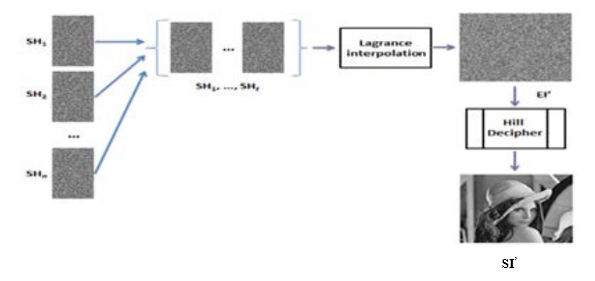


Hình 1.6: Giai đoạn chia sẻ ảnh

Mỗi ảnh chia SHi có kích thước bằng 1/2 kích thước của ảnh EI, trong đó không ảnh SHi nào hiển thị thông tin của ảnh gốc.

### Tiến trình khôi phục ảnh thông tin

Tiến trình khôi phục dữ liệu được thực hiện ở phía người nhận, có nghĩa là dữ liệu sau khi được mã hóa, chia sẻ sẽ gửi cho n người nhận qua phương tiện truyền thông. Tiến trình này bao gồm hai giai đoạn: khôi phục và giải mã dữ liệu, tiến trình có thể được tóm tắt như hình 1.7. Giai đoạn thứ nhất kết nối t ảnh chia sẻ dựa trên đa thức nội suy Lagrance, tiếp theo giai đoạn giải mã theo phương pháp Hill. Chi tiết của tiến trình được trình bày như sau:



Hình 1.7: Tiến trình khôi phục dữ liệu

### Giai đoạn khôi phục ảnh

Đối với lược đồ (t,n) - ngưỡng của Shamir, thông tin được phục hồi là sự hợp tác của t thành phần tham gia vào hệ thống; hay nói cách khác, phải tập hợp đủ ít nhất bất kỳ t ảnh chia sẻ của n mẫu chia. Khi đó ảnh mã hóa EI’ được khôi phục 100% chính xác bởi đa thức nội suy Lagrance, công thức (3).

Từng bước của giai đoạn mô tả như sau:

Bước 1: Lấy điểm ảnh đầu tiên chưa được sử dụng của mỗi t ảnh chia sẻ.

Bước 2: Áp dụng t điểm ảnh này vào đa thức (3) để tìm các hệ số a0, a1 ,…,at-1. Khi đó các hệ số a0 , …, at-1 sẽ tương ứng với các điểm ảnh của ảnh mã hóa EI’.

Bước 3: Lặp lại bước 1 và bước 2 cho đến khi tất cả các điểm ảnh của t ảnh chia sẻ được thực hiện

### Giai đoạn giải mã ảnh

Dùng phương pháp giải mã Hill đối với ảnh mã hóa EI’ để giải mã ảnh bí mật theo các bước được mô tả như sau:

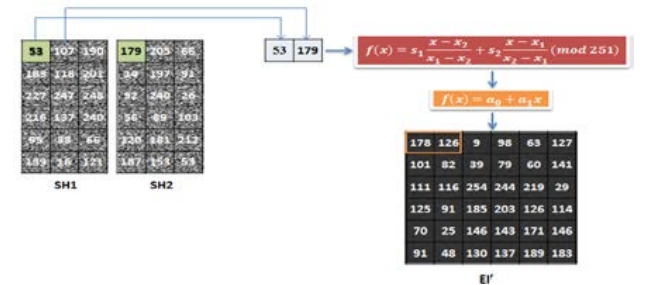
Bước 1: Chia ảnh mã hóa EI’ thành từng khối 2 điểm ảnh theo quy tắc trái-phải-trên-dưới.

Bước 2: Lấy lần lượt từng cặp điểm ảnh (P1EI’, P2EI’)

Bước 3: Áp dụng hàm giải mã Hill đối với 2 vừa lấy ở bước 2. Trong đó K-1 là ma trận nghịch đảo của K.

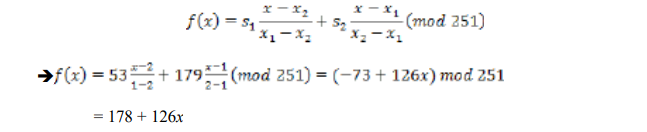


Bước 4: Lặp lại từ bước 2 đến 3 cho đến khi tất cả khối điểm ảnh được giải mã. Để diễn giải cho tiến trình khôi phục và giải mã ảnh thông tin gốc, ta sử dụng ví dụ bên trên, Hình 5. Giả sử 2 trong 4 ảnh chia sẻ được chọn là SH1 và SH2 để hợp tác cùng khôi phục ảnh theo lược đồ ngưỡng (2, 4) của Shamir.



Hình 1.8: Kết hợp bất kì của 2 trong 4 ảnh SHi

Lấy lần lượt từng điểm của mỗi ảnh chia sẻ (SH1, SH2), kết hợp nhau để khôi phục lại ảnh mã hóa EI’. Như vậy cặp điểm ảnh đầu tiên của ảnh mã hóa EI’ được phục hồi qua cặp điểm ảnh (1,53) và (2,179) và đa thức nội suy Lagrance (3):



Suy ra: cặp điểm ảnh thứ nhất của EI’ là (P1EI’, P2EI’ ) = (178, 126)

Tiến trình khôi phục ảnh EI’ được tiếp tục cho đến khi tất cả điểm ảnh của 2 ảnh chia sẻ SH1 va SH2 được thực hiện.

Sau khi nhận được ảnh mã hóa EI’, ta phải giải mã để có được ảnh thông tin mật ban đầu. Hình 8 trình bày sơ đồ giải mã ảnh thông tin dựa trên hệ mật mã Hill.



Hình 1.9: Giai đoạn giả mã ảnh thông tin gốc

# KẾT QUẢ NGHIÊN CỨU

## Giới thiệu

* Đề tài nghiên cứu: Chia sẻ bí mật thông tin dạng ảnh với hệ mã hóa Hill và lược đồ ngưỡng Shamir
* Nhiệm vụ, công việc chính khi thực hiện bài tập lớn:

## Nội dung:

1. Tìm hiểu mật mã Hill.
2. Nghiên cứu bài toán chia sẻ bí mật ngưỡng Shamir.
3. Ứng dụng lược đồ chía sẻ bí mật vào chia sẻ thông tin dưới dạng ảnh.
4. Demo chương trình.

## Yêu cầu

1. Đọc tài liệu và hiểu được vấn đề đặt ra, nắm được các phương pháp mã hóa Hill một cách thành thạo.
2. Hiểu được lược đồ chia sẻ bí mật Lagrange(sơ đồ ngưỡng Shamir)
3. Đọc hiểu được 1 số tài liệu chuyên môn bằng tiếng anh
4. Nắm vững một ngôn ngữ lập trình cơ bản và giải 1 bài toàn có tính ứng dụng vào thực tiễn.

## Các bước triển khai đề tài bao gồm:

1. Nghiên cứu nội dung các thuật toán (Thuật toán Euclid, Module, Multiplicative Inverse, Hill, đa thức Lagrange)
2. Thiết kế giao diện, cài đặt chương trình demo

* Hình thức sản phẩm: ứng dụng phầm mềm

## Nội dung thuật toán

### Tập các số nguyên (Set of Integers)

Tập hợp các số nguyên, được ký hiệu là **Z**, là tập các số nguyên (không chứa số phân số) có miền xác định từ -∞ đến +∞ **Z** = {…, -2, -1, 0, 2, …}

Tập các số dư nguyên, được ký hiệu là **Zn** , là kết quả của phép toán modulo với n.

Khi đó **Zn** = {0, 1, …, n-1}

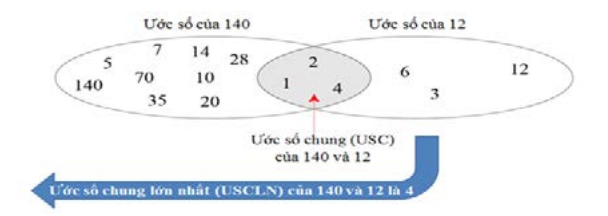
Ví dụ:

Z2 = {0, 1}; Z6 = {0, 1, …, 5} ; Z11 = {0, 1, …, 10}

### Thuật toán Euclid

Hai số nguyên dương a và b có thể có nhiều ước số, nhưng chỉ có một số chung là lớn nhất.

Ví dụ:



Như vậy, tìm USCLN của 2 số a và b là liệt kê tất cả ước số của mỗi số a và b, sau đó tìm tập giao USC của 2 số và rút ra USCLN cùa 2 số a và b. Rõ ràng với cách này là không thực tế khi 2 số a, b lớn. Thật may mắn, cách đây hơn 2000 năm, nhà toán học tên là Euclid đã phát triển một thuật toán có thể giúp ta tìm USCLN của 2 số nguyên dương. Thuật toán Euclid để tìm USCLN(a,b) có thể được mô tả như sau:

USCLN(a,b)

{ r1=a; r2=b;

While (r2>0)

{ q=a/b;

r=r1-q\*r2;

r1=r2;

r2=r;

}

return r1;

}

### Nghịch đảo nhân (Multiplicative Inverse)

Trong tập Zn , hai số a và b được gọi là nghịch đảo nhân của nhau nếu:

a ×b ≡ 1 (mod n)

Ví dụ:

Nếu giá trị modulus n = 10, thì nghịch đảo nhân của 3 trong Z10 là 7, vì:

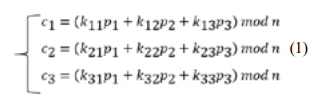
(3 × 7) mod 10 = 1

### Mật mã Hill

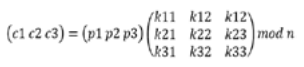
Mật mã Hill do nhà toán học người Mỹ tên là Lester S. Hill đề xuất năm 1929. Ý tưởng của giải thuật mã hóa Hill là lần lượt lấy m ký tự liên tiếp trong bản rõ (Plaintext) để thay thế bởi m ký tự mã hóa. Tiến trình thay thể được xác định bởi m phương trình tuyến tính.

### Quá trình mã hóa Hill (Encoded Hill Cipher)

Giả sử m=3, ta có hệ 3 phương trình tuyến tính như sau:



hay



hay C = E(P,K) = PK (mod n); trong đó E là hàm mã hóa, P là ký tự rõ, K là ma trận khóa, C là mật mã.

### Quá trình giải mã Hill (Decoded Hill Cipher)

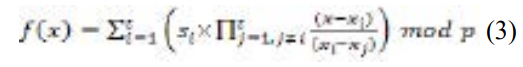
P’ = D(K, C) = CK-1 mod n = PKK-1 = PI = P, Trong đó: K-1là ma trận nghịch đảo của ma trận K, I là ma trận đơn vị.

### Lược đồ ngưỡng Shamir

Khái niệm chia sẻ bí mật đầu tiên được giới thiệu vào năm 1979 bởi Shamir [1] và được gọi là lược đồ ngưỡng (t, n); trong đó n là số thành viên tham gia vào hệ thống, t là số thành viên tối thiểu (ngưỡng) cùng tham gia để phục hồi thông tin mật. Lược đồ Shamir dựa trên hàm đa thức bậc t−1 được định nghĩa như sau:



Để mã hóa chia sẻ bí mật s thành n mảnh/phần chia (shadow/share), ký hiệu là: s1 , s2 , …, sn , với p là số nguyên tố và các hệ số a1 , a2 , …, at-1 được lựa chọn ngẫu nhiên sao cho ai ∈[0, p−1]. Chọn n số nguyên x1 , x2 , …, xn khác nhau từng đôi một, tượng trưng cho n thành phần tham gia vào hệ thống. Các phần chia s1 , s2 , …, sn được tính si = f(xi), với i = 1..n . Khi đó mỗi thành phần tham gia vào hệ thống được chia sẻ cặp (xi, si). s chỉ có thể được khôi phục khi tập hợp đủ ít nhất t phần chia (share) (có ít nhất t cặp (xi , si )) và áp dụng vào đa thức nội suy Lagrange:



Ví dụ:

Lược đồ ngưỡng (t, n)=(2,3) với giá trị bí mật s=3, chọn p=11, từ (2), ta có:

f(x) = 3+2x mod 11.

Nếu ta chọn x1 =1, x2 =2, x3 =3, f(1)=1, f(2)=2, f(3)=3, theo đó 3 phần chia tương ứng là: (1,5), (2,7), (3,9).

Thật ra chúng là những điểm trên đồ thị của phương trình f(x) = 3+2x mod 11.

Hình 2.1: Đồ thị f(x) = 3+2x

Để khôi phục lại giá trị bí mật s, ta phải tập đủ ít nhất 2 phần trong 3 phần chia. Chẳng hạn, ta có 2 phần chia: (1,5) và (2,7).

Từ công thức (3):



Ta được:



suy ra, s=3.

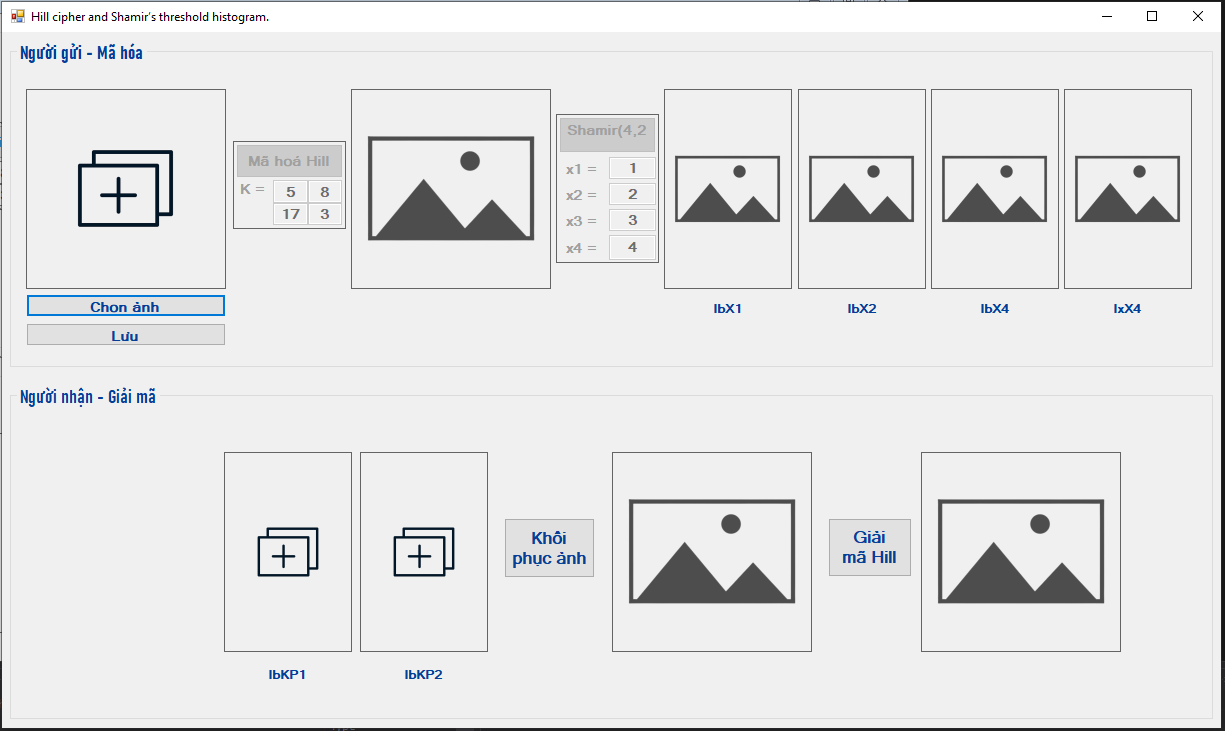
## Thiết kế, cài đặt chương trình demo thuật toán

### Môi trường cài đặt

Ngôn ngữ lập trình: C#

IDE hỗ trợ soạn thảo: Visul studio 2019 community

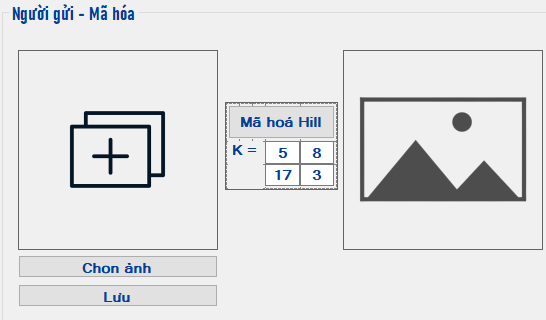
### Giao diện chương trình demo



Hình 2.2: Giao diện chương trình demo

* Mã hóa Hill

Để có thể tiến hành mã hóa người dùng cần cung cấp thông tin cần mã hóa và khóa K(Khóa mã hóa Hill).



Hình 2.3: Chức năng mã hóa Hill

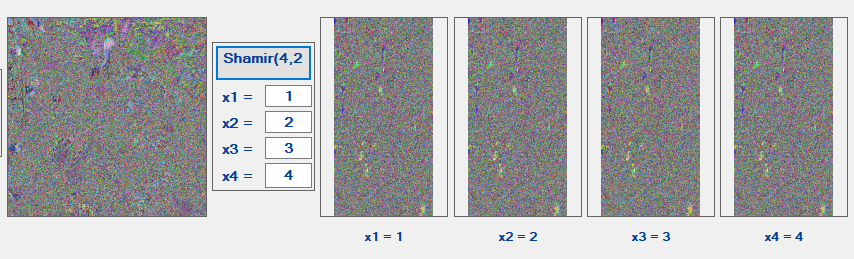
Sau đó click nút mã hóa để tiến hành mã hóa



Hình 2.4: Kết quả sau khi nhập thông tin và mã hóa Hill

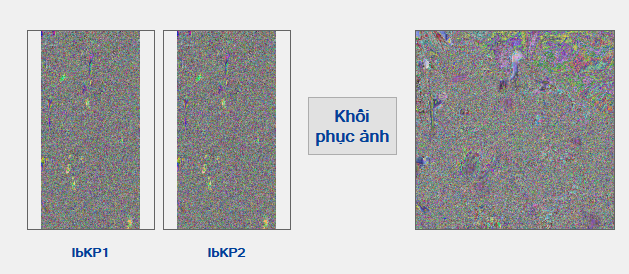
* Chia sẻ Shamir

Sau khi mã hóa Hill thành công, người dùng click vào nút Shamir để tiến hành chia sẻ lược đồ ngưỡng Shamir



Hình 2.5: Lược đồ ngưỡng Shamir

* Khôi phục ảnh



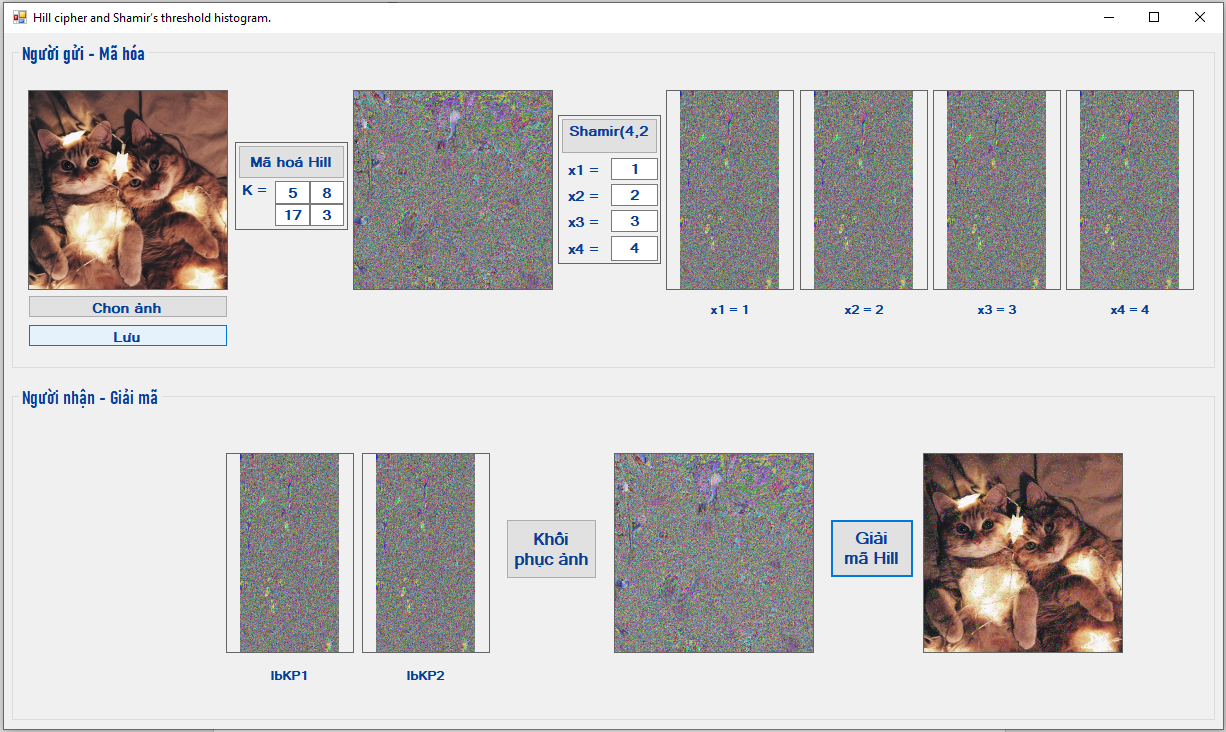
Hình 2.6: Khôi phục ảnh

* Giải mã Hill



Hình 2.7: Giải mã Hill

* Tổng quát quá trình người gửi – người nhận



Hình 2.8: Quá trình mã hóa và giải mã

## Thực hiện bài toán

### Các chức năng chính

Các chức năng trong chương trình demo:

1. Mã hóa ảnh
2. Chia sẻ ảnh
3. Khôi phục ảnh
4. Giải mã ảnh
5. Mã hóa kí tự trong bảng chữ cái tiếng anh Z26
6. Giải mã kí tự trong bảng chữ cái tiếng anh Z26

### Phân công công việc

|  |  |
| --- | --- |
| **Tên sinh viên** | **Tên công việc** |
| Nguyễn Ngọc Khánh | - Tìm hiểu về ảnh số  + Khải niệm ảnh số, ma trận ảnh, ảnh RGB, ảnh xám  + Cách đọc ảnh, ghi ảnh  + Cách chuyển đổi ảnh sang mảng, mảng sang ảnh  - Tìm hiểu về hệ mã hóa Hill  - Tìm hiểu về lược đồ ngưỡng Shamir  + Thuật toán Euclid  + Phép toán Modulo  + Qui trình chia ảnh thành cách ảnh con  + Qui trình khôi phục ảnh |
| Đỗ Tuấn Long | - Tìm hiểu về hệ mã hóa Hill  + Khái niệm  + Nguyên lý mã hóa, giải mã  + Thuật toán Euclid  + Phép toán Modulo  + Qui trình mã hóa Hill  + Qui trình giải mã |
| Nguyễn Minh Long | - Tìm hiểu về hệ mã hóa Hill  + Khái niệm  + Nguyên lý mã hóa, giải mã  + Thuật toán Euclid  + Phép toán Modulo  + Mã hóa Hill |
| Ngọ Trọng Long | - Tìm hiểu về lược đồ ngưỡng Shamir  + Thuật toán Euclid  + Phép toán Modulo  + Qui trình chia thông tin thành cách mảnh  + Qui trình khôi phục thông tin |

# KẾT LUẬN

## Nội dung đã thực hiện

### Kiến thức cần nắm được

Sau quá trình nghiêm cứu đề tài cần phải nắm được những kỹ thuật, kiến thức cơ bản:

* **Các khái niệm cơ bản cần nắm được:**
* Tổng quan về ảnh số
* Một số phép toán, thuật toán cơ sở phụ trợ(Modulo, Euclid, nghịch đảo nhân,…)
* Qui trình mã hóa, giải mã Hill
* Tổng quan về lược đồ ngưỡng Shamir
* **Về các kỹ thuật cần nắm được những kỹ thuật sau:**
* Đọc ảnh RGB lấy ma trận
* Sử dụng thuật toán Hill tiến hành mã hóa, giải mã
* Sử dụng lược đồ ngưỡng Shamir để chia ảnh và khôi phục ảnh
* **Về các ngôn ngữ sử dụng:**
* Nắm được cách cài đặt các ngôn ngữ sử dụng để làm chương trình demo như C#, C++, Java, Java script
* Nắm được một số hàm và biến sử dụng trong chương trình khi code.
* Nắm được cách thiết kế giao diện cho chương trình.
* Nắm vững được thuật toán và triển khai thuật toán vào chương trình để tiến hành code.
* Phân tích để có thể tiến hành tạo các hàm và lớp để sử dụng.
* Kỹ năng làm việc nhóm hiệu quả thông qua việc phân công các công việc cụ thể và deadline giúp công việc đi đúng tiến độ hơn.
* Hiểu biết hơn về ứng dụng mã hóa và giải mã đặc trưng vân tay.
* Có kỹ năng chuẩn bị trước để không bị rơi vào thế bị động.
* Hiểu biết hơn về kiến thức chưa từng gặp
* Về kỹ năng chuyên môn:
* Thu thập, phân tích các dữ liệu một cách hiệu quả.
* Thiết kế, triển khai dự án một cách hiệu quả
* Đánh giá, đảm bảo đúng tiến độ.
* Vận dụng các kiến thức có sẵn của các thành viên để giải quyết dự án một cách nhanh chóng.
* Nắm được các kiến thức về an toàn bảo mật và lập trình.
* **Kỹ năng mềm:**
* Sử dụng công cụ một cách có hiệu quả, tối ưu.
* Có phương pháp làm việc khoa học, chuyên nghiệp.
* Khả năng tự học, tự nghiên cứu giải quyết vấn đề.
* Biết xây dựng và thực hiện tinh thần làm việc theo nhóm.
* Trao đổi và trau dồi kiến thức cho các thành viên trong nhóm qua các công cụ.
* Năng động trong phân công công việc nhóm, tích cực hoàn thành các công việc đã được phân công.
* Tích cực đóng góp các ý kiến để xây dựng hướng đi một cách đúng đắn nhất.
* Tích vực học hỏi các thành viên trong nhóm để cùng nhau phát triển và hoàn thiện.
* Chủ động với các công việc được phân công tránh rơi vào thế bị động – không biết nên thực hiện bước nào trước, bước nào sau.
* Giúp đỡ các thành viên còn lại khi đã hoàn thành xong công việc đã được phân công.
* Kỹ năng làm việc nhóm, kỹ năng tổ chức – phân công công việc, kỹ năng tôn trọng và giúp đỡ lẫn nhau trong công việc làm việc nhóm, có trách nhiệm với công việc và nhiệm vụ mình được giao trong bài tập lớn.

## Bài học kinh nghiệm

Qua những nội dung nhóm 09 đã thực hiện có thể thấy rằng tầm quan trọng của việc bảo mật thông tin. Thông tin là một loại hàng hóa, tài sản quan trọng đối với mỗi cá nhân, cơ quan tập thể. Việc bảo mật thông tin là vô cùng quan trọng, đặc biệt là những thông tin có liên quan đến an ninh – quốc phòng của một quốc gia.

## Hướng phát triển.

Đề tài Ứng dụng mã hóa và giải mã bằng hệ mã hóa Hill và lược đồ ngưỡng Shamir là một đề tài rất thực tiễn với xã hội hiện nay, nhằm giải quyết và đáp ứng nhu cầu một cách hiệu quả về mã hóa thông tin hiện nay. Với thời đại 4.0 hiện nay việc ứng dụng công nghệ thông tin vào đời sống ngày càng trở nên phổ biến đặc biệt trong trao đổi, giao dịch qua Internet… Nhằm mang lại tính bảo mật cao và tiết kiệm được thời gian với độ chính xác và thuận tiện cho người dùng. Vì đây là một đề tài có tính thực tiễn cao, do vậy nhóm đã thiết kế một số ứng dụng demo bằng một số ngôn ngữ khác nhau để tiến hành mã hóa và giải mã bằng Hill, lược đồ ngưỡng Shamir nhằm mang lại sự bảo mật cao cho người sử dụng.

Các hướng phát triển:

- Cho phép chia sẻ Shamir với *t, n* do người dùng cài đăt.

- Mở rộng thêm chức năng mã hóa đoạn văn bản.

Tài liệu tham khảo

[1] Võ Phước Hưng – Số 21, tháng 3/2016 – Một giải pháp mới chia sẻ bảo mật thông tin

[2] Duda, Richard O., Peter E. Hart, and David G. Stork – 2012 - Digital image processing- John Wiley & Sons

[3] Dalal - 2005 - Navneet, and Bill Triggs - [Histograms of oriented gradients for human detection](https://hal.inria.fr/file/index/docid/548512/filename/hog_cvpr2005.pdf) - Computer Vision and Pattern Recognition (CVPR).

[4] Pedro F. Felzenszwalb, and Daniel P. Huttenlocher - 2004 - [Hill](http://cvcl.mit.edu/SUNSeminar/Felzenszwalb_IJCV04.pdf) cipher - Intl. journal of computer vision 59.2.

=====\*\*\*\*\*=====

**TRÂN TRỌNG CẢM ƠN !**