**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

**TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

**----------------------------------------**

****

BÁO CÁO HỌC PHẦN

**Xây dựng sản phẩm tạo chữ ký số và giải mã chữ ký số sử dụng hệ mật ElGamal**

|  |  |
| --- | --- |
| **GVHD:** | TS. Lê Thị Anh |
| **Nhóm:** | 14 |
| **Sinh viên:** | Nguyễn Tất Vựng 2022605314 Nguyễn Phúc Xuân 2023603756  Lê Đình Vượng 2023602938  Phạm Trường Vũ 2023603133  Võ Tá Anh Vũ 2021601955 |
| **Lớp:** | 20242IT6029001 |

**Hà Nội - 2025**LỜI CẢM ƠN

Không cô đố mày làm nên, nhờ có sự định hướng tận tình và chỉ bảo sâu sắc của cô **TS. Lê Thị Anh** mà tập thể nhóm chúng em mới có thể hoàn thành bài tập lớn này. Ngay từ những bước đầu tiên, cô đã đồng hành, dìu dắt và truyền cho chúng em nguồn cảm hứng cũng như niềm tin vững chắc. Suốt quá trình thực hiện, sự kiên nhẫn chỉ dẫn, sự định hướng rõ ràng cùng sự đồng hành bền bỉ của cô đã trở thành động lực lớn lao để nhóm chúng em nỗ lực không ngừng và đạt được thành quả này.

Bằng tất cả lòng biết ơn chân thành, tập thể nhóm xin gửi lời tri ân sâu sắc nhất đến cô – người đã không chỉ truyền đạt tri thức mà còn chắp cánh cho những khát vọng bay xa. Chúng em nguyện sẽ luôn ghi nhớ những bài học quý báu ấy và lấy đó làm hành trang vững chắc trên con đường tương lai.

**Nhóm thực hiện**

Nhóm 14

MỤC LỤC

[MỤC LỤC 3](#_Toc201024524)

[DANH MỤC CÁC CHỮ CÁI VIẾT TẮT 5](#_Toc201024525)

[DANH MỤC HÌNH ẢNH 6](#_Toc201024526)

[CHƯƠNG 1. TỔNG QUAN VỀ AN TOÀN THÔNG TIN 7](#_Toc201024527)

[1.1. Tổng quan về an toàn và bảo mật thông tin. 7](#_Toc201024528)

[1.1.1. Khái niệm. 7](#_Toc201024529)

[1.1.2. Các mối đe dọa đối với một hệ thống. 7](#_Toc201024530)

[1.1.3. Các loại tấn công an toàn thông tin. 8](#_Toc201024531)

[1.2. Giới thiệu về các hệ mật mã. 8](#_Toc201024532)

[1.2.1. Khái niệm. 8](#_Toc201024533)

[1.2.2. Các thành phần của hệ mã. 8](#_Toc201024534)

[1.2.3. Phân loại hệ mật mã. 9](#_Toc201024535)

[1.3. Kiến thức cơ sở. 10](#_Toc201024536)

[1.3.1. Cơ sở toán học. 10](#_Toc201024537)

[1.3.2. Thuật toán. 11](#_Toc201024538)

[1.3.3. Ngôn ngữ lập trình. 11](#_Toc201024539)

[1.4. Nội dung nghiên cứu. 12](#_Toc201024540)

[CHƯƠNG 2. KẾT QUẢ NGHIÊN CỨU 13](#_Toc201024541)

[2.1. Hệ mật ElGamal. 13](#_Toc201024542)

[2.1.1. Giới thiệu về hệ mật ElGamal. 13](#_Toc201024543)

[2.1.2. Giới thiệu về thuật toán ElGamal. 13](#_Toc201024544)

[2.2. Tìm hiểu về chữ ký số và chữ ký số ElGamal. 16](#_Toc201024545)

[2.2.1. Tìm hiểu về chữ ký số. 16](#_Toc201024546)

[2.2.2. Tìm hiểu về chữ ký số ElGamal. 16](#_Toc201024547)

[2.2.3. Độ an toàn của chữ ký ElGamal. 18](#_Toc201024548)

[2.3. Cơ sở toán học. 19](#_Toc201024549)

[2.3.1. Số học đồng dư. 19](#_Toc201024550)

[2.3.2. Tìm phần tử nghịch đảo. 20](#_Toc201024551)

[2.3.3. Kiểm tra số nguyên tố. 21](#_Toc201024552)

[2.4. Tìm hiểu về hàm băm. 22](#_Toc201024553)

[2.4.1. Giới thiệu hàm băm Hash 22](#_Toc201024554)

[2.4.2. Ứng dụng của hàm băm Hash. 23](#_Toc201024555)

[CHƯƠNG 3. KIẾN THỨC LĨNH HỘI VÀ BÀI HỌC KINH NGHIỆM 24](#_Toc201024556)

[3.1. Nội dung đã thực hiện 24](#_Toc201024557)

[3.2. Hướng phát triển 24](#_Toc201024558)

[TÀI LIỆU THAM KHẢO 25](#_Toc201024559)

**DANH MỤC CÁC CHỮ CÁI VIẾT TẮT**

|  |  |
| --- | --- |
| OSI | Open Systems Interconnection |
| ITU | International Telecommunication Union |
| IDE | Integrated Development Environment |
| UI | User Interface |
| MVC | Model-View-Controller |

**DANH MỤC HÌNH ẢNH**

[Hình 1.1: Quá trình mã hóa và giải mã. 10](#_Toc201023891)

[Hình 1.2: Mô hình mã hóa đối xứng. 10](#_Toc201023892)

[Hình 1.3: Mô hình mã hóa bất đối xứng. 11](#_Toc201023893)

[Hình 2.1: Các bước thực hiện tạo khóa. 15](#_Toc201023894)

[Hình 2.2: Các bước thực hiện mã hóa. 16](#_Toc201023895)

[Hình 2.3: Các bước thực hiện giải mã. 16](#_Toc201023896)

[Hình 2.4: Minh họa thương số và phần dư. 20](#_Toc201023897)

[Hình 2.5:Một số tính chất của đồng dư. 21](#_Toc201023898)

[Hình 2.6: Các phép toán của đồng dư. 21](#_Toc201023899)

[Hình 2.7: Ví dụ về hàm băm. 23](#_Toc201023900)

# TỔNG QUAN VỀ AN TOÀN THÔNG TIN

Tổng quan về an toàn và bảo mật thông tin.

### Khái niệm.

*Hệ thống thông tin* (IS – Information System): Hệ thống thông tin là một hệ thống tích hợp các thành phần (phần cứng, phần mềm, dữ liệu, v.v.) nhằm thu thập, lưu trữ, xử lý và chuyển giao thông tin, tri thức, và các sản phẩm số.

*An toàn thông tin*:Là việc bảo vệ thông tin khỏi truy nhập, sử dụng, tiết lộ, sửa đổi hoặc phá hủy trái phép. An toàn thông tin cũng bao gồm việc bảo vệ các thành phần hoặc hệ thống được sử dụng để quản lý, lưu trữ, xử lý và trao đổi thông tin.

*Bảo mật hệ thống thông tin*: Là việc bảo vệ hệ thống thông tin chống lại việc truy cập, sử dụng, chỉnh sửa, phá hủy, làm lộ và làm gián đoạn thông tin cũng như hoạt động của hệ thống một cách trái phép.

### Các mối đe dọa đối với một hệ thống.

Có ba hình thức đe dọa chủ yếu đối với một hệ thống:

* *Phá hoại*: Kẻ thù phá hỏng thiết bị phần cứng hoặc phần mềm hoạt động trên hệ thống.
* *Sửa đổi*: Tài sản của hệ thống bị sửa đổi trái phép, làm cho hệ thống không thực hiện đúng chức năng của nó. Ví dụ, thay đổi mật khẩu hoặc quyền người dùng trong hệ thống làm họ không thể truy cập vào hệ thống để làm việc.
* *Can thiệp*: Tài sản bị truy cập bởi những người không có thẩm quyền, các truyền thông thực hiện trên hệ thống bị ngăn chặn hoặc sửa đổi.

Các biện pháp ngăn chặn:

* *Điều khiển thông qua phần mềm*: Sử dụng các cơ chế an toàn bảo mật của hệ điều hành và các thuật toán mật mã học.
* *Điều khiển thông qua phần cứng*: Sử dụng các cơ chế bảo mật và thuật toán mật mã học được cứng hóa để sử dụng.
* *Điều khiển thông qua các chính sách của tổ chức*: Ban hành các quy định của tổ chức nhằm đảm bảo tính an toàn bảo mật của hệ thống.

### Các loại tấn công an toàn thông tin.

Có hai loại tấn công an ninh chính trong cả kiến trúc bảo mật X.800 (cho hệ thống OSI do ITU quy định) và tiêu chuẩn RFC 4949 (Request for Comment, bao gồm các thuật ngữ bảo mật Internet):

* *Tấn công thụ động*: Là cuộc tấn công cố gắng tìm hiểu hoặc sử dụng thông tin từ hệ thống nhưng không ảnh hưởng đến tài nguyên của hệ thống.
* *Tấn công chủ động*: Là cuộc tấn công mà kẻ tấn công cố gắng thay đổi tài nguyên hệ thống hoặc ảnh hưởng đến hoạt động của các hệ thống đó.

Các loại tấn công này đều đe dọa đến tính bảo mật và an toàn của hệ thống thông tin, đòi hỏi các biện pháp bảo vệ hiệu quả để đảm bảo tính toàn vẹn, bảo mật và sẵn sàng của dữ liệu cũng như các tài nguyên hệ thống.

## Giới thiệu về các hệ mật mã.

### Khái niệm.

Hệ mật mã là một tập hợp các thuật toán và giao thức được sử dụng để bảo vệ thông tin bằng cách biến đổi thông tin gốc (plaintext) thành một dạng khó hiểuvà ngược lại. Quá trình này chính là việc thực hiện mã hóa và giải mã.

* *Mã hóa*: Quá trình biến đổi thông tin từ dạng bản rõ (plaintext), dễ đọc và hiểu, sang dạng bản mã (ciphertext), khó hiểu và chỉ người có khóa giải mã mới có thể đọc được. Mục đích của mã hóa là che giấu và bảo vệ thông tin khi lưu trữ hoặc truyền tải.
* *Giải mã*: Quá trình ngược lại của mã hóa, biến đổi thông tin từ dạng bản mã trở lại dạng bản rõ. Điều này cho phép người nhận có thể đọc và hiểu được thông tin ban đầu.

### Các thành phần của hệ mã.

Một hệ mật mã là một hệ bao gồm 5 thành phần (P, C, K, E, D) thoả mãn:

* P (Plaintext): Không gian bản rõ, là tập hợp hữu hạn các bản rõ có thể.
* C (Ciphertext): Không gian bản mã, là tập hợp những bản mã có thể.
* K (Key): Không gian khoá, là tập hợp các khoá có thể.
* E (Encryption): Không gian hàm mã hoá, là tập hợp các quy tắc mã hoá có thể.
* D (Decryption): Không gian hàm giải mã, là tập hợp các quy tắc giải mã có thể.

A diagram of a flowchart

Description automatically generated

Hình 1.1: Quá trình mã hóa và giải mã.

### Phân loại hệ mật mã.

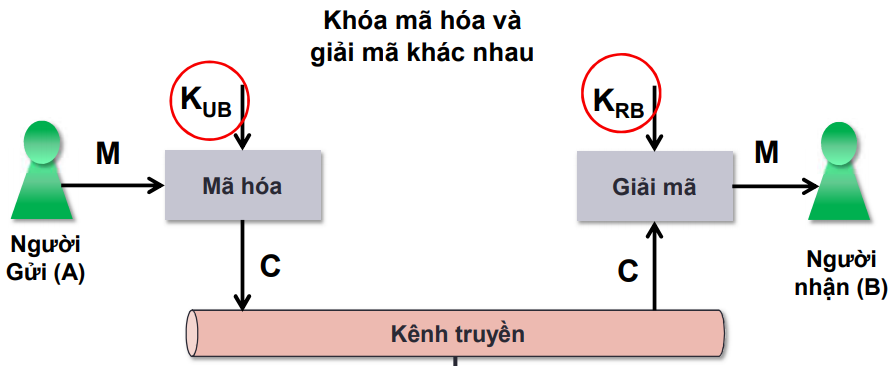
*Hệ mật mã đối xứng*: sử dụng cùng một khóa cho cả quá trình mã hóa và giải mã dữ liệu. Điều này có nghĩa là khóa phải được giữ bí mật tuyệt đối để đảm bảo an toàn cho thông tin. Một số thuật toán nổi tiếng trong hệ mật mã đối xứng bao gồm: DES, 3-DES, AES, ...

A screenshot of a phone

AI-generated content may be incorrect.

Hình 1.2: Mô hình mã hóa đối xứng.

*Hệ mật mã bất đối xứng*: sử dụng hai khóa khác nhau, một khóa để mã hóa và một khóa khác để giải mã. Hai khóa này tạo nên một cặp khóa liên kết ngược nhau, và không thể suy luận được khóa này từ khóa kia. Khóa dùng để mã hóa là khóa công khai, khóa dùng để giải mã là khóa bí mật. Một số thuật toán nổi tiếng trong hệ mật mã bất đối xứng bao gồm: RSA, ElGamal, …



Hình 1.3: Mô hình mã hóa bất đối xứng.

## Kiến thức cơ sở.

### Cơ sở toán học.

Cơ sở toán học là một phần không thể thiếu của các hệ mật mã, dưới đây là một số cơ sở toán học:

*- Các phép toán đồng dư*: là một phần quan trọng của lý thuyết số và toán học nói chung, đặc biệt trong lĩnh vực toán học tổ hợp, lý thuyết số, và mật mã học. Phép toán đồng dư liên quan đến việc so sánh hai số nguyên để xem chúng có cùng một phần dư khi chia cho một số nguyên dương cho trước hay không. Đây là khái niệm cơ bản trong hệ thống số học đồng dư, được sử dụng để giải quyết nhiều vấn đề toán học phức tạp một cách đơn giản hơn.

*- Tìm phần tử nghịch đảo*: là một khái niệm quan trọng trong lĩnh vực số học và mật mã học, đặc biệt trong các hệ thống mật mã dựa trên số nguyên tố và phép toán modulo. Phần tử nghịch đảo là một số nguyên trong hệ số nguyên tố mà khi nhân với một số khác trong hệ đó, kết quả sẽ cho phần dư là 1 khi chia cho một số nguyên dương nào đó. Việc tìm phần tử nghịch đảo là quan trọng trong quá trình mã hóa và giải mã trong các thuật toán mật mã khóa công khai, như ElGamal, RSA, DSA, …

### Thuật toán.

Nhiều thuật toán được áp dụng trong việc tối ưu và sử dụng các hệ mật, nổi bật với một số thuật toán sau:

- *Thuật toán về kiểm tra số nguyên tố*: để xác định xem một số nguyên dương có phải là số nguyên tố hay không. Số nguyên tố là các số chỉ có hai ước số dương là 1 và chính nó.

- *Thuật toán Euclid mở rộng*: để tìm ước số chung lớn nhất (GCD - Greatest Common Divisor), được sử dụng để tính toán các hệ số Bézout, là một phương pháp quan trọng trong toán học. Nó cho phép tính toán ước số chung lớn nhất của hai số nguyên dương a và b cùng với các hệ số nguyên x, y. Đây được xem là một thuật toán nổi bật cần dùng áp dụng cho nhiều hệ mật.

### Ngôn ngữ lập trình.

Hiện có rất nhiều ngôn ngữ lập trình để phù hợp để sử dụng trong mật mã và bảo mật các tài liệu quan trọng, dưới đây là một vài ngôn ngữ lập trình nổi bật:

* Python là ngôn ngữ lập trình bậc cao, nổi bật với cú pháp đơn giản và dễ học. Với khả năng ứng dụng linh hoạt, Python được sử dụng rộng rãi trong các lĩnh vực như phát triển phần mềm, xử lý dữ liệu, trí tuệ nhân tạo và bảo mật thông tin. Ngôn ngữ này cung cấp nhiều thư viện mạnh mẽ hỗ trợ các tác vụ phức tạp, đặc biệt trong mật mã học và bảo mật.
* C**++** là một ngôn ngữ lập trình bậc trung, nổi bật với hiệu suất cao và khả năng tối ưu hóa bộ nhớ. C++ cung cấp nhiều thư viện hỗ trợ cho các thuật toán mật mã, đặc biệt là thư viện Crypto++, cho phép triển khai các thuật toán như ElGamal, RSA, SHA và các phép toán trên số nguyên lớn. Với khả năng kiểm soát chi tiết và tốc độ xử lý nhanh, C++ rất phù hợp và sử dụng tốt khi xây dựng các hệ thống cần xử lý dữ liệu lớn hoặc yêu cầu tính bảo mật cao.

## Nội dung nghiên cứu.

Trong bối cảnh an toàn thông tin ngày càng trở nên quan trọng, việc xây dựng các hệ thống bảo mật đáng tin cậy là một yêu cầu tất yếu. Việc sử dụng các hệ mật vào trong bảo mật dữ liệu cũng trở thành một phần tất yếu để đảm bảo sự bí mật. Do vậy việc xây dựng một sản phẩm tạo chữ ký số và giải mã chữ ký số bằng hệ mật ElGaml sẽ là một phần đóng góp cho việc an toàn và bảo mật thông tin và sản phẩm sẽ đi sau vào nghiên cứu:

* Tìm hiểu về hệ mật ElGamal: Tìm hiểu nguyên lý và đặc điểm của hệ mật ElGamal.
* Chữ ký số ElGamal: Cách thức hoạt động và quy trình tạo, xác minh chữ ký số ElGamal.
* Độ an toàn của chữ ký ElGamal: Đánh giá tính bảo mật và các ưu, nhược điểm của chữ ký số ElGamal.
* Tìm hiểu về hàm băm: Tổng quan về hàm băm, vai trò, các tính chất quan trọng, một số hàm băm được ứng dụng nhiều và việc ứng dụng hàm băm vào trong bảo mật thông tin và xác thực dữ liệu.

# KẾT QUẢ NGHIÊN CỨU

## Hệ mật ElGamal.

### Giới thiệu về hệ mật ElGamal.

ElGamal là một hệ mật mã khóa công khai, nghĩa là người gửi có thể mã hóa thông điệp bằng khóa công khai của người nhận, và chỉ người nhận với khóa bí mật mới có thể giải mã được.

Hệ thống này được xây dựng dựa trên độ khó của bài toán logarit rời rạc – một bài toán mà hiện tại chưa có thuật toán giải hiệu quả trong thời gian hợp lý.

ElGamal là một biến thể của sơ đồ Diffie-Hellman – sơ đồ nổi tiếng trong việc trao đổi khóa, và được đề xuất bởi Taher ElGamal vào năm 1984.

So với RSA, ElGamal không dính đến các vấn đề bản quyền, nên được sử dụng tự do hơn trong thực tiễn.

### Giới thiệu về thuật toán ElGamal.

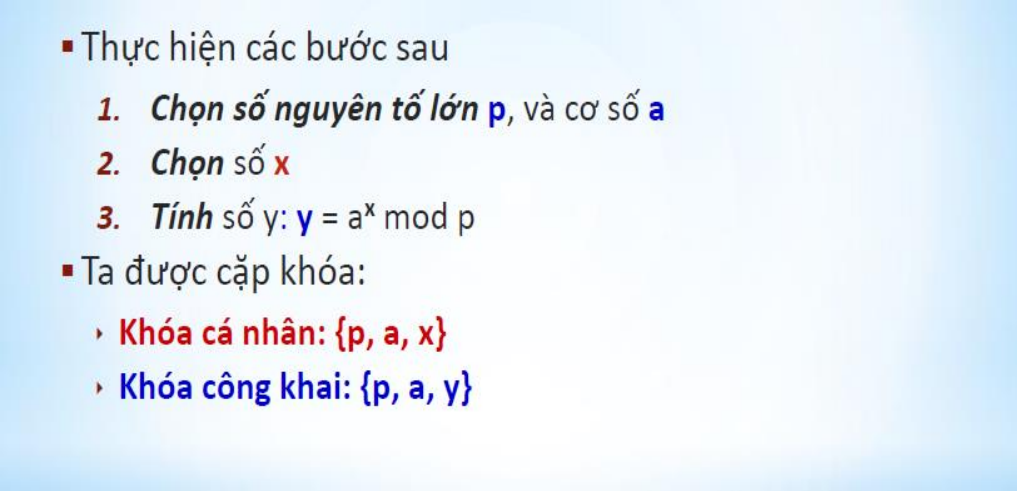
Thuật toán của ElGamal ElGamal sẽ gồm 3 phần:

* Tạo khóa.
* Mã hóa.
* Giải mã.

#### Tạo khóa.

Để tạo được cặp khóa là khóa bí mật và khóa công khai dùng cho mã hóa và giải mã thì sẽ bao gồm 5 bước:

* Bước 1: Chọn một số nguyên tố lớn p (số này xác định không gian làm việc).
* Bước 2: Chọn a là một căn nguyên thủy của p (a < p , a và p là số nguyên tố cùng nhau gcd(a, p) = 1.
* Bước 3: Chọn một số bí mật XA, thỏa mãn 1 ≤ XA ≤ p - 2. Đây là một phần của khóa bí mật.
* Bước 4: Bước 4: Tính YA = a XA mod p
* Bước 5: Xác định khóa:
* Khóa công khai : {q,a,Ya}
* Khóa bí mật : {XA}



Hình 2.1: Các bước thực hiện tạo khóa.

#### Mã hóa.

Sau khi tạo khóa thành công, ta sử dụng khóa đó để có thể mã hóa và nó gồm 5 bước như sau:

* Bước 1: Thông điệp ban đầu M, sao cho 0 < M < q.
* Bước 2: Chọn ngẫu nhiên k (k <q).
* Bước 3: Tính K =( YA )k mod q.
* Bước 4: Tính thành phần thứ nhất C1 = ak mod q .
* Bước 5: Tính thành phần thứ hai C2 = K .M mod q.

Và sau khi mã hóa xong ta sẽ thu được bản mã là cặp {C1, C2}.

A screenshot of a computer

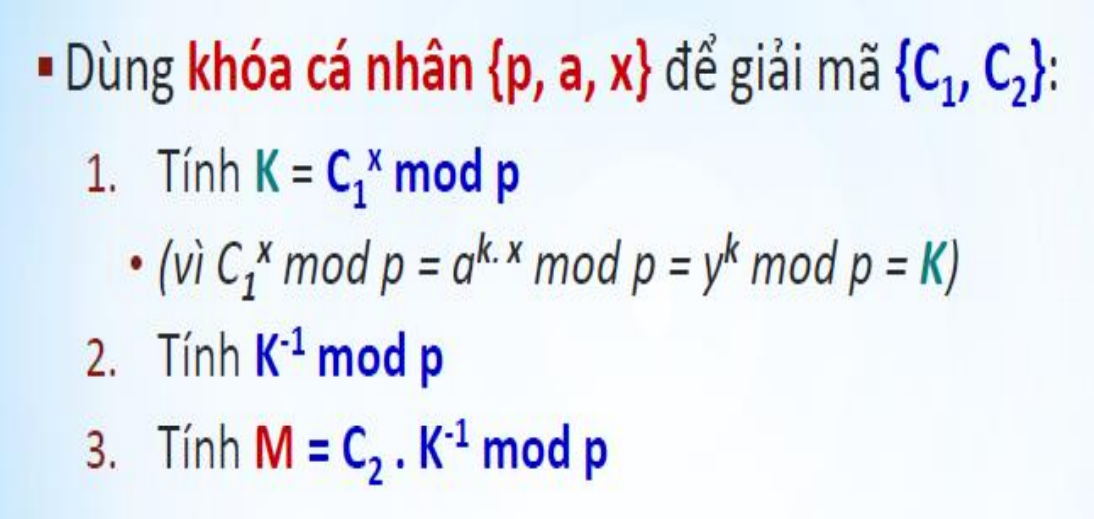
AI-generated content may be incorrect.

Hình 2.2: Các bước thực hiện mã hóa.

#### Giải mã.

Sau khi mã hóa thành công, ta sử dụng bản mã thu được ở trước để giải mã và nó bao gồm các bước sau:

* Bước 1: Bản mã {C1 , C2}.
* Bước 2: K = (C1) XA mod q.
* Bước 3: M = C2 K-1 mod q.



Hình 2.3: Các bước thực hiện giải mã.

## Tìm hiểu về chữ ký số và chữ ký số ElGamal.

### Tìm hiểu về chữ ký số.

Chữ ký số là chữ ký điện tử sử dụng thuật toán khóa không đối xứng, gồm khóa bí mật và khóa công khai, trong đó khóa bí mật được dùng để ký số và khóa công khai được dùng để kiểm tra chữ ký số.

Chữ ký số bảo đảm tính xác thực, tính toàn vẹn và tính chống chối bỏ nhưng không bảo đảm tính bí mật của thông điệp dữ liệu.

Chữ ký số có những đặc điểm và vai trò quan trọng sau:

* Tính xác thực: Chữ ký số giúp xác định danh tính của người ký, đảm bảo rằng thông điệp dữ liệu thực sự được tạo ra hoặc chấp thuận bởi danh tính người tạo ra.
* Tính toàn vẹn: Chữ ký số đảm bảo rằng nội dung của thông điệp dữ liệu không bị thay đổi kể từ khi được ký, giúp người nhận tin tưởng vào tính chính xác của thông tin.
* Tính chống chối bỏ: Chữ ký số là bằng chứng pháp lý, giúp người ký không thể phủ nhận việc đã ký thông điệp dữ liệu.
* Thay thế chữ ký tay và con dấu: Trong nhiều trường hợp, chữ ký số có giá trị pháp lý tương đương với chữ ký tay hoặc con dấu trong các giao dịch điện tử, giúp tiết kiệm thời gian và chi phí.

### Tìm hiểu về chữ ký số ElGamal.

Sơ đồ chữ ký ElGamal là sơ đồ chữ ký số dựa trên các tính chất đại số của lũy thừa mô đun, cùng với bài toán logarit rời rạc.

Sau đây ta sẽ mô tả sơ đồ chữ kí số Elgamal đã từng giới thiệu trong một báo cáo năm 1985. Bản cải tiến của sơ đồ này đã được Viện Tiêu chuẩn và Công nghệ Quốc gia Mỹ (NIST) chấp nhận làm chuẩn chữ kí số. Sơ đồ Elgamal được thiết kế với mục đích dành riêng cho chữ kí số. Khác với sơ đồ RSA dùng cho cả mã khoá công khai lẫn chữ kí số.

Sơ đồ Elgamal được thiết kế với mục đích dành riêng cho chữ kí số, điểm mạnh của nó là cùng số nguyên tố p trong cùng một sơ đồ thì với k là ngẫu nhiên nên ta có thể có nhiều chữ kí số, không tất định giống như hệ thống mã khoá công khai Elgamal, ở sơ đồ chữ kí RSA ta chỉ thấy trên cùng một sơ đồ với cùng một số nguyên tố p thì ta chỉ có một chữ kí số. Điều này có nghĩa là có nhiều chữ kí hợp lệ trên bức điện cho trước bất kì. Thuật toán xác minh phải có khả năng chấp nhận bất kì chữ kí hợp lệ nào khi xác thực chữ kí đó.

Sơ đồ chữ ký Elgamal :

1. **Tạo cặp khóa (bí mật, công khai) (x, y):**

Chọn số nguyên tố **p** sao cho bài toán logarit rời rạc trong **Zp** là “khó” giải.

Chọn phần tử nguyên thuỷ **g** ∈ **Z­p**\*. Đặt **P** = **Zp\***, **A** = **Zp**\* x **Zp-1**.

Chọn khóa bí mật là **x**∈ **Zp**\*. Tính khóa công khai y ≡ **gx** mod **p**.

Định nghĩa tập khóa:

* + Khóa công khai = { y }.
  + Khóa bí mật = { x }.
  + p, g là các tham số công khai.

1. **Ký số:**

Thông điệp ban đầu: M

Chọn ngẫu nhiên số k thỏa mãn điều kiện: 1< k < p -1 và gcd (k, p-1) = 1.

Tính thành phần thứ nhất: r = gk mod p.

Tính thành phần thứ hai: s = k-1 \* (H(M) – x\*r) mod(p-1) .

Trong đó H(.) là hàm băm.

1. Kiểm tra chữ ký:

Tính thành phần thứ nhất: u = gH(M) mod p.

Tính thành phần thứ hai: v = (yr \* rs) mod p.

Kiểm tra: nếu (u = v) thì chữ ký (r, s) là hợp lệ ngược lại là giả mạo.

### Độ an toàn của chữ ký ElGamal.

Chữ ký ElGamal là một hệ thống chữ ký số an toàn, dựa trên bài toán logarit rời rạc – một bài toán được coi là khó giải trong lĩnh vực mật mã học hiện đại. Tuy nhiên, tính an toàn của chữ ký ElGamal phụ thuộc vào việc lựa chọn đúng các tham số và cách sử dụng. Dưới đây là các tình huống ảnh hưởng đến độ an toàn của hệ thống:

a) Trường hợp kẻ tấn công tính được khóa bí mật x hoặc số ngẫu nhiên k.

Để phá vỡ chữ ký ElGamal, kẻ tấn công cần phải tính được một trong hai giá trị:

* Khóa bí mật x từ khóa công khai y = g^x mod p.
* Số ngẫu nhiên k từ chữ ký r = g^k mod p.

Cả hai bài toán trên đều là bài toán logarit rời rạc. Đây là bài toán rất khó giải khi số nguyên tố p đủ lớn. Hiện tại chưa có thuật toán hiệu quả nào giải được bài toán này trong thời gian ngắn. Do đó, khả năng tấn công theo hướng này là rất thấp.

b) Trường hợp cố gắng khôi phục thông điệp từ chữ ký m.

Giả sử kẻ tấn công không có khóa bí mật x hoặc số ngẫu nhiên k, họ sẽ phải giải:

* m = c × y^(-k) mod p (nếu không biết x)
* m = c × r^(-x) mod p (nếu không biết k)

Các công thức trên cho thấy muốn tìm lại được thông điệp m, kẻ tấn công vẫn cần biết x hoặc k, nghĩa là phải giải bài toán logarit rời rạc như đã nói ở trên. Do đó, cách tấn công này cũng không khả thi nếu các tham số được chọn đúng cách.

c) Nguy cơ khi sử dụng lại số ngẫu nhiên k

Một trong những điểm yếu quan trọng của ElGamal là nếu sử dụng lại cùng một số k để ký nhiều thông điệp, hệ thống sẽ bị lộ thông tin. Giả sử hai thông điệp m₁ và m₂ được ký với cùng k, và bản mã tương ứng là:

* (c₁, r) với m₁.
* (c₂, r) với m₂ (cùng giá trị r do cùng k).

Khi đó có thể tính: m₁ ≡ (m₂ × c₁ × c₂⁻¹) mod p.

Nếu kẻ tấn công biết một trong hai thông điệp, họ có thể dễ dàng tính ra thông điệp còn lại. Đây là lỗ hổng nghiêm trọng nếu k không được chọn ngẫu nhiên cho mỗi lần ký.

***Kết luận về độ an toàn:***

Hệ mật Elgamal được xây dựng dựa trên bài toán logarit rời rạc. Do đó, tính an toàn của hệ mật tùy thuộc vào độ phức tạp của bài toán logarit. Khi chọn số nguyên tố p đủ lớn, thuật toán mã hóa Elgamal chưa có phương pháp thám mã hiệu quả nếu các tham số được lựa chọn đúng.

## Cơ sở toán học.

### Số học đồng dư.

Định nghĩa: Cho một số nguyên *a* và số nguyên dương *n* bất kỳ, thực hiện phép chia *a* cho *n* thì thu được thương số *q* và phần dư *r* thỏa mãn mối quan hệ sau:

A math equations on a white background

AI-generated content may be incorrect.

Hình 2.4: Minh họa thương số và phần dư.

Tóm lại, cho một số nguyên *a* và số nguyên dương *n* thì ta định nghĩa *a mod n* là phần dư của phép chia *a* cho *n*

*Ví dụ: 13 mod 4 = 1 và -13 mod 4 = 3*

Hai số nguyên a và b được gọi là đồng dư modulo với n nếu *(a mod n) = (b mod*

*n)* và được ký hiệu như sau: *a ≡ b(mod n).*

Ví dụ: 13 ≡ 5 (mod 4) vì 13 mod 4 = 1 và 5 mod 4 = 1

7 ≡ -13 (mod 4) vì 7 mod 4 = 3 và - 13 mod 4 = 3

Một số tính chất của đồng dư:

A table with math equations

AI-generated content may be incorrect.

Hình 2.5:Một số tính chất của đồng dư.

Các phép toán trên module:

A white paper with black text

AI-generated content may be incorrect.

Hình 2.6: Các phép toán của đồng dư.

### Tìm phần tử nghịch đảo.

Để tìm phần tử nghịch đảo của một số nguyên a trong hệ ℤₙ (modulo n), ta cần xác định một số x sao cho: a × x ≡ 1 (mod n).  
Khi đó, x là phần tử nghịch đảo của a, ký hiệu là a⁻¹. Điều kiện cần để tồn tại nghịch đảo là gcd(a, n) = 1. Với n nhỏ, có thể tìm nghịch đảo bằng cách thử từng giá trị. Tuy nhiên, khi n lớn cách hiệu quả hơn là dùng thuật toán Euclid mở rộng. Thuật toán này không chỉ tìm gcd(a, n) mà còn cho ta hệ số x và y sao cho:  
a·x + n·y = gcd(a, n)  
Nếu gcd(a, n) = 1, thì x chính là nghịch đảo của a mod n.

Thuật toán Euclid mở rộng (EXTENDED EUCLID) được trình bày như sau:

1. (A1, A2, A3) ← (1, 0, n); (B1, B2, B3) ← (0, 1, a)

2. Lặp lại cho đến khi B3 = 0:

a. Nếu B3 = 1, thì trả về B2 chính là phần tử nghịch đảo a⁻¹ mod n

b. Tính thương nguyên: Q = A3 div B3

c. Tính (R1, R2, R3) = (A1 - Q×B1, A2 - Q×B2, A3 - Q×B3)

d. Gán lại: (A1, A2, A3) ← (B1, B2, B3); (B1, B2, B3) ← (R1, R2, R3)

3. Nếu B3 = 0 và A3 ≠ 1 → không tồn tại phần tử nghịch đảo (a và n không nguyên tố cùng nhau).

Thuật toán này trả về giá trị B2 là phần tử nghịch đảo a⁻¹ mod n nếu tồn tại. Đây là công cụ rất quan trọng trong các hệ mật mã như RSA và ElGamal, nơi việc tính nghịch đảo trong tập số lớn là điều kiện bắt buộc để thực hiện giải mã hay ký số.

### Kiểm tra số nguyên tố.

Trong một số thuật toán mã hóa ta cần chọn một hoặc một số nguyên tố rất lớn một cách ngẫu nhiên. Như vậy, ta cần phải giải quyết bài toán xác định liệu xem một số lớn có phải là số nguyên tố hay không. Như vậy, việc áp dụng thuật toán kiểm tra một số nhỏ là nguyên tố hay không để kiểm tra một số rất lớn là không phù hợp. Để kiểm tra một số lớn có phải là nguyên tố hay không ta dựa vào hai tính chất của số nguyên tố.

Tính chất thứ nhất: Nếu *p* là số nguyên tố và *a* là một số nguyên dương nhỏ hơn *p* thì *a2 mod p = 1* khi và chỉ khi *a mod p = 1* hoặc *a mod p = p – 1*.

Thật vậy, vì theo tính chất số học của modulo (a mod p) × (a mod p) =a2 mod p. Do đó, nêu *a mod p = 1* thì *a2 mod p = 1* và *a mod p = p – 1* thì:

*a2 mod p = (p – 1)2 mod p = (p2 – 2.p + 1) mod p = 1*.

Tính chất thứ hai: Nếu p là số nguyên tố lớn hơn 2 thì ta có thể phân tích

p - 1 = 2k × q, với k > 0 và q là số lẻ. Gọi a là một số nguyên bất kỳ trong

phạm vi 1 < a < p – 1. Khi đó, một trong hai điều kiện sau đây được thỏa

mãn:

* aq  ≡ 1(modp)
* Một trong các số aq, a2q,a4q … , a 2(k-1)q đồng dư với *1* modulo *p*

Thuật toán Miller-Rabin nhận tham số đầu vào là một số nguyên và trả về kết quả số cần kiểm tra là hợp số không phải số nguyên tố hoặc có thể là số nguyên tố.

## Tìm hiểu về hàm băm.

### Giới thiệu hàm băm Hash

Hàm băm (Hash function) là một hàm toán học chuyển đổi một thông điệp đầu vào có độ dài bất kỳ thành một dãy bit có độ dài cố định (tuỳ thuộc vào thuật toán băm). Dãy bit này được gọi là thông điệp rút gọn (message digest) hay giá trị băm (hash value), đại diện cho thông điệp ban đầu.

*Hàm băm:* là hàm một chiều mà nếu đưa một lượng dữ liệu bất kì qua hàm này sẽ cho ra một chuỗi có độ dài cố định ở đầu ra.

**A diagram of a person's body

AI-generated content may be incorrect.**

Hình 2.7: Ví dụ về hàm băm.

*Mã hóa SHA*: mã hóa SHA có tên đầy đủ là Secure Hash Algorithm hay còn gọi là thuật giải băm an toàn. Đây là tổ hợp 5 thuật giải băm mật được phát triển bởi NSA (National Security Agency) – Cục An ninh Quốc gia Mỹ và được xuất bản thành chuẩn của chính phủ Mỹ bởi NIST (National Institute of Standards and Technology) – Viện Công nghệ và chuẩn quốc gia Mỹ.

Thuật giải SHA đã được chấp nhận bởi FIPS (Federal Information Processing Standards) – Tiêu chuẩn Xử lý Thông tin Liên bang. Mã hóa SHA được dùng để chuyển đổi từ một đoạn dữ liệu nhất định thành một đoạn dữ liệu có chiều dài không đổi nhưng với xác suất khác biệt cao.

Năm thuật giải SHA chuẩn bao gồm:

* SHA-1: Trả lại kết quả dài 160 bit.
* SHA-224: Trả lại kết quả dài 224 bit.
* SHA-256: Trả lại kết quả dài 256 bit.
* SHA-384: Trả lại kết quả dài 384 bit.
* SHA-512: Trả lại kết quả dài 512 bit.

### Ứng dụng của hàm băm Hash.

*Lưu trữ mật khẩu*: Hàm băm mã hóa mật khẩu người dùng trước khi lưu trữ, giúp bảo mật thông tin.

*Kiểm tra toàn vẹn dữ liệu*: Dùng để kiểm tra tính toàn vẹn của tệp hoặc dữ liệu trong quá trình truyền tải, đảm bảo không bị thay đổi.

*Chữ ký số*: Hàm băm kết hợp với khóa bí mật để tạo chữ ký số, bảo vệ tính toàn vẹn và xác thực của tài liệu.

Ngoài ra, hàm băm còn được ứng dụng trong blockchain, cơ sở dữ liệu, xác thực thông điệp, tạo mã kiểm tra lỗi và tìm kiếm dữ liệu. Nhờ vào các tính chất như tính một chiều, khả năng phân tán và nhanh chóng, hàm băm là công cụ quan trọng trong bảo mật và quản lý dữ liệu.

# KIẾN THỨC LĨNH HỘI VÀ BÀI HỌC KINH NGHIỆM

## Nội dung đã thực hiện

**Bài học kinh nghiệm**

Do thời gian có hạn nên đề tài của nhóm vẫn chưa được hoàn chỉnh nhất, chưa tìm hiểu được hết những vấn đề và bài toán phát sinh trong đề bài. Kèm theo có nhiều những mặt hạn chế còn tồn tại ở trong ứng dụng, từ đó nhóm đã rút ra được các bài học quan trọng để cải thiện dự án trong tương lai.

## Hướng phát triển

TÀI LIỆU THAM KHẢO

**Tiếng Việt**

**Tiếng Anh**