BỘ GIÁO DỤC VÀ ĐÀO TẠO

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

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

**🙞 🕮 🙜**



**ĐỒ ÁN**

**AN TOÀN BẢO MẬT THÔNG TIN**

**Đề tài**

**HÀM BĂM VÀ CÁC ỨNG DỤNG**

**Giáo viên hướng dẫn: Sinh viên thực hiện:**

**ThS.Phan Bích Chung Nguyễn Hoàng Điển B2113329**

**Lê Ngọc Tân B2113319**

**Phan Minh Tài B2113341**

Cần Thơ, 03/2025

**NHẬN XÉT CỦA GIẢNG VIÊN**

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

**LỜI CẢM ƠN**

Để có được bài báo cáo này, em xin được bày tỏ lòng biết ơn chân thành và sâu sắc đến Cô Phan Bích Chung – người đã trực tiếp tận tình hướng dẫn giúp đỡ em. Trong suốt quá trình thực hiện niên luận, nhờ những sự chỉ bảo và hướng dẫn quý giá đó mà bài báo cáo này được hoàn thành một cách tốt nhất.

|  |  |
| --- | --- |
|  | Cần Thơ, ngày 08 tháng 03 năm 2025  Người viết  Nguyễn Hoàng Điển  Phan Minh Tài  Lê Ngọc Tân |

|  |  |  |
| --- | --- | --- |
| **Họ và tên** | **Công việc** | **Mức độ hoàn thành** |
| Nguyễn Hoàng Điển – B2113329 | * Tìm hiểu về giải thuật MD5 | 100% |
| Lê Ngọc Tân – B2113319 | * Tìm hiểu về giải thuật SHA-256 | 100% |
| Phan Minh Tài – B2113341 | * Tìm hiểu về giải thuật Argon2 | 100% |

**BẢNG PHÂN CHIA CÔNG VIỆC**

**CAM KẾT**

Nhóm chúng em cam kết đồ án này là do nhóm chúng em thực hiện, không sử dụng sản phẩm của người khác. Nếu sử dụng sản phẩm của người khác, chúng em sẽ nhận toàn bộ trách nhiệm.

Cần Thơ, ngày 08 tháng 03 năm 2025

Ký tên

Nguyễn Hoàng Điển

Lê Ngọc Tân

Phan Minh Tài

**MỤC LỤC**

[PHẦN NỘI DUNG 9](#_Toc192434485)

[CHƯƠNG 1 9](#_Toc192434486)

[TỔNG QUAN VỀ AN TOÀN BẢO MẬT THÔNG TIN VÀ MẬT MÃ 9](#_Toc192434487)

[1. Giới thiệu môn học 9](#_Toc192434488)

[2. Tổng quan về an toàn thông tin hiện nay 9](#_Toc192434489)

[2.1. Sự gia tăng của các cuộc tấn công mạng. 9](#_Toc192434490)

[2.2. Các phương pháp bảo vệ dữ liệu quan trọng. 10](#_Toc192434491)

[2.3. Vai trò của mật mã trong bảo mật dữ liệu. 11](#_Toc192434492)

[2.4.Xu hướng phát triển của các thuật toán mật mã và bảo mật hiện đại. 11](#_Toc192434493)

[CHƯƠNG 2 13](#_Toc192434494)

[TÌM HIỂU VỀ CÁC HÀM BĂM 13](#_Toc192434495)

[1. Giới thiệu chung về hàm băm 13](#_Toc192434496)

[1.1. Hàm băm là gì? 13](#_Toc192434497)

[1.2. Các loại hàm băm 13](#_Toc192434498)

[1.3. Tiêu chí đánh giá hàm băm tốt 13](#_Toc192434499)

[2. Hàm băm MD5 13](#_Toc192434500)

[2.1. Giới thiệu 13](#_Toc192434501)

[2.1.1. MD5 (Message Digest Algorithm 5) là gì? 13](#_Toc192434502)

[2.1.2. Lịch sử phát triển 15](#_Toc192434503)

[2.1.3. Hạn chế và lỗ hỏng của MD5 15](#_Toc192434504)

[2.2. Nguyên lý thuật toán và các hoạt động MD5 16](#_Toc192434505)

[2.3. Các ứng dụng sử dụng giải thuật MD5 19](#_Toc192434506)

[2.3.1. Kiểm tra tính toàn vẹn dữ liệu 19](#_Toc192434507)

[2.3.2. Quản lý bộ nhớ đệm và chỉ mục dữ liệu 20](#_Toc192434508)

[2.3.3. Xác định các tệp tin trong các hệ thống lưu trữ đám mây 20](#_Toc192434509)

[3. Hàm băm SHA-256 21](#_Toc192434510)

[3.1. Giới thiệu 21](#_Toc192434511)

[3.2. Nguyên lý thuật toán và các hoạt động SHA-256 21](#_Toc192434512)

[3.3. Các ứng dụng sử dụng giải thuật SHA-256 hiện tại 23](#_Toc192434513)

[3.3.1. Blockchain và tiền mã hóa 23](#_Toc192434514)

[3.3.2. Tăng cường bảo mật Blockchain bằng thuật toán SHA-256 23](#_Toc192434515)

[4. Hàm băm Argon2 24](#_Toc192434516)

[4.1. Giới thiệu 24](#_Toc192434517)

[4.2. Nguyên lý thuật toán và các hoạt động Argon2 24](#_Toc192434518)

[4.2.1. Các tham số đầu vào 24](#_Toc192434519)

[4.2.2. Hàm băm Blake2b 25](#_Toc192434520)

[4.2.3. Hàm nén 25](#_Toc192434521)

[4.2.4. Cấu trúc và hoạt động của Argon2 26](#_Toc192434522)

[4.2.5. Ưu nhược điểm của 2 phiên bản 28](#_Toc192434523)

[4.3. Các ứng dụng sử dụng giải thuật Argon2 hiện tại 28](#_Toc192434524)

[CHƯƠNG 3 29](#_Toc192434525)

[ỨNG DỤNG QUẢN LÝ NHÂN SỰ SỬ DỤNG HÀM BĂM ĐỂ BẢO VỆ THÔNG TIN MẬT KHẨU VÀ THÔNG TIN NGƯỜI DÙNG 29](#_Toc192434526)

[1.Giới thiệu về hệ thống 29](#_Toc192434527)

[2. Giới thiệu giao diện 29](#_Toc192434528)

[PHẦN KẾT LUẬN 32](#_Toc192434529)

[1. Kết quả đạt được 32](#_Toc192434530)

[2. Hạn chế 32](#_Toc192434531)

[3. Hướng phát triển 32](#_Toc192434532)

[TÀI LIỆU THAM KHẢO 33](#_Toc192434533)

**DANH MỤC HÌNH**

[Hình 2.1 Sơ đồ tổng quát MD5 14](#_Toc192434551)

[Hình 2.2 Sơ đồ minh họa kết quả đầu vào đầu ra của giải thuật MD5 14](#_Toc192434552)

[Hình 2.3 Độ dài của tin nhắn sau khi đệm (tính bằng bit) 15](#_Toc192434553)

[Hình 2.4. Sơ đồ hoạt động tổng quá cho quá trình băm MD5 16](#_Toc192434554)

[Hình 2.5 Minh học cho quá trình thực hiện bước 1 thêm bit đệm của MD5 17](#_Toc192434555)

[Hình 2.6 Cấu trúc của hàm F trong quy trinh hoạt động ở hình 2.4 18](#_Toc192434556)

[Hình 2.7 Sơ đồ quy trình hoạt động của MD5 19](#_Toc192434557)

[Hình 2.8. Các giá trị băm khởi tạo được chọn bằng cách lấy 32 bit 22](#_Toc192434558)

[Hình 2.9. Thuật toán vòng lặp xử lý trên từng word 22](#_Toc192434559)

[Hình 2.10. Thuật toán vòng lặp xử lý trên từng block 23](#_Toc192434560)

[Hình 2.11. Hàm nén G của Argon2 [11] 25](#_Toc192434561)

[Hình 2.12. Một chu kỳ lắp đầy bộ nhớ với p lane và 4 slice của Argon2 [11] 26](#_Toc192434562)

[Hình 2.13. Cách cập nhật giá trị của các khối trong ma trận ở lần chạy đầu tiên [11] 27](#_Toc192434563)

[Hình 3.1. Giao diện đăng ký của hệ thống 29](#_Toc192434564)

[Hình 3.2. Giao diện đăng nhập 30](#_Toc192434565)

[Hình 3.3. Giao diện trang chủ của hệ thống 30](#_Toc192434566)

[Hình 3.4. Giao diện chức năng quản lý tài khoản 31](#_Toc192434567)

**DANH MỤC CÁC TỪ VIẾT TẮT**

|  |  |  |
| --- | --- | --- |
| **STT** | **Từ viết tắt** | **Diễn giải** |
| 1 | SHA | Secure Hash Algorithm |
| 2 | MD5 | Message-Digest algorithm 5 |
| 3 | PRNG | Pseudo-random number generator |

# PHẦN NỘI DUNG

# CHƯƠNG 1

# TỔNG QUAN VỀ AN TOÀN BẢO MẬT THÔNG TIN VÀ MẬT MÃ

## 1. Giới thiệu môn học

Trong kỷ nguyên số hiện nay, thông tin và dữ liệu đóng vai trò quan trọng trong hầu hết các lĩnh vực của đời sống, từ kinh doanh, tài chính, y tế cho đến chính phủ và quốc phòng. Việc bảo vệ dữ liệu trước các mối đe dọa từ bên ngoài, đặc biệt là các cuộc tấn công mạng ngày càng gia tăng, trở thành một trong những thách thức lớn đối với các tổ chức và cá nhân.

Môn học "An toàn bảo mật thông tin và mật mã" nhằm cung cấp những kiến thức nền tảng về các phương pháp bảo mật dữ liệu, các thuật toán mã hóa, cũng như các kỹ thuật phòng chống tấn công mạng. Thông qua môn học này, người học sẽ hiểu rõ về tầm quan trọng của bảo mật thông tin, cách thức vận hành của các hệ thống mã hóa, cũng như các phương pháp bảo vệ dữ liệu trong thực tế.

Ngoài ra, môn học còn tập trung vào các xu hướng phát triển của công nghệ bảo mật hiện đại, giúp người học cập nhật các kỹ thuật mới nhất trong lĩnh vực an ninh mạng và mã hóa dữ liệu.

## 2. Tổng quan về an toàn thông tin hiện nay

### 2.1. Sự gia tăng của các cuộc tấn công mạng.

Trong những năm gần đây, các cuộc tấn công mạng đã trở thành một trong những mối đe dọa nghiêm trọng nhất đối với các hệ thống thông tin trên toàn cầu. Theo báo cáo của các tổ chức an ninh mạng, số lượng các vụ tấn công mạng đã tăng mạnh cả về quy mô lẫn mức độ tinh vi. Các loại tấn công phổ biến bao gồm:

* Tấn công DDoS (Distributed Denial of Service): Tin tặc sử dụng hàng ngàn, thậm chí hàng triệu thiết bị để gửi yêu cầu đến một máy chủ, làm quá tải hệ thống và khiến nó không thể hoạt động.
* Tấn công phần mềm độc hại (Malware): Các phần mềm độc hại như virus, trojan, ransomware có thể xâm nhập vào hệ thống và đánh cắp, mã hóa hoặc phá hủy dữ liệu quan trọng.
* Tấn công lừa đảo (Phishing): Kẻ tấn công giả mạo các tổ chức uy tín để lừa người dùng cung cấp thông tin cá nhân như tài khoản ngân hàng, mật khẩu, hoặc thông tin đăng nhập hệ thống.
* Khai thác lỗ hổng bảo mật (Exploiting Vulnerabilities): Tin tặc tìm kiếm và khai thác các lỗ hổng trong phần mềm hoặc hệ điều hành để xâm nhập vào hệ thống và đánh cắp dữ liệu.

Sự phát triển của công nghệ trí tuệ nhân tạo (AI) và điện toán đám mây đã giúp cải thiện đáng kể khả năng phòng thủ trước các cuộc tấn công mạng, nhưng đồng thời cũng tạo ra những phương thức tấn công mới với mức độ tinh vi cao hơn.

### 2.2. Các phương pháp bảo vệ dữ liệu quan trọng.

Để bảo vệ thông tin trước các nguy cơ tấn công mạng, nhiều phương pháp và kỹ thuật bảo mật đã được phát triển. Dưới đây là một số phương pháp quan trọng được áp dụng rộng rãi:

* Mã hóa dữ liệu (Encryption): Dữ liệu được chuyển đổi sang định dạng không thể đọc được nếu không có khóa giải mã phù hợp. Các thuật toán mã hóa phổ biến như AES, RSA, và ECC đóng vai trò quan trọng trong bảo vệ thông tin.
* Xác thực và quản lý danh tính (Authentication & Identity Management): Hệ thống xác thực đa yếu tố (MFA), sinh trắc học (vân tay, nhận diện khuôn mặt), và chứng chỉ số giúp kiểm soát quyền truy cập và ngăn chặn truy cập trái phép.
* Tường lửa (Firewall) và hệ thống phát hiện xâm nhập (IDS/IPS): Các công cụ này giúp giám sát và ngăn chặn các kết nối đáng ngờ, đảm bảo rằng hệ thống mạng không bị xâm nhập trái phép.
* Sao lưu dữ liệu (Data Backup): Thực hiện sao lưu định kỳ giúp khôi phục dữ liệu trong trường hợp hệ thống bị tấn công bởi ransomware hoặc lỗi kỹ thuật.
* Đào tạo và nâng cao nhận thức về bảo mật: Một trong những nguyên nhân chính dẫn đến các vụ tấn công thành công là do sự bất cẩn của con người. Việc đào tạo nhân viên về các kỹ thuật tấn công mạng phổ biến và cách phòng tránh là một biện pháp quan trọng.

Những phương pháp trên không chỉ giúp bảo vệ dữ liệu trước các mối đe dọa, mà còn đảm bảo tính toàn vẹn và tính sẵn sàng của hệ thống thông tin.

### 2.3. Vai trò của mật mã trong bảo mật dữ liệu.

Mật mã học (Cryptography) đóng vai trò cốt lõi trong bảo vệ thông tin, giúp ngăn chặn truy cập trái phép và bảo đảm rằng dữ liệu chỉ có thể được đọc bởi những người có quyền. Mật mã được ứng dụng trong nhiều lĩnh vực quan trọng như:

* Bảo mật giao dịch tài chính: Hệ thống ngân hàng trực tuyến sử dụng mã hóa để bảo vệ thông tin khách hàng và ngăn chặn gian lận trong giao dịch.
* Bảo vệ dữ liệu cá nhân: Các ứng dụng nhắn tin như WhatsApp, Signal sử dụng mã hóa đầu cuối (end-to-end encryption) để bảo vệ nội dung tin nhắn khỏi sự theo dõi của bên thứ ba.
* Xác thực danh tính và bảo vệ quyền truy cập: Các hệ thống đăng nhập hiện đại sử dụng mật mã để mã hóa thông tin đăng nhập, ngăn chặn tấn công dò mật khẩu (brute force attack).

Mật mã không chỉ giúp bảo vệ dữ liệu trong quá trình truyền tải mà còn đảm bảo tính toàn vẹn của dữ liệu khi lưu trữ.

### 2.4.Xu hướng phát triển của các thuật toán mật mã và bảo mật hiện đại.

Trước sự gia tăng của các cuộc tấn công mạng và sự phát triển nhanh chóng của công nghệ, các thuật toán mật mã và cơ chế bảo mật liên tục được nâng cấp để đảm bảo an toàn cho dữ liệu. Một số xu hướng nổi bật trong lĩnh vực bảo mật thông tin bao gồm:

* Mật mã hậu lượng tử (Post-Quantum Cryptography): Khi máy tính lượng tử phát triển, các thuật toán mã hóa truyền thống như RSA và ECC có thể bị phá vỡ. Do đó, các nhà nghiên cứu đang phát triển các thuật toán mật mã kháng lượng tử để đảm bảo tính bảo mật trong tương lai.
* Xác thực không mật khẩu (Passwordless Authentication): Công nghệ như sinh trắc học (vân tay, nhận diện khuôn mặt) và khóa bảo mật FIDO2 đang dần thay thế các phương pháp xác thực truyền thống.
* Trí tuệ nhân tạo trong an ninh mạng: AI và Machine Learning được sử dụng để phát hiện và ngăn chặn các mối đe dọa mạng trong thời gian thực.
* Bảo mật dữ liệu trên đám mây: Các kỹ thuật như mã hóa homomorphic và chia sẻ bí mật (Shamir's Secret Sharing) giúp tăng cường bảo mật cho các dịch vụ đám mây.

Những xu hướng này sẽ đóng vai trò quan trọng trong việc đảm bảo an toàn thông tin trong thời đại số hóa và IoT.

# CHƯƠNG 2

# TÌM HIỂU VỀ CÁC HÀM BĂM

## 1. Giới thiệu chung về hàm băm

### 1.1. Hàm băm là gì?

Hàm băm (Hash Function) là một công cụ cơ bản trong an toàn thông tin. Hàm băm là một thuật toán nhận đầu vào có kích thước bất kỳ và xuất ra một “mã băm” có độ dài cố định, khó dự đoán trước. Ý tưởng cơ bản là mã băm đóng vai trò như một hình ảnh đại diện ngắn gọn cho một chuỗi đầu vào và có thể được sử dụng như thể nó có thể nhận diện duy nhất với chuỗi đó. Nghĩa là, đầu ra của hàm băm đóng vai trò như một dấu vân tay số cho đầu vào và nên giống nhau mỗi khi cùng một thông điệp được băm. Chúng ta sử dụng các hàm băm để giúp cung cấp tính toàn vẹn dữ liệu trong mã xác thực thông điệp (MACs), để tạo ra các bản tóm tắt thông điệp cho việc sử dụng với các sơ đồ chữ ký số và để tạo ra mã phát hiện can thiệp (MDCs) trong xác thực thể và các sơ đồ thiết lập khóa [1].

### 1.2. Các loại hàm băm

Hàm băm (Hash Function) có nhiều loại khác nhau, mỗi loại phù hợp với các mục đích sử dụng cụ thể trong bảo mật, mã hóa dữ liệu, và blockchain. Trong bài báo cáo sẽ tìm hiểu MD5 (Message Digest Algorithm 5), SHA-256 (Secure Hash Algorithm) và Argon2

### 1.3. Tiêu chí đánh giá hàm băm tốt

Sự phát triển nhanh chóng của các công nghệ kỹ thuật số mới đã làm tăng nhu cầu về bảo mật thông tin trong giao tiếp. Các hàm băm đang được sử dụng rộng rãi trong các ứng dụng và giao thức bảo mật khác nhau như chữ ký số, mã xác thực tin nhắn, SSL, TLS, v.v. để đảm bảo tính toàn vẹn và tính xác thực của thông tin. Các hàm băm mật mã là các hàm nén một tin nhắn đầu vào có độ dài tùy ý thành một đầu ra có độ dài cố định ngắn , giá trị băm, khả năng chống va chạm, khả năng chống tiền ảnh và khả năng chống tiền ảnh thứ hai là ba thuộc tính bảo mật quan trọng của một hàm băm [2].

## 2. Hàm băm MD5

### 2.1. Giới thiệu

#### 2.1.1. MD5 (Message Digest Algorithm 5) là gì?

MD5 (Message-Digest algorithm 5) là một hàm băm để mã hóa với giá trị băm là 128bit. Từng được xem là một chuẩn trên Internet, MD5 đã được sữ dụng rông rải trong các chương trình an ninh mạng, và cũng thường được dùng để kiểm tra tính nguyên vẹn của tập tin.

Hàm băm MD5 tạo ra bản tóm tắt tin nhắn 128 bit từ một tin nhắn đầu vào có độ dài bit hữu hạn, tùy ý. Thuật toán xử lý các khối 512 bit được chia thành các từ 32 bit. Nó bao gồm 64 vòng. Nó là phần mở rộng của thuật toán MD4, nhanh hơn một chút và chỉ bao gồm 48 vòng, nhưng cũng được chứng minh là kém an toàn hơn. Thuật toán MD5 hoạt động như sau. [3]

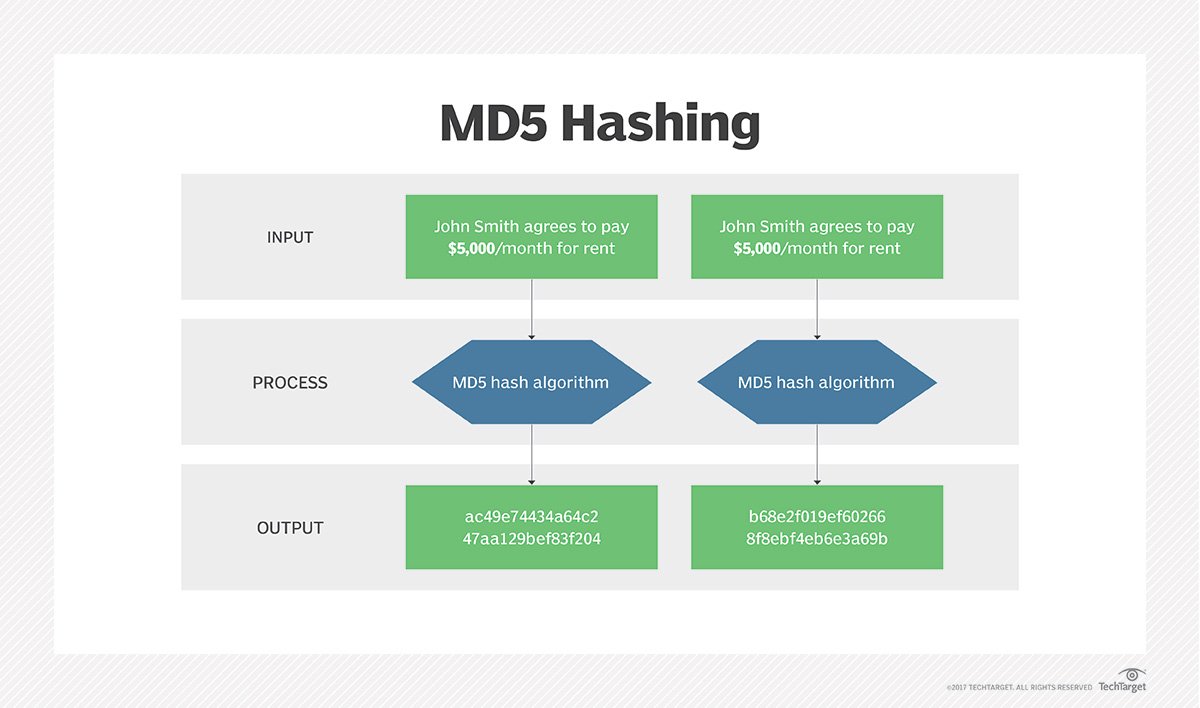
128 bits

MD5

Text 512 bit

Hình 2.1 Sơ đồ tổng quát MD5

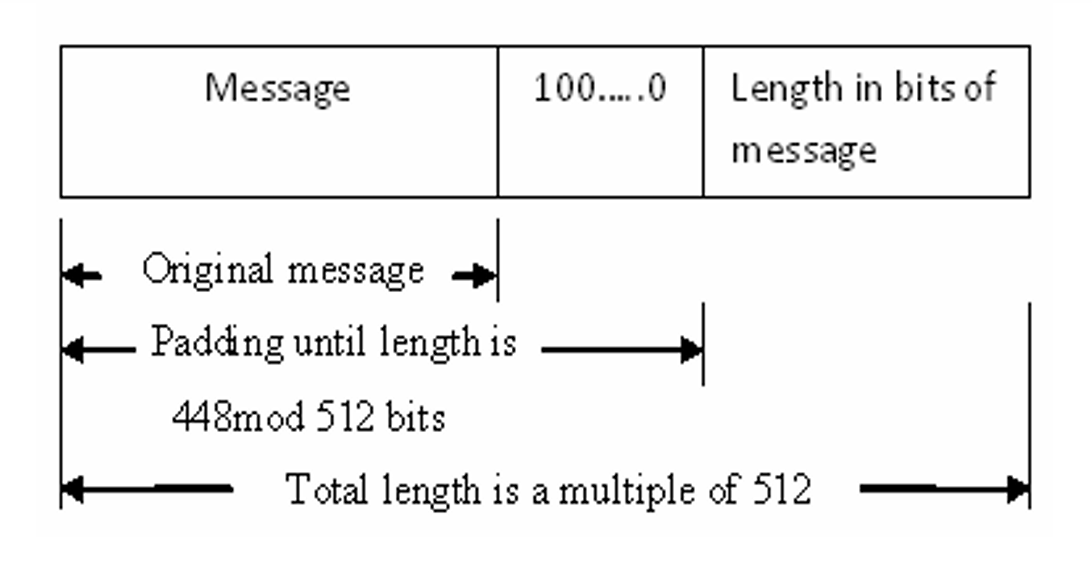
Thuật toán MD5 được thiết kế khá nhanh trên các máy 32-bit [4]. Ngoài ra, thuật toán MD5 không yêu cầu bất kỳ bảng thay thế lớn nào; Thuật toán có thể được mã hóa khá nhỏ gọn. Thuật toán MD5 là một phần mở rộng của thuật toán thông báo MD4 MD5 chậm hơn một chút so với MD4, nhưng có thiết kế "bảo thủ" hơn.



Hình 2.2 Sơ đồ minh họa kết quả đầu vào đầu ra của giải thuật MD5

#### 2.1.2. Lịch sử phát triển

MD5 được phát triển bởi Giáo sư Ronald Rivest vào năm 1991 [5], như một phần của họ các thuật toán băm MD (Message Digest). MD5 được thiết kế để khắc phục những hạn chế của MD4, đồng thời cải thiện hiệu suất và tính bảo mật. Tuy nhiên, các nhà nghiên cứu đã phát hiện ra nhiều lỗ hổng nghiêm trọng trong MD5, khiến nó dần bị thay thế bởi các thuật toán băm an toàn hơn như SHA-256 và SHA-3.



Hình 2.3 Độ dài của tin nhắn sau khi đệm (tính bằng bit)

Vào năm 1996, người ta phát hiện lỗ hổng trong MD5, trong khi vẫn chưa biết nó có phải là lỗi hổng nghiêm trọng hay không. Những chuyên gia đã nghĩa đến những giải thuật khác, như SHA-1.

Năm 2004 nhiều lỗi hổng bị khám phá nhiều hơn khiến cho việc bảo mật của giải thuật này bị đặt nghi vấn.

#### 2.1.3. Hạn chế và lỗ hỏng của MD5

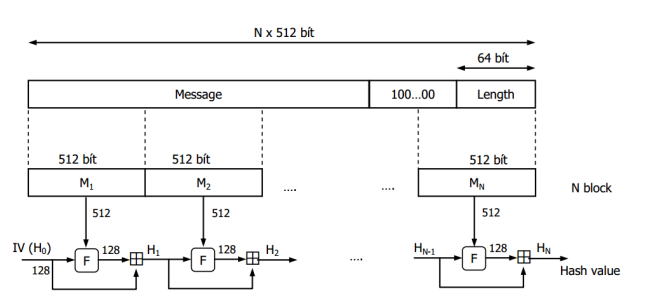
MD5 là giải thuật đời đầu nên còn nhiều lỗ hỏng trong bảo mật:

* Dễ bị tấn công va chạm (Collision Atack), hàm băm lý tưởng phải đảm bảo rằng hai đầu vào khác nhau sẽ cho giá trị băm khác nhau nhưng có cùng một giá trị băm [5]. Năm 2004, Các nhà nghiên cứu đã công bố phương pháp tạo ra va chạm cho MD5 chỉ trong vài giờ. Năm 2012, Google chứng minh rằng MD5 có thể bị lợi dụng để tạo ra chứng chỉ giả mạo.
* Dễ bị tấn công dò ngược (Preimage Attack), nếu một hàm băm mạnh sẽ rất khó để tìm ra đầu vào ban đầu từ giá trị băm. Nhưng MD5 có độ dài là 128 bit, kiến nó dễ bị tấn công brute-force. Với phần cứng hiện đại, kẻ tấn công có thể thử nhiều đầu vào để tìm ra giá trị ban đầu nhanh chóng.
* Không an toàn trước tấn công mở rộng độ dài bằng kỹ thuật length extension, cho phép hacker mở rộng dữ liệu băm mà không cần biết dữ liệu gốc.
* Tốc độ tính toán nhanh khiến dễ bị brute-force hoặc dictionary attack [5] (dùng danh sách mật khẩu phổ biến để thử). Hiện nay có cơ sở dữ liệu “rainbow table” chứa hành triệu hash MD5 của các mật khẩu phổ biến, giúp hacker dễ dàng dò ra mật khẩu.
* Không phù hợp lưu trữ mật khẩu, MD5 không có **cơ chế salt** (thêm dữ liệu ngẫu nhiên trước khi băm), nên nếu hai người dùng có cùng mật khẩu thì giá trị băm của họ cũng giống nhau.

### 2.2. Nguyên lý thuật toán và các hoạt động MD5

Giải thuật MD5 chính hoạt động trên trang thái 128-bit, được chia thành 4 từ mã 32-bit, với ký hiệu A, B, C và D. Chúng được khởi tạo với các hằng số cố định. Giải thuật chính sau đó sẽ xử lý các khối tin 512-bit, mỗi khối xác định một trạng thái. Quán trình xử lý khối tin bao gồm bốn giai đoạn giống nhau, gọi là vòng, mỗi vòng có 16 tác vụ giống nhau dự trên hàm phi tuyến F, cộng module và dịch trái.

Input: xâu đầu vào x có độ dài tối đa 2^64 -Output: chuỗi băm 128 bit -Sơ đồ thuật toán:

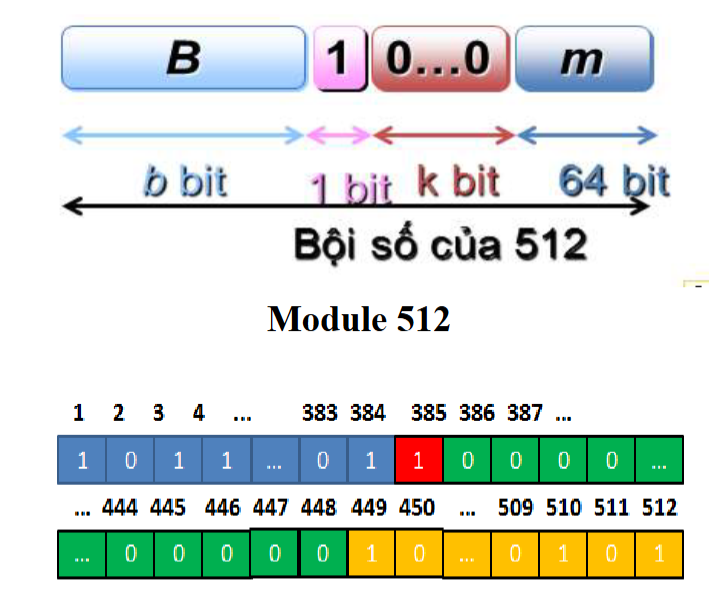


Hình 2.4. Sơ đồ hoạt động tổng quá cho quá trình băm MD5

Thuật toán MD5 có các bước thự hiện như sau [4] [2] [6]:

* *Bước 1*: Thêm các bit đệm vào chuỗi.

Thông điệp được đệm sao cho độ dài của nó đồng dạng với 448, modulo 512. Một bit "1" duy nhất được thêm vào thông điệp, và sau đó bit "0" được thêm vào sao cho độ dài tính bằng bit bằng 448 modulo 512 [4].



Hình 2.5 Minh học cho quá trình thực hiện bước 1 thêm bit đệm của MD5

- *Bước 2*: Thêm chiều dài

Biểu diễn 64 bit về độ dài của tin nhắn được thêm vào kết quả. Giai đoạn này để làm cho độ dài tin nhắn trở thành bội số chính xác của độ dài 512 bit.

- *Bước 3*: Phân chia thông điệp

MD5 xử lý chuỗi đầu vào trong các khối 512 bit, được chia thành 16 khối phụ 32 bit. Đầu ra của thuật toán được tập hợp gồm bốn khối 32 bit, nối với nhau để tạo thành giá trị băm 128 bit duy nhất.

- *Bước 4*: Khởi tạo bộ đệm MD

Bộ đệm bốn từ (A, B, C, D) được sử dụng để tính toán bản tóm tắt tin nhắn. Ở đây, mỗi từ A, B, C, D là một thanh ghi 32 bit. Các thanh ghi này được khởi tạo thành các giá trị sau theo hệ thập lục phân (hexadecimal) [4]:

+ Từ A: 01 23 45 67

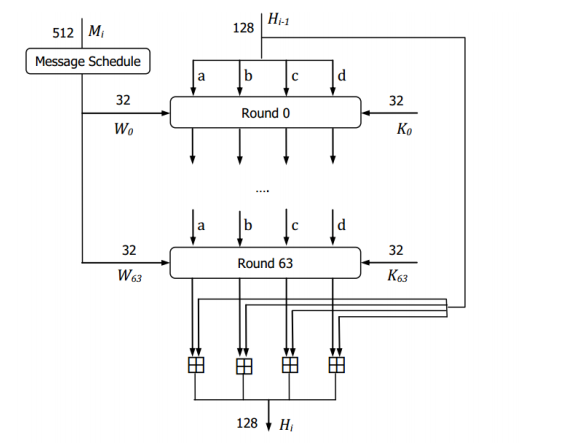
+ Từ B: 89 ab cd ef

+ Từ C: fe dc ba 98

+ Từ D: 76 54 32 10

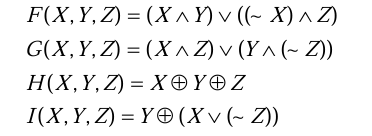
- *Bước 5:* Xử lý tin nhắn trong 16-blocks

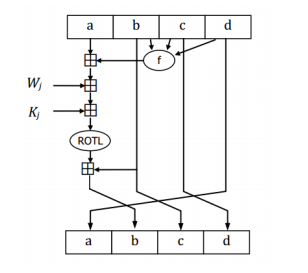
Vòng lặp chính của thuật toán bắt đầu và tiếp tục cho nhiều khối 512 bit như trong thông báo. Bốn khối được sao chép vào các biến khác nhau: a nhận A, b nhận B, c nhận C và d nhận D. Vòng lặp chính có bốn vòng, tất cả đều rất giống nhau. Mỗi chuỗi sử dụng một phép toán khác nhau 16 lần [6]. Mỗi phép toán thực hiện một hàm phi tuyến tính trên ba trong số a, b, c và d. Sau đó, nó thêm kết quả đó vào bên phải một số bit biến đổi và thêm kết quả vào một trong số a, b, c và d. Cuối cùng, kết quả thay thế một trong số a, b, c và d [2].



Hình 2.6 Cấu trúc của hàm F trong quy trinh hoạt động ở hình 2.4

Bốn hàm phi tuyến [6]:





Hình 2.7 Sơ đồ quy trình hoạt động của MD5

- Bước 6: Đầu ra

Thông báo thông báo được tạo ra dưới dạng đầu ra là A, B, C, D. Nghĩa là, đầu ra bắt đầu bằng byte bậc thấp của A và kết thúc bằng byte bậc cao của D.

### 2.3. Các ứng dụng sử dụng giải thuật MD5

Mặc dù không còn phù hợp cho các ứng dụng bảo mật, thuật toán MD5 vẫn có giá trị trong các ứng dụng không yêu cầu tính bảo mật cao, như **kiểm tra tính toàn vẹn của dữ liệu, quản lý bộ nhớ đệm và xác minh tệp tin trong các hệ thống lưu trữ đám mây.** Tuy nhiên, để đảm bảo tính bảo mật trong các hệ thống hiện đại, các thuật toán băm mạnh hơn như SHA-256, bcrypt hoặc Argon2 nên được ưu tiên sử dụng thay thế cho MD5.

#### 2.3.1. Kiểm tra tính toàn vẹn dữ liệu

Một trong những ứng dụng phổ biến nhất của MD5 hiện nay là kiểm tra tính toàn vẹn của dữ liệu. Khi một tệp tin được truyền tải qua mạng hoặc tải xuống từ một nguồn bên ngoài, nó có thể bị hỏng hoặc bị thay đổi ngoài ý muốn. Do đó, MD5 thường được sử dụng để tạo một mã kiểm tra (checksum) nhằm xác minh rằng dữ liệu nhận được có giống với dữ liệu gốc hay không.

Trong thực tế, nhiều trang web cung cấp file tải xuống thường đi kèm với mã hash MD5 để người dùng kiểm tra. Người dùng có thể tính toán giá trị MD5 của tệp tải về và so sánh với giá trị do nhà cung cấp cung cấp. Nếu hai giá trị này trùng khớp, điều đó chứng tỏ file không bị sửa đổi trong quá trình tải xuống.

Ngoài ra, trong các hệ thống truyền tải dữ liệu lớn, MD5 cũng được sử dụng để kiểm tra xem dữ liệu có bị lỗi trong quá trình truyền hay không. Một số hệ thống sao lưu dữ liệu cũng sử dụng MD5 để đảm bảo rằng các bản sao lưu không bị thay đổi hoặc mất mát dữ liệu.

#### 2.3.2. Quản lý bộ nhớ đệm và chỉ mục dữ liệu

Mặc dù MD5 không còn phù hợp để bảo vệ mật khẩu trong các hệ thống hiện đại do các lỗ hổng bảo mật, nhưng nó vẫn được sử dụng trong một số ứng dụng liên quan đến lưu trữ và xác thực dữ liệu.

Trong một số hệ thống quản lý dữ liệu, MD5 được sử dụng để tạo mã định danh duy nhất cho các tệp tin hoặc bản ghi trong cơ sở dữ liệu. Ví dụ, một hệ thống quản lý tài liệu có thể sử dụng MD5 để tạo mã hash cho mỗi tài liệu, giúp dễ dàng truy xuất và kiểm tra tính duy nhất của dữ liệu.

Bên cạnh đó, MD5 cũng được sử dụng trong bộ nhớ đệm (cache) của các hệ thống web. Các hệ thống này thường tạo một mã hash MD5 từ nội dung của trang web hoặc dữ liệu API, sau đó sử dụng mã này để kiểm soát việc lưu trữ dữ liệu trong bộ nhớ đệm, giúp tăng tốc độ truy xuất dữ liệu.

#### 2.3.3. Xác định các tệp tin trong các hệ thống lưu trữ đám mây

Các dịch vụ lưu trữ đám mây như Google Drive, Dropbox và AWS S3 sử dụng MD5 để xác minh tính toàn vẹn của tệp tin được tải lên và lưu trữ. Khi người dùng tải một tệp lên, hệ thống tính toán giá trị MD5 của tệp và so sánh với mã MD5 được lưu trữ trước đó để phát hiện sự thay đổi. Điều này giúp đảm bảo rằng tệp tin không bị hỏng hoặc thay đổi trong quá trình truyền tải.

Ngoài ra, MD5 còn hỗ trợ trong việc phát hiện trùng lặp tệp tin trong các hệ thống lưu trữ đám mây. Nếu một tệp tin có cùng giá trị MD5 với tệp đã tồn tại, hệ thống có thể tránh lưu trữ tệp tin đó hai lần, giúp tối ưu hóa dung lượng lưu trữ.

## 3. Hàm băm SHA-256

### 3.1. Giới thiệu

Thuật toán băm SHA-256 xử lý trên dữ liệu có độ rộng giới hạn trong 264 bit, bản tóm tắt được tạo ra từ SHA-256 luôn có kích thước gói gọn trong 256 bit, độ bảo mật cao phù hợp ứng dụng trong nhiều hệ thống. [7]

### 3.2. Nguyên lý thuật toán và các hoạt động SHA-256

Hoạt động của thuật toán băm bao gồm hai bước: tiền xử lý và tính toán. Bước tiền xử lý tiếp nhận dữ liệu vào và đệm, sao cho độ dài của dữ liệu được đệm là bội số của 512. Giả sử một tin nhắn M có độ dài l bit, quá trình tiền xử lý sẽ đệm bit “1” vào cuối M cùng với k bit “0” tiếp sau đó. Theo tiêu chuẩn được đề ra bởi [8], giá trị k là nghiệm nhỏ nhất của phương trình (1).

l + 1 + k ≡ 448 mod 512 (1)

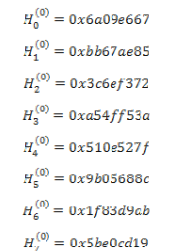
Sau khi đệm k bit “0”, tiền xử lý sẽ thêm 64 bit độ dài l tính theo hệ nhị phân vào cuối để M đạt độ dài là bội số của 512. Mỗi đoạn 512-bit của M được xem là một block, ngõ ra của bước tiền xử lý là các block 512-bit sau khi đệm: M(1), M(2), M(3) ,..., M(i) .

Ở bước tính toán, các block M(1), M(2), M(3) ,..., M(i) được đưa vào lần lượt để xử lý. Mỗi block 512- bit M(i) sẽ được phân tách thành 16 word 32-bit M1 (i) , M2 (i) ,..., M15 (i), đồng thời mở rộng số lượng 32-bit word này lên 64 word theo công thức sau: Wt = Mt (i) (0 ≤ t ≤ 15) (2) Wt = 𝜎1 256 (Wt-2) + Wt-7 + 𝜎0 256 (Wt-15) + Wt-16 (16 ≤ t ≤ 63) (3) Theo đó, hàm 𝜎(𝑥) là một hàm logic được quy định trong [1] như sau:

𝜎0 256(𝑥) = ROTR7 (x) ⊕ ROTR18 (x) ⊕ SHR3 (x) (4)

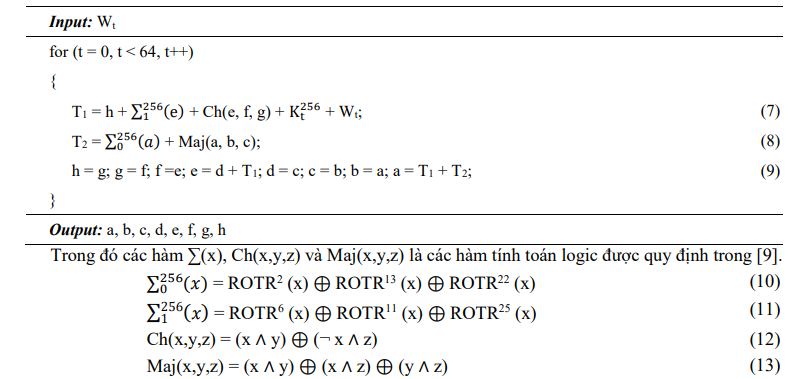
𝜎1 256(𝑥) = ROTR17 (x) ⊕ ROTR19 (x) ⊕ SHR10 (x) (5)

Các word Wt này sẽ được đưa vào vòng lặp để tính toán ra các giá trị trung gian a, b, c, d, e, f, g, h. Các giá trị này sẽ lưu giá trị băm cần để tính toán cho block kế tiếp. Đối với block đầu tiên, các giá trị trung gian sẽ được gán bằng giá trị băm khởi tạo H(0)0, H(0)1,..., H(0)7. Trong SHA-256, các giá trị băm khởi tạo được chọn bằng cách lấy 32 bit đầu của phần thập phân căn bậc 2 tám số nguyên tố đầu tiên [8].



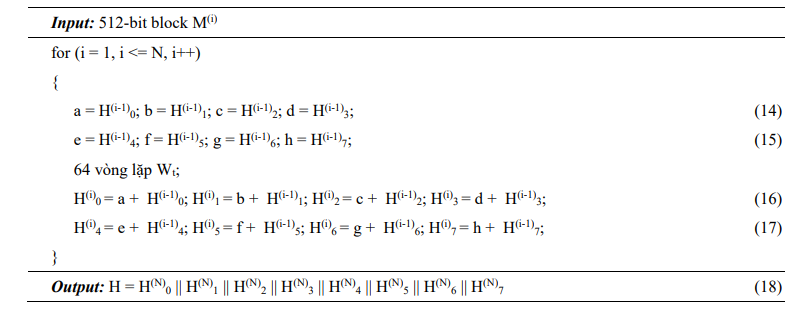
Hình 2.8. Các giá trị băm khởi tạo được chọn bằng cách lấy 32 bit

Có 64 vòng lặp tương ứng với việc tính toán trên 64 word của mỗi block. Mỗi vòng lặp thực thi một loạt các phép tính toán phức tạp trong bảng 1.



Hình 2.9. Thuật toán vòng lặp xử lý trên từng word

Như vậy, với N block 512-bit sau khi đệm, bước tính toán sẽ làm các công việc trong bảng 2 để đưa ra giá trị băm cuối cùng.



Hình 2.10. Thuật toán vòng lặp xử lý trên từng block

### 3.3. Các ứng dụng sử dụng giải thuật SHA-256 hiện tại

#### 3.3.1. Blockchain và tiền mã hóa

Một hệ thống tiền điện tử phi tập trung, hoạt động trên nền tảng blockchain để cho phép các giao dịch trực tiếp giữa người dùng mà không cần trung gian như ngân hàng. Bitcoin sử dụng cơ chế "Proof of Work" (bằng chứng công việc) để xác thực giao dịch và đảm bảo tính toàn vẹn của hệ thống, đồng thời ngăn ngừa các cuộc tấn công như "chi tiêu gấp đôi". Quá trình khai thác (mining) giải quyết các bài toán mật mã phức tạp và tạo ra các block mới. Bitcoin bảo mật và ẩn danh các giao dịch, đồng thời mở ra một mô hình tài chính phân quyền và phi tập trung [9]

#### 3.3.2. Tăng cường bảo mật Blockchain bằng thuật toán SHA-256

SHA-256 đóng vai trò quan trọng trong nhiều ứng dụng bảo mật, đặc biệt là trong công nghệ blockchain, nơi nó được sử dụng để đảm bảo tính toàn vẹn và xác thực của các bản ghi giao dịch. Tuy nhiên, với sự phát triển của các cuộc tấn công mật mã, SHA-256 đã đối mặt với một số thách thức về bảo mật. Để giải quyết vấn đề này, một phương pháp cải tiến SHA-256 đã được đề xuất, trong đó mở rộng cơ chế tạo ra hàm băm 288 bit và giảm số vòng lặp xuống 44 thay vì 64, nhưng vẫn duy trì tính khuếch tán dữ liệu thông qua quá trình lặp phức tạp. Phương pháp này giúp đảm bảo rằng ngay cả những thay đổi nhỏ trong dữ liệu đầu vào cũng tạo ra sự thay đổi rõ rệt trong hàm băm đầu ra, bảo vệ các tính chất mật mã học của thuật toán. Hàm băm SHA288 cải thiện khả năng chống xung đột và chống tìm kiếm ngược, đồng thời có thời gian thực thi nhanh hơn so với SHA-256. Các thử nghiệm thực tế đã chứng minh tính bảo mật và hiệu suất vượt trội của nó, làm tăng cường các biện pháp bảo mật trong các ứng dụng thực tế như blockchain và tiền mã hóa [10].

## 4. Hàm băm Argon2

### 4.1. Giới thiệu

Argon2 là một hàm băm **memory-hard**, được thiết kế để mã hóa các thông điệp bí mật như mật khẩu bằng cách yêu cầu sử dụng một lượng lớn bộ nhớ trong quá trình tính toán. Điều này làm tăng chi phí tính toán cho các đối thủ tấn công, đặc biệt là những kẻ sử dụng phần cứng chuyên dụng. Argon2 có hai phiên bản chính là Argon2d và Argon2i. [11]

Argon2 được phát triển bởi Alex Biryukov, Daniel Dinu, và Dmitry Khovratovich, và đã giành chiến thắng trong cuộc thi Password Hashing Competition (PHC) năm 2015.

Mục tiêu của hàm băm Argon2 là chống lại các cuộc tấn công từ điển, brute force, và các cuộc tấn công sử dụng phần cứng chuyên dụng.

### 4.2. Nguyên lý thuật toán và các hoạt động Argon2

#### 4.2.1. Các tham số đầu vào

Hàm băm Argon2 có kiểu đầu vào là primary inputs và secondary inputs. Primary inputs là có tham số đầu vào bắt buộc phải truyền khi sử dụng hàm băm Argon2. Secondary inputs là các tham số không bắt buộc, sẽ lấy giá trị mặc định nếu không truyền vào. [11]

* Primary inputs:
* **P**: Thông điệp muốn băm (0 đến bytes).
* **S**: Giá trị salt ngẫu nhiên thêm vào mật khẩu (8 đến bytes).
* Secondary inputs:
* ***p***: Số lượng luồng (threads) song song dùng để tính toán (1 đến 64 luồng).
* **m**: Kích thước bộ nhớ để áp dụng giải thuật (0 đến KB).
* **t**: Số lần lặp (1 đến lần).
* **τ**: Độ dài của tag (kết quả băm từ 4 đến bytes).
* **v**: Số phiên bản của Argon2 (1 bytes).
* **K**: Khóa bí mật nếu muốn tăng thêm độ bảo mật (0 đến 32 bytes).
* **X**: Dữ liệu liên quan (0 đến bytes)

#### 4.2.2. Hàm băm Blake2b

Blake2b là một hàm băm chia dữ liệu đầu vào thành các khối 128 bytes và sử dụng một quá trình lặp lại 12 vòng (rounds) để trộn dữ liệu đầu vào và tạo ra kết quả cuối cùng. Mỗi vòng gồm 8 lần biến đổi, mỗi lần biến đổi sẽ thực hiện 3 phép tính chính: cộng, XOR và xoay bit để trộn dữ liệu.

Hàm băm này sẽ băm các giá trị đầu vào của giải thuật Argon2, để chuẩn bị cho bước tiếp theo.

#### 4.2.3. Hàm nén

Hàm nén G là trái tim của Argon2. Nó được xây dựng dựa trên hàm vòng (round function) của Blake2b. Hàm nén G nhận hai khối 1024 byte làm đầu vào và trả về một khối 1024 byte [11]. Quá trình tính toán của hàm nén G bao gồm các bước sau:

A diagram of a diagram

AI-generated content may be incorrect.

Hình 2.11. Hàm nén G của Argon2 [11]

1. **XOR hai khối đầu vào**: R=X⊕Y
2. **Áp dụng hàm vòng Blake2b**: Hàm vòng này được áp dụng theo hàng và cột trên ma trận 8x8 của các thanh ghi 16 byte.
3. **Kết quả cuối cùng**: output=Z⊕R, trong đó Z là kết quả của hàm vòng.

#### 4.2.4. Cấu trúc và hoạt động của Argon2

Argon2 hoạt động theo nguyên tắc extract-then-expand:

* **Extract**: Trích xuất entropy từ thông điệp P và salt S bằng cách băm chúng cùng với các tham số khác.
* **Expand**: Mở rộng entropy bằng cách điền đầy bộ nhớ và thực hiện các phép tính trên các khối bộ nhớ.

A diagram of a diagram

AI-generated content may be incorrect.

Hình 2.12. Một chu kỳ lắp đầy bộ nhớ với p lane và 4 slice của Argon2 [11]

Các bước hoạt động như sau:

1. **Khởi tạo**

Tính toán giá trị khởi tạo H0 bằng cách băm tất cả các tham số đầu vào bằng hàm Blake2b:

H0=H(d, τ, m, t, v, ⟨P⟩, P, ⟨S⟩, S, ⟨K⟩, K, ⟨X⟩, X)

Trong đó d, ⟨P⟩, ⟨S⟩, ⟨K⟩, ⟨X⟩ là độ dài của τ, P, S, K, X tương ứng.

1. **Điền đầy bộ nhớ**

Bộ nhớ được tổ chức thành một ma trận các khối (blocks) có kích thước 1024 byte, chia thành p lanes và q columns. Sau đó ma trận sẽ được làm đầy t lần, các khối B[i][j] trong ma trận khối sẽ được tính thông qua hàm nén G.

+ với **t** = 1 (lần tính toán làm đầy ma trận đầu tiên) thì hai khối đầu tiên của mỗi dòng sẽ có cách tính giống nhau:

A close up of text

AI-generated content may be incorrect.

Hình 2.13. Cách cập nhật giá trị của các khối trong ma trận ở lần chạy đầu tiên [11]

Ở hai khối đầu tiên B[i][0] và khối B[i][1] của mỗi dòng sẽ được tính qua hàm G nhận 2 tham số. Tham đầu tiên là giá trị băm H0 (bằng cách băm tất cả các tham số đầu vào bằng hàm Blake2b), tham số thứ hai là ghép một giá trị 4 bytes với một giá trị 4 bytes khác (ví dụ hàng I cột 0 là ghép 4 bytes giá trị 0 với 4 bytes giá trị I), 1016 bytes còn lại sẽ là bit 0 để đủ điều kiện 1024 bytes làm đầu vào cho hàm nén [11].

Ở các khối B[i][j] còn lại sẽ được tính qua hàm G nhận tham số thứ nhất là giá trị của khối đứng trước trong hàng, tham số thứ hai sẽ là giá trị của khối mà được xác định phụ thuộc vào phiên bản:

* **Argon2d**: Sử dụng truy cập bộ nhớ phụ thuộc dữ liệu. Chỉ số khối tham chiếu được tính toán dựa trên giá trị của khối trước đó.
* **Argon2i**: Sử dụng truy cập bộ nhớ độc lập dữ liệu. Chỉ số khối tham chiếu được tính toán bằng cách sử dụng một hàm PRNG (pseudo-random number generator).

+ Với t > 1, thì hai khối đầu tiên của mỗi hàng sẽ được tính giống như các cách tính các khối khác, khối đứng trước khối đầu tiên là khối cuối cùng.

1. **Tính toán tag (kết quả) cuối cùng**

Sau khi điền đầy bộ nhớ, tag cuối cùng được tính toán bằng cách XOR các khối cuối cùng của tất cả các lane:

​=B[0][q−1]⊕B[1][q−1]⊕⋯⊕B[p−1][q−1]

Tag cuối cùng được tạo ra bằng cách băm  bằng hàm Blake2b trả về một giá trị băm 64 byte, nhưng chỉ 32 byte đầu tiên được sử dụng. Quá trình này lặp lại cho đến khi tổng số byte đầu ra đạt được độ dài mong muốn τ (độ dài tag). [11]

#### 4.2.5. Ưu nhược điểm của 2 phiên bản

**Argon2d:**

Ưu điểm: Tốc độ nhanh hơn do sử dụng truy cập bộ nhớ phụ thuộc dữ liệu. Phù hợp cho các ứng dụng như tiền điện tử, nơi không có nguy cơ tấn công kênh bên (side-channel attacks). [11]

Nhược điểm: Dễ bị tấn công kênh bên do truy cập bộ nhớ phụ thuộc dữ liệu.

**Argon2i:**

Ưu điểm: Chống lại các cuộc tấn công kênh bên do sử dụng truy cập bộ nhớ độc lập dữ liệu, phù hợp cho băm mật khẩu và các ứng dụng yêu cầu bảo mật cao. [11]

Nhược điểm: Chậm hơn Argon2d do phải thực hiện nhiều lần truy cập bộ nhớ để đảm bảo tính độc lập dữ liệu.

### 4.3. Các ứng dụng sử dụng giải thuật Argon2 hiện tại

Hiện tại thuật toán băm Argon2 được sử dụng rỗng rãi cho các ứng dụng sau:

* **Hệ thống xác thực mật khẩu**: Argon2 thường được sử dụng trong các hệ thống xác thực mật khẩu để bảo vệ thông tin đăng nhập của người dùng. Khi người dùng đăng ký hoặc thay đổi mật khẩu, mật khẩu sẽ được mã hóa bằng Argon2 và lưu trữ trong cơ sở dữ liệu. Khi người dùng đăng nhập, hệ thống sẽ sử dụng Argon2 để so khớp mật khẩu đã nhập với mật khẩu đã được mã hóa.
* **Hệ thống quản lý danh mục mật khẩu**: Argon2 cũng được sử dụng trong các ứng dụng quản lý danh mục mật khẩu, nơi người dùng có thể lưu trữ và quản lý các mật khẩu cho các dịch vụ khác nhau. Argon2 được sử dụng để mã hóa và bảo vệ các mật khẩu này, đảm bảo rằng chỉ người dùng có quyền truy cập mới có thể xem và sử dụng chúng.
* **Tiền điện tử**: Argon2 được sử dụng trong các ứng dụng tiền điện tử để tăng chi phí tính toán cho các đối thủ sử dụng phần cứng chuyên dụng (ASIC, GPU).
* **Hệ thống mã hóa đám mây**: Argon2 cũng có thể được sử dụng trong các hệ thống mã hóa đám mây để bảo vệ dữ liệu trong quá trình truyền và lưu trữ. Khi dữ liệu được tải lên đám mây hoặc tải xuống từ đám mây, Argon2 có thể được sử dụng để mã hóa và giải mã dữ liệu, đảm bảo tính bảo mật của nó.

# CHƯƠNG 3

# ỨNG DỤNG QUẢN LÝ NHÂN SỰ SỬ DỤNG HÀM BĂM ĐỂ BẢO VỆ THÔNG TIN MẬT KHẨU VÀ THÔNG TIN NGƯỜI DÙNG

## 1.Giới thiệu về hệ thống

Hệ thống quản lý nhân sự là hệ thống cho phép người dùng quản lý nhân viên, chức vụ, phòng ban,…. cũng như là tính lương cho nhân viên.

Hệ thống được xây dựng trên ngôn ngữ C#, dữ liệu được lưu trữ ở SQL Server.

Để tăng bảo mật cho hệ thống, mật khẩu sẽ được lưu trữ trong cơ sở dữ liệu dưới dạng mã được băm bởi giải thuật SHA-256 nhằm ngừa thông tin tài khoản bị lộ khi hacker chiếm được cơ sở dữ liệu.

## 2. Giới thiệu giao diện

Giao diện đăng ký cho phép người dùng đăng ký tài khoản để có sử dụng hệ thống, giao diện đăng ký như sau:

A person holding a person's hand

AI-generated content may be incorrect.

Hình 3.1. Giao diện đăng ký của hệ thống

Giao diện đăng nhập:

A person holding a person's hand

AI-generated content may be incorrect.

Hình 3.2. Giao diện đăng nhập

Giao diện hệ thống:

A screenshot of a computer

AI-generated content may be incorrect.

Hình 3.3. Giao diện trang chủ của hệ thống

Khi một tài khoản đăng ký thành công thì mật khẩu của người dùng này sẽ được băm bằng giải thuật SHA256. Ví dụ trong chức năng quản lý tài khoản, sẽ thấy được các tài khoản trong hệ thống đã được băm:

A screenshot of a computer

AI-generated content may be incorrect.

Hình 3.4. Giao diện chức năng quản lý tài khoản

# PHẦN KẾT LUẬN

## Kết quả đạt được

Trong quá trình thực hiện đồ án, nhóm đã tìm hiểu và áp dụng thành công các thuật toán hàm băm như MD5, SHA-256 và Argon2 vào việc bảo mật thông tin. Bên cạnh việc phân tích nguyên lý hoạt động của từng thuật toán, nhóm cũng đã triển khai một ứng dụng thử nghiệm để minh họa cách sử dụng hàm băm trong việc bảo vệ mật khẩu và thông tin người dùng. Kết quả cho thấy việc áp dụng hàm băm giúp tăng cường tính bảo mật và đảm bảo tính toàn vẹn của dữ liệu.

## Hạn chế

Mặc dù đạt được những kết quả khả quan, nhưng đồ án vẫn còn một số hạn chế nhất định:

* Chưa thực hiện đánh giá hiệu suất chi tiết của từng thuật toán băm trong các trường hợp khác nhau.
* Chưa thử nghiệm trên một hệ thống thực tế với dữ liệu lớn để kiểm chứng khả năng mở rộng.
* Chưa đề cập đến các phương pháp tấn công hiện đại nhằm phá vỡ hàm băm và cách phòng tránh hiệu quả.

## Hướng phát triển

Tiến hành đánh giá hiệu suất của các thuật toán băm trên tập dữ liệu lớn nhằm so sánh tốc độ và độ an toàn.

Kết hợp các phương pháp bảo mật khác như mã hóa và xác thực hai yếu tố để nâng cao mức độ bảo vệ dữ liệu.

Giả lập cuộc tấn công vào các giải thuật mật mã băm đã tìm hiểu.

# TÀI LIỆU THAM KHẢO

|  |  |
| --- | --- |
| [1] | M. M. Aung, «On The Security of Hash Functions,» de *Department of Engineering Physics*, Mandalay Technological University, 2009. |
| [2] | S. Long, « A Comparative Analysis of the Application of Hashing Encryption Algorithms for MD5, SHA-1, and SHA-512,» ICEMCE, Changzhou University Huaide College, 2019. |
| [3] | A. Z. D. Pamula, «Hardware implementation of the MD5 algorithm.,» 9th IFAC Workshop on Programmable Devices and Embedded Systems, Silesian University of Technology, Gliwice, Poland , 2009. |
| [4] | S. M. N. K. Shweta Mishra, «Hashing Algorithm: MD5,» IJSRD - International Journal for Scientific Research & Development, Department of Computer Science & Engineering, 1Echelon Institute of Technology, Faridabad, India 2, 3 JB Knowledge Park, Faridabad, India, 2013. |
| [5] | Z. W. a. S. D. D. Mary Cindy Ah Kioon, «Security Analysis of MD5 algorithm in Password Storage,» Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation (ISCCCA-13), School of Computer and Communication Engineering,University of Science and Technology (USTB), Beijing Beijing 100083, China, 2015. |
| [6] | J. T. T. a. A. B. C. G. D Rachmawati1, « A comparative study of Message Digest 5(MD5) and SHA256 algorithm,» 2nd International Conference on Computing and Applied Informatics , Departemen Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Jl. Universitas No. 9-A, Medan 20155, Indonesia, 2017. |
| [7] | T. H. H. e. a. NGUYEN, «hiết Kế Và Đánh Giá Thuật Toán Băm Bảo Mật SHA-256 Trên Nền Tảng Phần Cứng ZynQ-702,» Journal of Technical Education Science, 2024. |
| [8] | N. I. o. S. a. T. NIST, «Secure Hash Standard (SHS),» Federal Information Processing Standards Publication, 2015. |
| [9] | S. NAKAMOTO, «. Bitcoin: A peer-to-peer electronic cash system,» 2008. |
| [10] | R. K. SALIH y A. H. KASHMAR, «Enhancing Blockchain Security by Developing the SHA256 Algorithm,» Iraqi Journal of Science, 2024. |
| [11] | Alex Biryukov, Daniel Dinu, Dmitry Khovratovich, «Argon2: new generation of memory-hard functions for password hashing and other applications,» Luxembourg, 2016. |
| [12] | A. M. S. R. P. M. Esmael V. Maliberan, «A New Approach in Expanding the Hash Size of MD5,» International Journal of Communication Networks and Information Security (IJCNIS), Graduate Programs, Technological Institute of the Philippines, Quezon City, Philippines, 2018. |
| [13] | Jean-Philippe Aumasson, Samuel Neves, Zooko Wilcox-O’Hearn, Christian Winnerlein, «BLAKE2: simpler, smaller, fast as MD5,» 2013. |