TRƯỜNG CAO ĐẲNG CÔNG THƯƠNG TP. HCM

Khoa Công Nghệ Thông Tin

**TÌM HIỂU VỀ HÀM BĂM**

****

**NHÓM:**

**TÊN THÀNH VIÊN:**

**1.NGUYỄN TIỂU ĐAN**

**2. NGUYỄN HOÀNG BẢO**

MỤC LỤC

1. Giới thiệu hàm Băm HASH1

1.1 Giới thiệu3

1.2 Tính chất cơ bản của hàm Hash3

* 1. Các hàm băm mật mã học4

1. Ứng dụng của hàm Băm Hash5
2. Thuật toán7

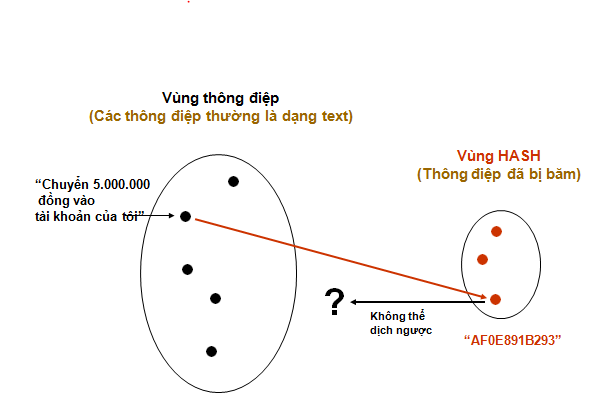
3.1 Giới thiệu hàm băm SHA-18

3.2 Thuật toán SHA-18

1. Demo chạy chương trình ứng dụng12
2. Tổng kết12
3. Tài liệu tham khảo12
4. **Giới thiệu hàm Băm HASH**
   1. **Giới thiệu:**

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 (hash function) 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.

****

* 1. **Tính chất cơ bản của hàm Hash**
* **Tính một chiều:** không thể suy ra dữ liệu ban đầu từ kết quả, điều này tương tự như việc bạn không thể chỉ dựa vào một dấu vân tay lạ mà suy ra ai là chủ của nó được.
* **Tính duy nhất**: xác suất để có một vụ va chạm (hash collision), tức là hai thông điệp khác nhau có cùng một kết quả hash là cực kỳ nhỏ
  1. **Danh sách các hàm băm mật mã học**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Thuật toán** | Kích thước  đầu ra | Kích thước trạng thái trong | Kích thước khối | Độ dài | Kích thước world | Xung đột |
| [**HAVAL**](http://vi.wikipedia.org/w/index.php?title=HAVAL&action=edit&redlink=1) | 256/224/192/160/128 | 256 | 1024 | 64 | 32 | Có |
| [**MD2**](http://vi.wikipedia.org/w/index.php?title=MD2&action=edit&redlink=1) | 128 | 384 | 128 | Không | 8 | khả năng lớn |
| [**MD4**](http://vi.wikipedia.org/w/index.php?title=MD4&action=edit&redlink=1) | 128 | 128 | 512 | 64 | 32 | Có |
| [**MD5**](http://vi.wikipedia.org/wiki/MD5) | 128 | 128 | 512 | 64 | 32 | Có |
| [**PANAMA**](http://vi.wikipedia.org/w/index.php?title=PANAMA&action=edit&redlink=1) | 256 | 8736 | 256 | No | 32 | Có lỗi |
| [**RIPEMD**](http://vi.wikipedia.org/w/index.php?title=RIPEMD&action=edit&redlink=1) | 128 | 128 | 512 | 64 | 32 | Có |
| [**RIPEMD-128/256**](http://vi.wikipedia.org/w/index.php?title=RIPEMD&action=edit&redlink=1) | 128/256 | 128/256 | 512 | 64 | 32 | Không |
| [**RIPEMD-160/320**](http://vi.wikipedia.org/w/index.php?title=RIPEMD&action=edit&redlink=1) | 160/320 | 160/320 | 512 | 64 | 32 | Không |
| [**SHA-0**](http://vi.wikipedia.org/w/index.php?title=SHA_hash_functions&action=edit&redlink=1) | 160 | 160 | 512 | 64 | 32 | Không |
| [**SHA-1**](http://vi.wikipedia.org/w/index.php?title=SHA_hash_functions&action=edit&redlink=1) | 160 | 160 | 512 | 64 | 32 | Có lỗi |
| [**SHA-256/224**](http://vi.wikipedia.org/w/index.php?title=SHA_hash_functions&action=edit&redlink=1) | 256/224 | 256 | 512 | 64 | 32 | Không |
| [**SHA-512/384**](http://vi.wikipedia.org/w/index.php?title=SHA_hash_functions&action=edit&redlink=1) | 512/384 | 512 | 1024 | 128 | 64 | Không |
| [**Tiger(2)-192/160/128**](http://vi.wikipedia.org/w/index.php?title=Tiger_(hash)&action=edit&redlink=1) | 192/160/128 | 192 | 512 | 64 | 64 | Không |
| [**VEST-4/8 (hash mode)**](http://vi.wikipedia.org/w/index.php?title=VEST&action=edit&redlink=1) | 160/256 | 256/384 | 8 | 80/128 | 1 | Không[[1]](http://vi.wikipedia.org/wiki/H%C3%A0m_b%C4%83m_m%E1%BA%ADt_m%C3%A3_h%E1%BB%8Dc#cite_note-0) |
| [**VEST-16/32 (hash mode)**](http://vi.wikipedia.org/w/index.php?title=VEST&action=edit&redlink=1) | 320/512 | 512/768 | 8 | 160/256 | 1 | Không |
| [**WHIRLPOOL**](http://vi.wikipedia.org/w/index.php?title=WHIRLPOOL&action=edit&redlink=1) | 512 | 512 | 512 | 256 | 8 | Không |

Trong đó hàm SHA-1 là một trong những hàm được sử dụng rộng rãi nhất ở Việt Nam.

1. **Ứng dụng của hàm Băm Hash**

* **Xác thực mật khẩu**

Mật khẩu thường không được lưu dưới dạng văn bản rõ (clear text), mà ở dạng tóm tắt. Để xác thực một người dùng, mật khẩu do người đó nhập vào được băm ra bằng hàm Hash và so sánh với kết quả băm được lưu trữ.

* **Xác thực thông điệp (Message authentication – Thông điệp tóm tắt -message digests)**

Giá trị đầu vào(tin nhắn, dữ liệu...) bị thay đổi tương ứng giá trị băm cũng bị thay đổi. Do vậy nếu 1 kẻ tấn công phá hoại, chỉnh sửa dữ liệu thì server có thể biết ngay lập tức.

* **Bảo vệ tính toàn vẹn của tập tin, thông điệp được gửi qua mạng**

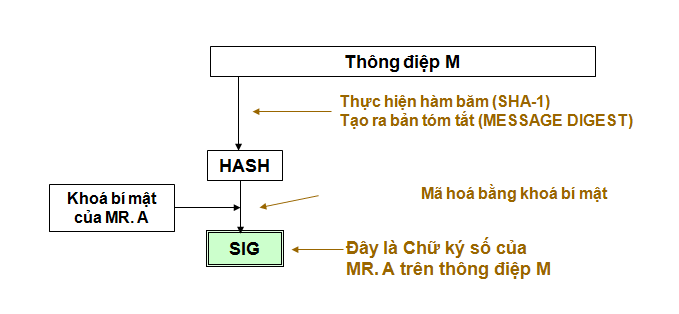
Hàm băm mật mã có tính chất là hàm 1 chiều. Từ khối dữ liệu hay giá trị đầu vào chỉ có thể đưa ra 1 giá trị băm duy nhất. Như chúng ta đã biết đối với tính chất của hàm 1 chiều. Một người nào đó dù bắt được giá trị băm họ cũng không thể suy ngược lại giá trị, đoạn tin nhắn băm khởi điểm.

Ví dụ: việc xác định xem một file hay một thông điệp có bị sửa đổi hay không có thể thực hiện bằng cách so sánh tóm tắt được tính trước và sau khi gửi (hoặc một sự kiện bất kỳ nào đó). Còn có thể dùng tóm tắt thông điệp làm một phương tiện đáng tin cậy cho việc nhận dạng file.

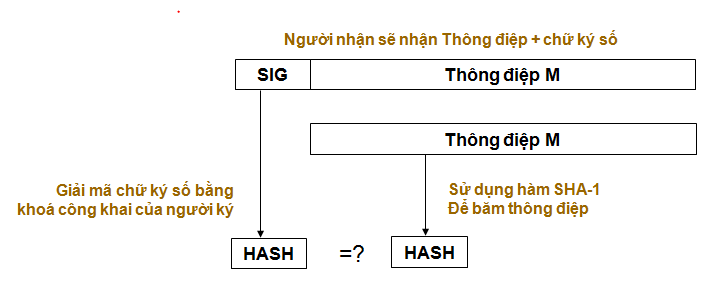
Hàm băm thường được dùng trong [bảng băm](http://vi.wikipedia.org/w/index.php?title=B%E1%BA%A3ng_b%C4%83m&action=edit&redlink=1) nhằm giảm chi phí tính toán khi tìm một khối dữ liệu trong một tập hợp. Giá trị băm đóng vai trò gần như một khóa để phân biệt các khối dữ liệu.

* **Tạo chữ ký điện tử (Digital signatures)**

**Chữ ký số có được bằng cách đem mã hoá bản tóm tắt của thông điệp bằng khoá bí mật của người ký**

****

**Chứng thực bằng chữ ký số**

****

Nếu kết quả băm giống nhau, Thông điệp được xác thực.

Tại sao?

Vì nếu bất kỳ BIT nào của M hay SIG bị thay đổi, kết quả băm sẽ khác

* **Đây là một ứng dụng cực kỳ quan trọng của hàm Hash, đặc biệt là trong thương mại điện tử.**

1. **Thuật toán Băm**

**Giới thiệu thuật toán hàm băm SHA-1**

* 1. **Giới thiệu hàm băm SHA-1**

Năm 1990, Ron Rivest đã sáng tạo ra hàm băm MD4. Sau đó năm 1992, ông cải tiến MD4 và phát triển một hàm băm khác: MD5. Năm 1993, Cơ quan An ninh Quốc gia Hoa Kỳ/Cục An ninh Trung ương (NSA) đã công bố,một hàm băm rất giống với MD5 được gọi là SHA. Vào năm 1995, sau việc khắc phục những lỗ hổng kỹ thuật, NSA đã thay đổi SHA trở thành một hàm băm mật mã khác gọi là SHA-1.

SHA-1 (Sercue Hash Algorithm) là thuật toán cũng được xây dựng trên thuật toán MD4, đang được sử dụng rộng rãi. Thuật toán SHA-1 tạo ra chuỗi mã băm có chiều dài cố định 160 bit từ chuỗi bit dữ liệu đầu vào x có chiều dài tùy ý.

* 1. **Thuật toán SHA-1**

I**nput**: thông điệp với độ dài tối đa 2^64 bits

**Output**: thông điệp rút gọn (message digest) có độ dài 160 bits

Giải thuật gồm 5 bước trên khối 512 bits

***Bước 1: Nhồi dữ liệu***

* Thông điệp được nhồi thêm các bit sao cho độ dài L mod 512 luôn đồng dư là 448.
* Thông điệp luôn luôn được nhồi thêm các bit.
* Số bit nhồi thêm phải nằm trong khoảng 1-512.
* Phần thêm vào cuối dữ liệu gồm 1 bit 1 và theo sau là các bit 0.

***Bước 2: Thêm độ dài****:*

* Độ dài khối dữ liệu ban đầu sẽ được biểu diễn dưới dạng nhị phân 64 bit và được thêm cuối chuỗi nhị phân mà ta thu được ở bước 1.
* Độ dài được biểu diễn dưới dạng nhị phân 64 bit không dấu
* Kết quả thu được từ 2 bước là một khối dữ liệu có độ dài là bội số của 512. ( Với cứ 512 bit là một khối dữ liệu)

***Bước 3: Khởi tạo bộ đệm MD ( MD buffer)***

Một bộ đệm 160 bit được dùng để lưu trữ các giá trị băm trung gian và kết quả. Bộđệm được biểu diễn bằng 5 thanh ghi 32-bit với các giá trị khởi tạo ở dạng big-endian (buyte có trọng số lớn nhất trong từ nằm ở địa chỉ thấp nhất) và có 2 bộ đệm. 5 thanh ghi của bộ đệm đầu tiên được đánh đặt tên là A, B,C,D,E và tương tự cho bộ đệm thứ 2 là. Có giá trị như sau ( Theo dạng Hex):

=67452301

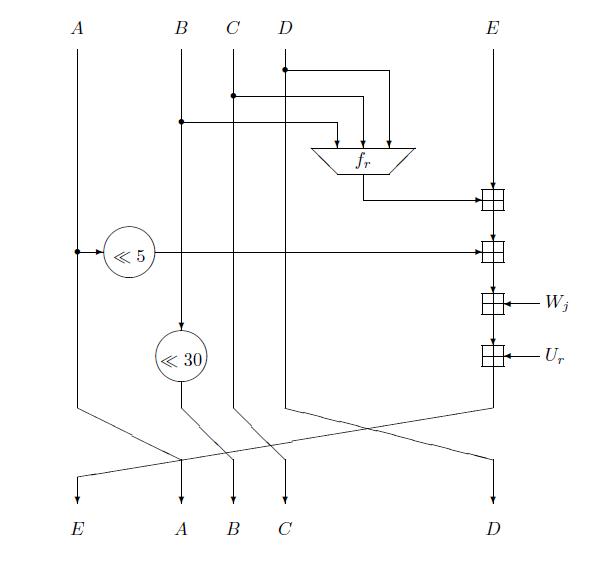
=EFCDAB89

= 98BADCFE

= 10325476

= C3D2E1F0

***Bước 4: Xử lý các khối dữ liệu 512 bit***



***Thuật toán SHA-1***

* Trọng tâm của giải thuật bao gồm 4 vòng lặp thực hiện tất cả 80 bước.
* 4 vòng lặp có cấu trúc như nhau, chỉ khác nhau ở hàm logic .

|  |  |  |
| --- | --- | --- |
| Bước | Hàm | Gía trị |
| (0≤ t ≤ 19) | =F(B, C,D) | (B AND C) OR ( (NOT B) AND D) |
| (20≤ t ≤ 39) | =F(B, C,D) | B XOR C XOR D |
| (40≤ t ≤ 59) | =F(B, C,D) | (B AND C) OR (B AND D) OR (C AND D) |
| (60≤ t ≤ 79) | =F(B, C,D) | B XOR C XOR D |

* Mỗi vòng có đầu vào gồm khối 512-bit hiện thời và một bộ điệm 160 bit A, C, B, D, E. Các thao tác sẽ cập nhật giá trị bộ đệm.
* Chia khối dữ liệu đã nhồi thêm (cuối bước 2) thành 16 nhóm (mỗi nhóm gồm 32 bit) và đặt theo thứ tự là: ,,....
* Mở rộng từ 16 nhóm 32bit lên đến 80 nhóm 32 bit bằng vòng lặp

For 16 to 79 let

= ( XOR XOR XOR )

* Gán A=, B=, C=, D=, E=.
* Mỗi vòng lặp sử dụng theo công thức chung với 1 hằng số = (0≤ t ≤ 79) như sau:

For t= 0 to 79 do

TEMP= (A)+(B,C,D)+ E + +

E=D; D=C; C= (B); B=A; A= TEMP

Với:

= 5A827999 (0≤ t ≤ 19)

= 6ED9EBA1(20≤ t ≤ 39)

= 8F1BBCDC(40≤ t ≤ 59)

= CA62C1D6(60≤ t ≤ 79) .

* Đầu ra của 4 vòng (bước 80) được cộng với giá trị của bộ đệm để tạo ra 1 chuỗi kết quả dài 160 bit.

= + A

= + E  
**Bước 5: Xuất kết quả**

* Sau khi thao tác trên toàn bộ N khối dữ liệu (blocks). Kết quả của khối thứ N là chuỗi băm 160 bit.

H=

1. **Demo chạy chương trình ứng dụng**
2. **Tổng kết**

|  |  |
| --- | --- |
| **Ưu điểm** | **Nhược điểm** |
| * Có nhiều ứng dụng quan trọng trong việc đảm bảo an toàn và bảo mật hệ thống thông tin. * Được sử dụng rộng rãi hiện nay. | * Kể từ năm 2005, các hàm mật mã của hash đã không còn đảm bảo được tính bảo mật tuyệt đối nữa. |