*Kỷ yếu Hội nghị Khoa học Quốc gia lần thứ IX "Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR'9)"; Cần Thơ, ngày 4-5/8/2016*
*DOI: 10.15625/vap.2016.00035*

# MÃ HÓA VÀ XÁC THỰC THÔNG ĐIỆP BẰNG THUẬT TOÁN MẬT MÃ VỚI KHÓA SỬ DỤNG MỘT LẦN

**Nguyễn Đức Toàn[1], Bùi Thế Hồng[2], Nguyễn Văn Tảo[3], Trần Mạnh Hường[4]**

[1] Trường Cao đẳng Công nghiệp Thực phẩm
[2] Trường Đại học Sư phạm Kỹ thuật Hưng Yên
[3] Trường Đại học Công nghệ Thông tin và Truyền thông Thái Nguyên
[4] Ban quản lý dự án VQ2 – Bộ tham mưu – Quân chủng Phòng không Không quân.

*ductoanndt@yahoo.com, hong@ioit.ac.vn, nvtao@ictu.edu.vn, quangtm102004@gmail.com*

***TÓM TẮT*** *— Trong bài báo này, chúng tôi đề xuất một thuật toán mã hóa và xác thực bản tin bằng khóa sử dụng một lần. Đây là phương pháp dựa trên sự kết hợp của phương thức xử lý mã OTP và mã khối, sử dụng hàm băm SHA256 để sinh khóa OTP ban đầu và thuật toán AES để sinh khóa OTP tiếp theo cho mỗi khối dữ liệu 256 bit. Thuật toán này sẽ tăng tốc độ mã hóa và giải mã, tăng tính bảo mật, giảm độ dài khóa bí mật, đồng thời thuật toán còn xác thực được nội dung và tác giả của bản tin và bảo mật được khóa ban đầu nhờ hệ mật khóa công khai RSA.*

***Từ khóa*** *— Hệ mật mã OTP, AES, hàm băm.*

## I. GIỚI THIỆU

Hệ mật mã với khóa sử dụng một lần OTP (One-Time Pad) là một hệ mật mã dòng, đã được chứng minh có độ an toàn hoàn hảo. Đặc điểm nổi bật của hệ mật mã này là mỗi khóa chỉ sử dụng đúng một lần và không bao giờ được dùng lại. Nhưng nhược điểm cơ bản của hệ mật mã này là độ dài của khóa phải bằng đúng độ dài của bản rõ và một yêu cầu rất khắt khe nữa là các khóa phải được sinh thực sự ngẫu nhiên. Đây là một điều kiện rất khó thực hiện ngay cả các chuỗi ngẫu nhiên được sinh tự động bằng máy tính cũng mới chỉ là giả ngẫu nhiên vì chúng phụ thuộc vào một cái nhân (seed) cho trước. Ngoài ra, độ dài của khóa cũng là một vấn đề khi bản rõ là một văn bản lớn. Vì thế đã có tác giả [x2] cho rằng OTP là phương pháp mã hóa bằng "giấy và bút chì".

Để vượt qua được những trở ngại nói trên, trong bài báo này chúng tôi đề xuất một lược đồ mã hóa với khóa sử dụng một lần cải tiến bao gồm ba quy trình sau:

- Mã hóa: Chia bản rõ thành các khối có kích thước bằng 256 bit. Nếu không chẵn thì phải chèn thêm cho đủ một khối. Băm bản rõ bằng một hàm băm an toàn với giá trị băm có kích thước bằng 256 bit. Giá trị băm này được chọn làm khóa OTP khởi đầu, gọi là $K_1$. Khóa này sẽ được sử dụng làm chữ ký số của người gửi đối với bản rõ. Sau đó, $K_1$ sẽ được XOR với khối bản rõ thứ nhất để tạo ra khối bản mã thứ nhất. Các khóa OTP tiếp theo, $K_i$ (i>=2) sẽ được sinh ra bằng cách mã hóa khối bản rõ thứ i-1 bằng hệ mật Rin256 với khóa $K_{i-1}$. Các khóa mới được sinh ra lại được XOR với khối bản rõ tương ứng để tạo ra các khối bản mã tiếp theo. Ghép tất cả các khối bản mã để thu được bản mã.

- Ký bản rõ và gửi bản mã: Khóa $K_1$ được mã hóa bằng khóa bí mật của người gửi. Sau đó, lại được mã hóa bằng khóa công khai của người nhận. Gửi cho bên nhận thông tin đã mã hóa này và bản mã.

- Xác thực và giải mã: Người nhận sử dụng khóa bí mật của mình và khóa công khai của người sử dụng để giải mã ra khóa $K_1$. Chia bản mã thành các khối có kích thước 256 bit sau đó làm tương tự như quá trình mã hóa để thu được bản rõ.

Trong lược đồ trên, chúng tôi đã lấy giá trị băm của bản rõ làm khóa OTP. Vì các bản rõ là những văn bản bất kỳ nên giá trị băm của chúng là các chuỗi bit ngẫu nhiên và gần như là duy nhất đối với mỗi bản rõ. Việc chọn giá trị băm làm khóa OTP đã làm giảm đáng kể độ dài của khóa đối với những bản rõ có dung lượng lớn. Các khóa tiếp theo được sinh ra bằng thuật toán AES trên khối bản rõ tương ứng với khóa được sinh ra trước đó vẫn đảm bảo được tính OTP.

Phần còn lại của bài báo được sắp xếp như sau. Mục II trình bày định nghĩa của một số khái niệm cần dùng trong bài báo. Mục III trình bày chi tiết lược đồ mã hóa kết hợp giữa mã khối và mã khóa sử dụng một lần. Mục IV mô tả quy trình cài đặt thử nghiệm và đánh giá kết quả.

## II. MỘT SỐ ĐỊNH NGHĨA

**Hệ mã với khóa sử dụng một lần OTP (One Time Pad)** Thuật ngữ "one-time pad" dùng để chỉ một phương pháp mã hóa trong đó mỗi byte của bản rõ được mã hóa bằng một byte của luồng khóa và mỗi byte khóa chỉ được sử dụng đúng một lần và không bao giờ được sử dụng lại. Trong hệ mã này độ dài của khóa phải bằng đúng độ dài của bản rõ và phải phải là một luồng được sinh ra thực sự ngẫu nhiên, tức là mọi byte của khóa có thể nhận bất kỳ giá trị nào trong khoảng từ 0 đến 255 với xác suất như nhau và độc lập với giá trị của tất cả các byte khóa khác. Trong hệ mã OTP,

bản rõ được biểu diễn dưới dạng một chuỗi nhị phân, luồng khóa cũng là một chuỗi nhị phân có độ độ dài bằng độ dài bản rõ.

Việc mã hóa bằng OTP thường được ký hiệu là $C_i = P_i \oplus K_i$ , (i = 1, 2, 3,...), trong đó $P_i$ là ký tự thứ i của bản rõ, $K_i$ là byte thứ i của khóa được sử dụng để mã hóa bản rõ này và $C_i$ là ký tự thứ i của bản mã kết quả, $\oplus$ là ký hiệu của phép cộng loại trừ (XOR), phép toán hay được dùng trong mã hóa OTP nhưng vẫn có thể được thay bằng phép toán khác. Quá trình giải mã được làm tương tự như mã hóa $P_i = C_i \oplus K_i$.

Tuy nhiên, phương pháp One-Time Pad không có ý nghĩa sử dụng thực tế vì chiều dài khóa bằng chiều dài bản tin, mỗi khóa chỉ sử dụng một lần, nên thay vì truyền khóa trên kênh an toàn thì có thể truyền trực tiếp bản rõ mà không cần quan tâm đến vấn đề mã hóa nữa. Bởi vậy chúng tôi đã sử dụng hàm băm để giảm độ dài của khóa bí mật.

**Hệ mã AES và Rindajen:** AES là một thuật toán mã hóa khối được chính phủ Hoa Kỳ áp dụng làm tiêu chuẩn mã hóa từ năm 2001. Tiền thân của AES là thuật toán Rijndael được thiết kế bởi J. Daemen và V. Rijmen. Hai thuật toán này vẫn thường được gọi thay thế cho nhau nhưng trên thực tế thì chúng không hoàn toàn giống nhau. AES chỉ làm việc với các khối dữ liệu (đầu vào và đầu ra) 128 bít và khóa có độ dài 128, 192 hoặc 256 bít trong khi Rijndael có thể làm việc với dữ liệu và khóa có độ dài 128, 192, 256 bít. Để cho thuận tiện, trong bài báo này chúng tôi dùng tên AES256 để chỉ thuật toán Rijndael với độ dài khối dữ liệu và khóa cùng là 256 bit.

**Mã hóa khóa công khai:** Mật mã hóa khóa công khai là một dạng mật mã hóa cho phép người sử dụng trao đổi các thông tin mật mà không cần phải trao đổi các khóa chung bí mật trước đó. Điều này được thực hiện bằng cách sử dụng một cặp khóa có quan hệ toán học với nhau là khóa công khai và khóa bí mật. Trong hai khóa, một dùng để mã hóa và khóa còn lại dùng để giải mã. Điều quan trọng đối với hệ thống là không thể tìm ra khóa bí mật nếu chỉ biết khóa công khai. Hệ thống mật mã hóa khóa công khai có thể sử dụng với các mục đích:

- Mã hóa: giữ bí mật thông tin và chỉ có người có khóa bí mật mới giải mã được.
- Tạo chữ ký số: cho phép kiểm tra một văn bản có phải đã được tạo với một khóa bí mật nào đó hay không.
- Thỏa thuận khóa: cho phép thiết lập khóa dùng để trao đổi thông tin mật giữa 2 bên.

**Chữ ký số khóa công khai:** là mô hình sử dụng các kỹ thuật mật mã để gắn với mỗi người sử dụng một cặp khóa công khai - bí mật và qua đó có thể ký các văn bản điện tử cũng như trao đổi các thông tin mật. Khóa công khai thường được phân phối thông qua chứng thực khóa công khai do một tổ chức được chính quyền tín nhiệm cấp. Quá trình sử dụng chữ ký số bao gồm 2 quá trình: tạo chữ ký và kiểm tra chữ ký. Có thể nói chữ ký số là thông tin đi kèm theo dữ liệu (văn bản, hình ảnh, video...) nhằm mục đích xác định người chủ của dữ liệu đó.

Để sử dụng chữ ký số thì văn bản cần phải được mã hóa bằng hàm băm (văn bản được "băm" ra thành chuỗi, thường có độ dài cố định và ngắn hơn văn bản) sau đó dùng khóa bí mật của người chủ khóa để mã hóa, khi đó ta được chữ ký số. Khi cần kiểm tra, bên nhận giải mã bằng khóa công khai của bên gửi để lấy lại chuỗi gốc được sinh ra qua hàm băm ban đầu, sau đó tự băm văn bản nhận được và so sánh hai chuỗi bít này với nhau. Nếu chúng trùng khớp thì bên nhận có thể tin tưởng rằng văn bản nhận được chính là văn bản xuất phát từ người sở hữu khóa bí mật.

## III. LƯỢC ĐỒ MÃ HÓA VÀ GIẢI MÃ

### A. Mô tả lược đồ

Lược đồ mã hóa và giải mã với khóa sử dụng một lần bao gồm ba quy trình được thực hiện giữa người gửi A và B. Giả thiết là A và B đều là thành viên của một hạ tầng khóa công khai nào đó. Các khóa công khai và bí mật của A được ký hiệu lần lượt là $K_{AU}$ và $K_{AR}$. Các khóa của B cũng được ký hiệu tương tự là $K_{BU}$ và $K_{BR}$. Lược đồ này bao gồm các quy trình sau:

1) **Mã hóa**: A chia bản rõ M thành các khối $M_i$ , (i=1, 2,..., n) có kích thước bằng 256 bit. Nếu không chẵn thì phải chèn thêm cho đủ một khối (cách chèn thêm giống như trong các hàm băm SHA). Băm bản rõ bằng hàm băm an toàn SHA256 với giá trị băm có kích thước bằng 256 bit. Giá trị băm này được chọn làm khóa OTP khởi đầu, gọi là K. Khóa này sẽ được sử dụng làm chữ ký số của người gửi đối với bản rõ. Sau đó, K sẽ được chọn làm khóa $K_1$ khởi đầu cho luồng khóa OTP. $K_1$ được XOR với khối bản rõ thứ nhất $M_1$ để tạo ra khối bản mã thứ nhất $C_1$ . Các khóa OTP tiếp theo, $K_i$ (i=2, 3, ..., n) sẽ được sinh ra bằng cách mã hóa khối bản rõ $M_{i-1}$ bằng hệ mật AES256 với khóa $K_{i-1}$. Các khóa mới được sinh ra lại được XOR với khối bản rõ tương ứng để tạo ra các khối bản mã tiếp theo. Ghép tất cả các khối bản mã để thu được bản mã.

| | |
|---|---|
| Bản rõ: | $M = M_1 M_2 \ldots M_n$ ; $\lvert M_i \rvert = 256$ bit, với i=1, 2,..., n |
| Khóa OTP ban đầu: | $K = K_1 = SHA256(M)$ |
| Các khóa OTP tiếp theo: | $K_i = AES256(M_{i-1}, K_{i-1})$, với i=2, 3,..., n |
| Các khối bản mã: | $C_i = M_i \oplus K_i$ , với i=1, 2,..., n |
| Bản mã: | $C = C_1 C_2 \ldots C_n$ |

2) **Ký bản rõ và truyền tin**: A ký bản rõ bằng cách mã hóa khóa OTP K (giá trị băm của M) bằng khóa bí mật $K_{AR}$. Sau đó lại mã hóa tiếp bằng khóa công khai $K_{BU}$ của B để đảm bảo chỉ B mới đọc được khóa K. A gửi cho B bản mã này và bản mã C:

$E(E(K, K_{AR}), K_{BU})$ C; trong đó E là thuật toán mã hóa khóa công khai RSA

3) **Xác thực và giải mã**: B nhận được chữ ký của A đã được mã hóa bằng khóa công khai của B và bản mã C'. B sử dụng khóa bí mật của mình và khóa công khai của A để giải mã ra một chuỗi bit tạm gọi là K'. Sau đó, B băm bản mã C'và thu được một chuỗi tạm gọi là K". So sánh K' với K". Nếu chúng trùng khớp nhau thì B khẳng định rằng A chính là người gửi tin cho mình, trong đó K ≡ K' ≡ K" là khóa OTP khởi đầu dùng để giải mã bản mã C ≡ C'. B chia bản mã C thành các khối có kích thước 256 bit sau đó làm tương tự như quá trình mã hóa của A để thu được bản rõ.

$$K' = D(D(E(E(K, K_{AR}), K_{BU}), K_{BR}), K_{AU})$$

$$K'' = SHA256(C')$$

Nếu K' ≡ K" thì thực hiện

$$C = C_1 C_2 \ldots C_n \text{ với } |C_i| = 256 \text{ bit, với } i=1, 2, \ldots, n$$

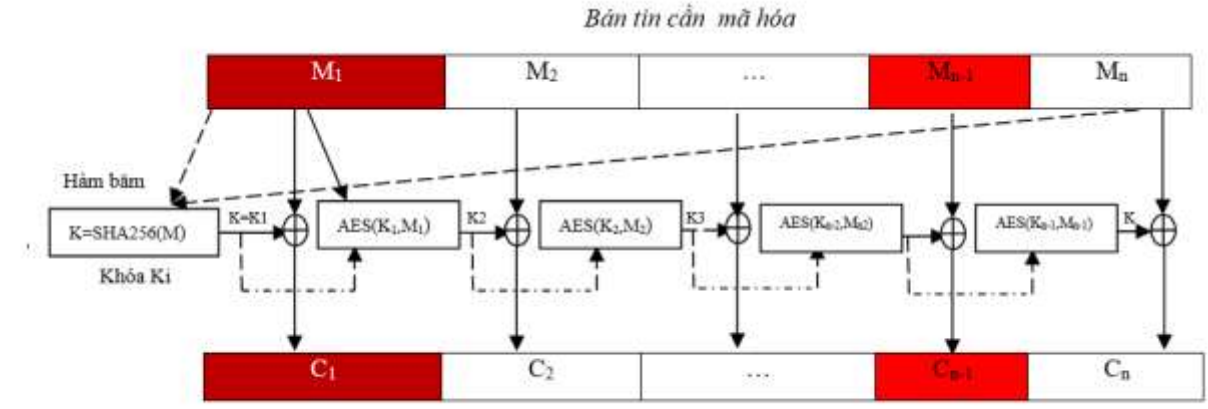
$$K_1 = K'$$

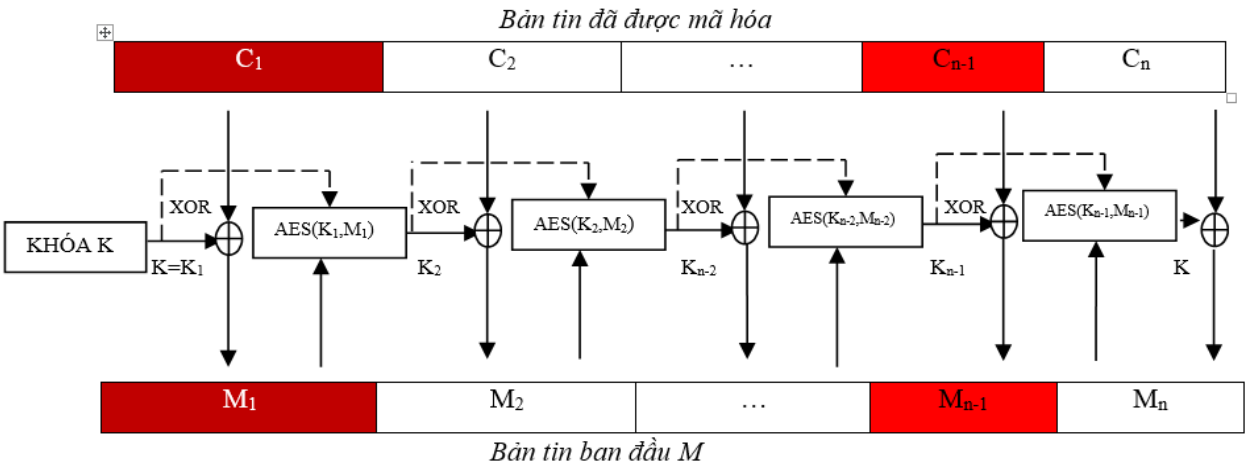
$$K_i = AES256(C_{i-1}, K_{i-1}), \text{ với } i=2, 3, \ldots, n$$

Các khối bản rõ: $M_i = C_i \oplus K_i$, với i=1, 2, ... , n

Bản rõ: $M = M_1 M_2 \ldots M_n$

Quy trình mã hóa được mô tả trong Hình 1 còn Hình 2 mô tả qui trình giải mã.




**B. Thuật toán giải mã**

***1) Mã hóa:*** **(bên A)**

- *Đầu vào: bản rõ M có độ dài tùy ý. M sẽ được chia thành các khối $M_i$ sao cho:*
  *$M = M_1 M_2 ... M_n$.*
- *Đầu ra: bản mã của M: $C = C_1 C_2 ... C_n$.*
- *Thuật toán:*

*// Sinh khóa $K_1$:*

*$K = SHA256(M)$;*

*$K_1 = K$;*

*// Mã hóa*

*$C_1 = K_1 \oplus M_1$;*

*For i=2; i<=n; i++*

*{*

*// Sinh các khóa $K_2, ..., K_n$:*

*$K_i = AES(K_{i-1}, M_{i-1})$;*

*// Mã hóa:*

*$C_i = M_i \oplus K_i$;*

*}*

*$C = C_1 C_2 ... C_n$;*

***2) Truyền tin:* (bên A)**

*// A ký thông điệp M (mã hóa K bằng khóa bí mật);*

*$E(K, K_{AR})$*

*// Mật hóa chữ ký bằng khóa công khai của B;*

*$E(E(K,K_{AR}),K_{BU})$*

*// Gửi bản mã C và bản mã chữ ký cho bên nhận;*

*send C and $E(E(K,K_{AR}),K_{BU})$ to B*

***3) Xác thực và giải mã (bên B)***

- Đầu vào: $C' = C'_1 C'_2 ... C'_n$; $E(E(K,K_{AR}),K_{BU})$
  Đầu ra: bản rõ $M = M_1 M_2 ... M_n$.
- Thuật toán:

// B giải mã chữ ký thu được khóa K'

K' = D(D(E(E(K, $K_{AR}$), $K_{BU}$), $K_{BR}$),$K_{AU}$)

// Băm bản mã C'

K" = SHA256(C')

// So sánh K' và K"

if (K'= = K") // A chính là người gửi bản mã C' = C và K=K' là khóa OTP khởi đầu

{

//Chia C thành các khối có độ dài bằng 256 bit

*$C = C_1 C_2 ... C_n$;*

*$M_1 = K_1 \oplus C_1$;*

For i=2; i<=n; i++

{

// sinh dòng khóa OTP

*$K_i = AES256(K_{i-1}, M_{i-1})$;*

*// giải mã $C_i$*

*$M_i = K_i \oplus C_i$ ;*

}

*$M = M_1 M_2 .... M_n$ ;*

}

*else print "có lỗi! đề nghị gửi lại";*

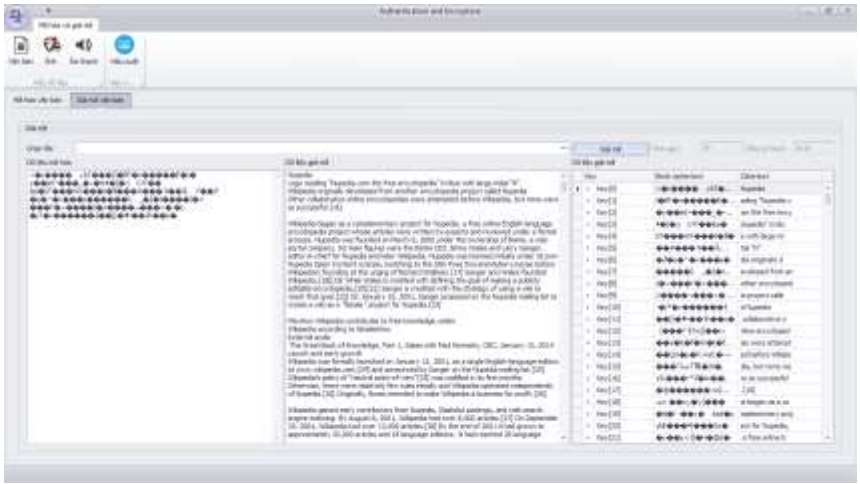
## IV. CÀI ĐẶT THỰC NGHIỆM

- Phần mã hóa văn bản:

Từ bản rõ mã hóa thành bản mã và các dữ liệu mã hóa, thời gian mã hóa và tổng số block.

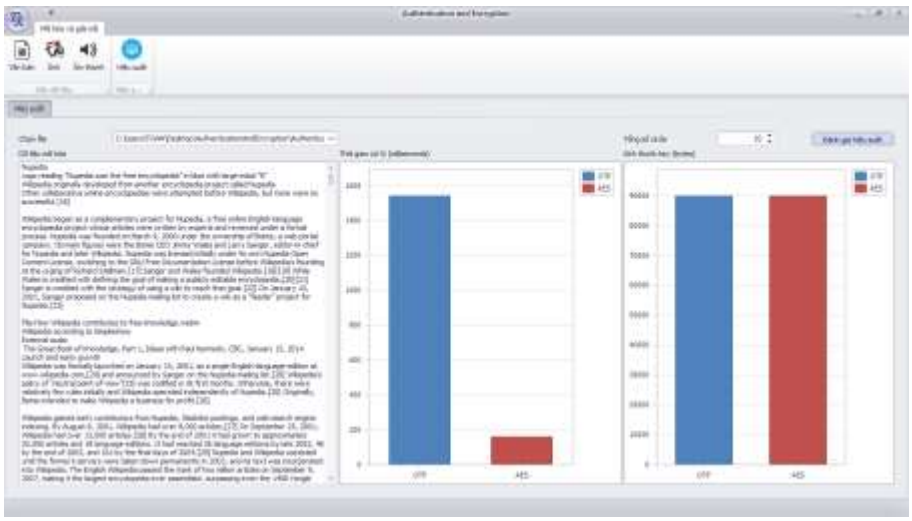
- Phần giải mã văn bản:

Từ bản mã thành bản rõ và các dữ liệu giải mã, thời gian mã hóa và tổng số block.



- Phần đánh giá và so sánh:

Với cùng dữ liệu đầu vào, với số vòng như nhau nhưng thời gian xử lý của OTP nhanh hơn hẳn so với các thuật toán khác.



## V. KẾT QUẢ THỰC NGHIỆM

Chương trình phần mềm minh họa thuật toán đã thực hiện các chức năng

- Mã hóa được thông điệp bất kỳ nhập vào, bằng thuật toán mật mã với khóa sử dụng 1 lần (OTP), thực hiện mã hóa, giải mã và hiển thị kết quả.
- So sánh với các thuật toán khác với tốc độ nhanh hơn khi có cùng thông điệp đầu vào.
- Nghiên cứu đưa ra các phương pháp sử dụng thuật toán OTP nhằm mục tiêu tăng độ an toàn, tăng tốc độ mã hóa và giải mã và giảm độ dài khóa, bổ sung thêm khả năng xác thực.

## VI. KẾT LUẬN VÀ ĐỊNH HƯỚNG NGHIÊN CỨU

Trong khuôn khổ bài báo này chúng tôi đã nghiên cứu và đưa ra một phương pháp mã hóa dựa trên sự kết hợp mật mã dòng, mật mã khối và hàm băm. Với phương pháp tương tự thuật toán có thể dễ dàng thực hiện với khối dữ liệu dài hơn, kết hợp thuật toán AES, RSA để tăng độ an toàn, xác thực được nội dung bản tin và bảo mật được khóa ban đầu.

## TÀI LIỆU THAM KHẢO

[1] http://www.pcworld.com.vn/articles/cong-nghe/cong-nghe/2006/02/1188574/khai-niem-ve-cryptography/

[2] https://vi.wikipedia.org/wiki/SHA.

[3] https://vi.wikipedia.org/wiki/Hàm_băm_mật_mã_học.

[4] Nguyễn Bình (1996), *Mật mã học: Lý thuyết và thực hành*, Cục Kỹ thuật - Viện Kỹ thuật thông tin.

[5] Dương Anh Đức, Trần Minh Triết (2005), *Mã hóa và ứng dụng*, Trường Đại học Khoa học Tự nhiên - Đại học Quốc gia TP.HCM.

[6] Richard A. Mollin, "An Introduction to Cryptography – 2nd ed", Taylor & Francis Group, LLC, 2007.

# ENCRYPTION AND AUTHENTICATION MESSAGE BY CIPHER ONE TIME PAD (OTP)

**Nguyen Duc Toan, Bui The Hong, Nguyen Van Tao, Tran Manh Huong**

***ABSTRACT*** *— In this paper, we propose an encryption algorithm and authentication with the key messages used once. This innovative method is based on a combination of treatment methods and OTP cipher code, using SHA256 hash function to the original OTP key generation and key generation algorithm AES for next OTP for each 256-bit block of data. This algorithm will speed encoding and decoding, increase security, reduce the length of the secret key, and authentication algorithm is also the author of the content and security bulletins and initial key thanks RSA public key.*

***Keywords*** *— OTP encryption system, AES, Hash.*