Đánh giá chính xác cận an toàn cho mã xác thực LightMAC

Nguyễn Tuấn Anh

***Tóm tắt*— LightMAC là mã xác thực thông điệp được Atul Luykx đề xuất sử dụng trong các môi trường có tài nguyên hạn chế và có cận an toàn không phụ thuộc vào độ dài thông điệp. Thuật toán LightMAC sinh ra nhãn xác thực có độ dài tùy theo yêu cầu của người sử dụng. Tuy nhiên, đánh giá an toàn trong [1] lại sử dụng trực tiếp kết quả dành cho độ dài nhãn xác thực bằng kích cỡ mã khối cơ sở của Dodis [2]. Trong bài báo này, đầu tiên, chúng tôi đánh giá cận an toàn của mã xác thực LightMAC trong trường hợp độ dài nhãn xác thực nhỏ hơn kích cỡ của mã khối cơ sở. Sau đó, sự phụ thuộc vào độ dài thông điệp trong cận an toàn của LightMAC được xem xét lại.**

***Abstract*— The message authentication code mode, LightMAC, which was proposed to use in resource-constrained environments by Atul Luykx has security bound independ on message length. The tag length in LightMAC algorithm depend on demand of user's. However, the security analysis's Atul [1] directly uses the Dodis's result [2] which presents for the case that tag length is the block size. In this paper, we first evaluate the security bound of LightMAC when tag length is less than the block size. Then, the dependence on the message length of LightMAC's security bound is reviewed.**

***Từ khóa*— hàm giả ngẫu nhiên; mã xác thực thông điệp; LightMAC.**

***Keywords*— pseudorandom function; message authentication code; LightMAC.**

I. GIỚI THIỆU

Các mã xác thực thông điệp thông thường như: CBC MAC, EMAC, CMAC, PMAC đều có cận an toàn phụ thuộc vào số lượng các thông điệp truy vấn và độ dài thông điệp. Cận an toàn cho các mã xác thực thông điệp này là $q^2l/2^n$ [3]; trong đó q là số truy vấn tối đa mà kẻ tấn công thực hiện, l là độ dài thông điệp theo khối, n là kích cỡ của mã khối cơ sở. Trong các môi trường xác thực thông thường, có nghĩa là mã xác thực sử dụng mã khối cơ sở có kích cỡ 128 bit ($n = 128$), và ta mong muốn rằng xác suất giả mạo của kẻ tấn công không vượt quá một phần một triệu [1], khi đó ta phải đảm bảo rằng:

$$\frac{q^2l}{2^{128}} \leq \frac{1}{2^{20}}, \text{hay } q^2l \leq 2^{108}.$$

Do đó, với mỗi khóa ta có thể xác thực được 2^{54} thông điệp, mỗi thông điệp gồm một khối. Tương tự, có những 2^{53} thông điệp, mỗi thông điệp gồm 4 khối, có thể được xác thực cho mỗi khóa. Ta quan sát thấy rằng, số lượng thông điệp được xác thực trong mỗi lần sử dụng khóa rất lớn. Điều này không gây ảnh hưởng lớn đến không gian dữ liệu được xác thực.

Tuy nhiên, trong các môi trường có tài nguyên hạn chế, tức là mã xác thực sử dụng mã khối cơ sở có kích cỡ là 32 bit hay 64 bit, thì số lượng thông điệp được xác thực đối với mỗi khóa sẽ bị giảm đi đáng kể. Thật vậy, tương tự như trên, ta xét số lượng thông điệp được xác thực cho mỗi khóa khi trong các ứng dụng dùng mã khối 32 bit ($n = 32$), và yêu cầu xác suất giả mạo của kẻ tấn công không vượt quá một phần một triệu [1]. Khi đó:

$$\frac{q^2l}{2^{32}} \leq \frac{1}{2^{20}}, \text{hay } q^2l \leq 2^{12}.$$

Từ ràng buộc trên, ta suy ra mỗi khóa chỉ có thể xác thực cho 64 thông điệp, mỗi thông điệp gồm 1 khối. Tương tự, chỉ có 32 thông điệp, mỗi thông điệp 4 khối có thể được xác thực cho mỗi khóa.

Để giải quyết được vấn đề này, năm 2015, tại hội nghị FSE, Atul Luykx và các cộng sự đã giới thiệu một mô hình xác thực thông điệp sử dụng mã khối hạng nhẹ với tên gọi là LightMAC [1] có cận an toàn không phụ thuộc vào độ dài thông điệp. Điều này cho phép LightMAC xác thực nhiều thông điệp hơn đối với mỗi khóa.

Bài báo được nhận ngày 3/10/2018. Bài báo được nhận xét bởi phản biện thứ nhất vào ngày 30/10/2018 và được chấp nhận đăng vào ngày 14/11/2018. Bài báo được nhận xét bởi phản biện thứ hai vào ngày 30/10/2018 và được chấp nhận đăng vào ngày 5/11/2018.

Các công trình liên quan. Đánh giá độ an toàn cho mã xác thực thông điệp LightMAC được Atul Luykz và các cộng sự trình bày trong [1]. Cách tiếp cận này dựa trên mô hình băm-rồi-mac của Dodis [2]. Tuy nhiên, kết quả của Dodis chỉ phát biểu cho trường hợp nhãn xác thực là toàn bộ đầu ra của hàm mã, trong khi mô hình của LightMAC phát biểu cho cả trường hợp đầu ra bị cắt ngắn. Do đó, cần phải có các đánh giá chính xác hơn cho LightMAC.

Đóng góp của chúng tôi. Trong bài báo này, chúng tôi đánh giá lại cận an toàn cho LightMAC trong trường hợp nhãn xác thực chỉ lấy $t < n$ bit đầu ra. Ngoài ra, chúng tôi cũng phân tích, so sánh mức độ phụ thuộc vào độ dài thông điệp của mã xác thực thông điệp này với các mã xác thực thông điệp trước đó.

Phần còn lại của bài báo được tổ chức gồm: Mục II trình bày các kiến thức cơ sở liên quan; Mục III sẽ đưa ra một số kết quả đã có; Cuối cùng trong Mục IV sẽ phân tích độ an toàn của LightMAC và đưa ra một số kết luận.

II. CÁC KIẾN THỨC CƠ SỞ

A. Một số ký hiệu

Ký hiệu I_n là tập các chuỗi bit có độ dài n; $I_{\leq n}$ là tập các chuỗi bit có độ dài không vượt quá n; I_* là tập các chuỗi bit có độ dài bất kỳ. $F_{n,m}$ là tập các hàm từ I_n vào I_m. Với số nguyên $1 \leq i \leq 2^s$, $\underline{i}_s$ biểu diễn cách viết lại i theo s bit. Với chuỗi M độ dài n bit, ký hiệu $\lfloor M \rfloor_t$ là t bit ít có ý nghĩa nhất của M. Ký hiệu $\xleftarrow{\$}$ là phép lấy ngẫu nhiên; trong khi $\xleftarrow{n-s}$ là phép chia thông điệp thành các khối $n-s$ bit, khối cuối nhỏ hơn hoặc bằng $n-s$ bit. Trong bài báo này ký hiệu $M10^*$ là phép đệm các bit có dạng 10…0 vào sau M sao cho $|M10^*| = n$.

B. Một số khái niệm, định nghĩa

Hàm được chọn ngẫu nhiên (tương ứng hoán vị được chọn ngẫu nhiên) ở đây được hiểu là hàm (tương ứng hoán vị) được lấy ngẫu nhiên từ $F_{n,m}$ (tương ứng $\mathrm{Perm}(n)$) phù hợp với một phân phối xác suất cố định. Hàm (hoán vị) ngẫu nhiên hoàn thiện là hàm (hoán vị) được lấy ngẫu nhiên đều từ tập $F_{n,m}$ ($\mathrm{Perm}(n)$).

Tiếp theo sẽ xem xét khái niệm lợi thế phân biệt. Theo đó, lợi thế phân biệt của một kẻ tấn công có được khi phân biệt một hàm được chọn ngẫu nhiên với một hàm ngẫu nhiên hoàn thiện. Ta viết A^g nếu như kẻ tấn công A được quyền truy cập vào bộ tiên tri là hàm g.

***Định nghĩa 1 (Definition 4.6,* [4]).** *Cho f là một hàm được chọn ngẫu nhiên. Gọi A là một kẻ tấn công phân biệt f và hàm ngẫu nhiên hoàn thiện f^*. Ta xét hai thí nghiệm sau:*

$\mathbf{Exp}_f^{\mathrm{prf}-1}(A)$	$\mathbf{Exp}_f^{\mathrm{prf}-0}(A)$
$b \leftarrow A^f$	$b \leftarrow A^{f^*}$
Trả về b	*Trả về* b

Lợi thế của một kẻ tấn công A trong việc phân biệt giữa f với một hàm ngẫu nhiên hoàn thiện là:

$$\mathrm{Adv}_f^{\mathrm{prf}}(A) := \left|\Pr\left[\mathrm{Exp}_f^{\mathrm{prf}-1}(A) = 1\right] - \Pr\left[\mathrm{Exp}_f^{\mathrm{prf}-0}(A) = 1\right]\right|.$$

Hàm lợi thế trong tấn công phân biệt hàm f với một hàm ngẫu nhiên hoàn thiện là:

$$\mathrm{Adv}_f^{\mathrm{prf}}(q,\tau) = \max_{A \in \mathcal{A}(q,\tau)} \mathrm{Adv}_f^{\mathrm{prf}}(A),$$

trong đó $\mathcal{A}(q,\tau)$ là tập các bộ phân biệt giả ngẫu nhiên chạy trong thời gian τ sử dụng tối đa q truy vấn.

Tương tự, có định nghĩa $\mathrm{Adv}_f^{\mathrm{prp}}(A)$ khi hàm f là một hoán vị được chọn ngẫu nhiên.

Một hàm được chọn ngẫu nhiên f được gọi là giả ngẫu nhiên nếu như $\mathrm{Adv}_f^{\mathrm{prf}}(A)$ không đáng kể với mọi kẻ tấn công A có năng lực thực tế.

Định nghĩa 2. (Definition 1,* [2], *hàm băm hầu ϵ 2-phổ quát) *Một hàm băm $H: \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ là hầu ϵ 2-phổ quát nếu như mọi $x, y \in \mathcal{X}, x \neq y$ và $h_k(\cdot) \stackrel{\mathrm{def}}{=} H(k,\cdot)$*

$$\Pr\left[k \xleftarrow{\$} \mathcal{K}; h_k(x) = h_k(y)\right] \leq \epsilon.$$

Trong bài báo này, sẽ thống nhất gọi "ϵ-phổ quát" thay cho "hầu ϵ 2-phổ quát".

***Tính chất 1. (tr 5,* [2]).** *Xét $H: \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ là một hàm băm ϵ-phổ quát. Gọi $x_1, \ldots, x_q$ là q thông điệp khác nhau. Khi đó:*

$$\Pr\left[k \xleftarrow{\$} \mathcal{K}; \exists 1 \leq i < j \leq q: h_k(x_i) = h_k(x_j)\right] \leq \epsilon \cdot q^2.$$

Tiếp theo, bài báo trình bày định nghĩa mã xác thực thông điệp và mô hình an toàn của nó. Để thuận tiện cho các phân tích và đánh giá ở

các phần sau, những khái niệm sau đây được nhắc lại.

Định nghĩa 3. (xem Definition 4.1,* [5]*) *Một mã xác thực thông điệp (MAC) gồm có 3 thuật toán thời gian đa thức* ***(Gen, Mac, Vrfy)*** *thỏa mãn:*

1. *Thuật toán sinh khóa* ***Gen*** *là phép chọn khóa K ngẫu nhiên từ tập khóa $\mathcal{K}$.*
2. *Thuật toán sinh nhãn* ***Mac*** *(có thể xác suất) lấy đầu vào là K và thông điệp $M \in I_*$ và đưa ra nhãn T. Ta ký hiệu $T \leftarrow \boldsymbol{Mac}_K(M)$.*
3. *Thuật toán xác thực* ***Vrfy*** *tất định lấy đầu vào là khóa K, thông điệp M và nhãn T. Thuật toán đưa ra một bit b, với $b = 1$ nghĩa là hợp lệ còn $b = 0$ thì ngược lại. Ta viết lại $b \coloneqq \boldsymbol{Vrfy}_K(M,T)$.*

Với mọi khóa K được sinh bởi ***Gen*** *và mọi $M \in I_*$ thì luôn có $\boldsymbol{Vrfy}_K\big(M, \boldsymbol{Mac}_K(M)\big) = 1$.*

Mã xác thực thông điệp an toàn nghĩa là không có một kẻ tấn công hiệu quả nào có thể giả mạo một giá trị nhãn cho thông điệp mới M bất kỳ, mà chưa từng được sử dụng để trao đổi trước đây.

Thí nghiệm xác thực thông điệp $\mathbf{Exp}^{\mathbf{MAC}}(A)$

1. Chạy thuật toán **Gen** sinh ra khóa K.
2. Kẻ tấn công A thực hiện tối đa q_{mac} truy vấn lên bộ tiên tri $\mathbf{Mac}_K(\cdot)$. Gọi $Q = \{M_i\}$ là tập tất cả các truy vấn mà A yêu cầu lên bộ tiên tri.
3. Kẻ tấn công đưa ra tối đa q_{vrfy} truy vấn xác thực lên bộ tiên tri $\mathbf{Vrfy}_K(\cdot)$. A thành công khi và chỉ khi (1) $\mathbf{Vrfy}_K(M,T) = 1$ với cặp truy vấn xác thực (M,T) nào đó và (2) $M \notin Q$. Trong trường hợp này thí nghiệm đưa ra 1, ngược lại thí nghiệm đưa ra 0.

Định nghĩa 4. (Xem Definition 4.2* [5]*) *Xét Π =****(Gen, Mac, Vrfy)*** *là một mã xác thực thông điệp và A là một thuật toán thời gian đa thức xác suất được quyền truy cập lên bộ tiên tri $\boldsymbol{Mac}_K(\cdot)$ và $\boldsymbol{Vrfy}_K(\cdot)$ sau đó trả về một bit như trong thí nghiệm trên.*

Lợi thế giả mạo của A được định nghĩa là

$$\mathrm{Adv}_{\Pi}^{\mathrm{MAC}}(A) = \Pr[\mathrm{Exp}^{\mathrm{MAC}}(A) = 1].$$

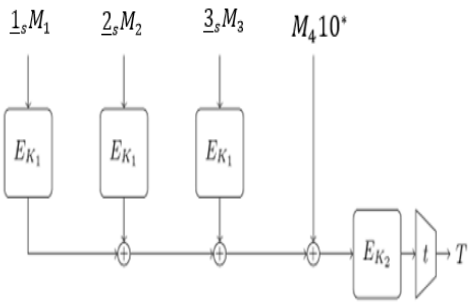
Hàm lợi thế trong tấn công giả mạo là

$$\mathrm{Adv}_{\Pi}^{\mathrm{MAC}}(q_{mac}, q_{vrfy}, \tau) = \max_{A \in \mathcal{A}(q_{mac}, q_{vrfy}, \tau)} \mathrm{Adv}_{\Pi}^{\mathrm{MAC}}(A),$$

trong đó giá trị max lấy trên tất cả kẻ tấn công chạy với thời gian τ, sử dụng nhiều nhất q_{mac} truy vấn Mac và q_{vrfy} truy vấn xác thực.

C. Thuật toán LightMAC

Trong [1] đã giới thiệu thuật toán LightMAC. Mô tả ngắn gọn về thuật toán này được trình ở Hình 1 và Thuật toán 1 dưới đây.



Hình 1. Mô tả thuật toán LightMAC cho thông điệp $M = M_1 \| M_2 \| M_3 \| M_4$ với $M_1, M_2, M_3 \in I_{n-s}$ và $M_4 \in I_{\leq (n-s)}$

Trong đó, $E: I_k \times I_n \to I_n$ là một mã khối, s và t lần lượt là các số nguyên không lớn hơn $n/2$ và n. LightMAC lấy đầu vào là hai khóa K_1, K_2 được chọn đều và độc lập từ tập $\{0,1\}^k$, và thông điệp M có độ dài tối đa $2^s(n-s)$ bit. Thuật toán trả về một đầu ra có độ dài t bit. Cặp thông điệp-nhãn khi đó sẽ là (M,T).

Thuật toán 1. $\mathrm{LightMAC}_{K_1,K_2}(M)$
Input: $K_1, K_2 \in I_k, M \in I_{\leq 2^s(n-s)}$
Output: $T \in I_t$
1. $V \leftarrow 0^n$
2. $M_1 M_2 \dots M_l \xleftarrow{n-s} M$ \\chia M thành các khối $n-s$ bit
3. for $i = 1$ to $l-1$ do
4. $\quad V \leftarrow V \oplus E_{K_1}(\underline{i}_s M_i)$
5. end
6. $V \leftarrow V \oplus (M_l 10^*)$
7. $T \leftarrow \lfloor E_{K_2}(V) \rfloor_t$
8. return T

III. CÁC KẾT QUẢ ĐÃ CÓ

***Định lý 1. (Theorem 2,* [1]).** *Lợi thế giả mạo lên LightMAC của một kẻ tấn công bất kỳ chạy trong thời gian τ thực hiện tối đa q_{mac} truy vấn MAC và q_{vrfy} truy vấn xác thực với độ dài thông điệp tối đa là $2^s(n-s)$ bit, không vượt quá*

$$\left(1+\frac{2}{2^{n/2}-1}+\frac{1}{(2^{n/2}-1)^2}\right)\cdot\left(\frac{q_{mac}^2}{2^n}+\frac{q_{vrfy}}{2^t}\right)+$$
$$\mathrm{Adv}_E^{\mathrm{prp}}(q_{mac}\cdot(2^s-1),\tau_1)+\mathrm{Adv}_E^{\mathrm{prp}}(q_{mac},\tau_2)+\mathrm{Adv}_E^{\mathrm{prp}}(q_{vrfy}2^s,\tau_3),$$

trong đó, n là kích cỡ khối, $\tau_1\in\tau+O\big(q_{mac}\cdot(2^s-1)\big),\tau_2\in\tau+O(q_{mac})$, và $\tau_3\in\tau+O\big(q_{vrfy}2^s\big)$.

Để chứng minh Định lý 1, Atul Luykx đã sử dụng hai Mệnh đề sau:

Mệnh đề 1. *(Proposition 1,* [2]*) (Độ an toàn của băm-rồi-mac) Gọi $H:\mathcal{K}\times\mathcal{X}\to\mathcal{Y}$ là một hàm băm ϵ-phổ quát và $f(\cdot)$ là một hoán vị ngẫu nhiên hoàn thiện trên $\mathcal{Y}$. Xét lược đồ MAC với khóa bí mật $K\xleftarrow{\$}\mathcal{K}$ với nhãn xác thực cho thông điệp $M\in\mathcal{X}$ được tính bởi:*

$$MAC(K,M)=f\big(h_K(M)\big).$$

Gọi A là một kẻ tấn công thực hiện tối đa q_{mac} truy vấn Mac và tối đa q_{vrfy} truy vấn xác thực. Nếu $1/(|\mathcal{Y}|-q_{mac})\le\epsilon$ thì xác suất giả mạo thành công của A không vượt quá:

$$\epsilon\cdot q_{mac}^2+\epsilon\cdot q_{vrfy}.$$

Mệnh đề 2. *(Proposition 1,* [1]*). Đặt $m=2^s(n-s)$. Gọi $M_1M_2\dots M_l\xleftarrow{n-s}M$ với $M\in I_{\le m}$ và định nghĩa F là:*

$$F(M)=M_l10^*\oplus_{i=1}^{l-1}\pi\big(\underline{i}_sM_i\big),$$

Trong đó π là hoán vị ngẫu nhiên hoàn thiện trên I_n, khi đó xác suất để hai thông điệp khác nhau $M,M'\in I_{\le m}$ va chạm là:

$$\Pr[F(M)=F(M')]\le\frac{1}{2^n-l_1-l_2+1},$$

Trong đó l_1 và l_2 lần lượt là độ dài của M và M' theo khối $(n-s)$-bit làm tròn (khối cuối cùng có thể chưa đủ $n-s$ bit, nhưng ta xem như nó là một khối đủ $n-s$ bit).

Tuy nhiên, chúng tôi nhận thấy rằng cách đánh giá của Atul Luykx là dễ gây hiểu nhầm. Bởi vì kết quả trong Mệnh đề 1 chỉ phát biểu cho trường hợp nhãn xác thực là toàn bộ đầu ra của hàm E_{K_2}, trong khi đó LightMAC chỉ lấy t bit.

IV. PHÂN TÍCH CẬN AN TOÀN CỦA LIGHTMAC

Trong phần này, chúng tôi sẽ đánh giá lại cận an toàn cho LightMAC trong trường hợp độ dài nhãn xác thực là t bit ($t<n$).

Đầu tiên, chúng tôi đưa ra mệnh đề sau về độ an toàn của mô hình băm-rồi-mac đối với trường hợp đầu ra của hàm băm bị cắt ngắn.

Mệnh đề 3. *Gọi $H:\mathcal{K}\times\mathcal{X}\to\mathcal{Y}$ là một hàm băm ϵ-phổ quát và $f(\cdot)$ là một hoán vị ngẫu nhiên hoàn thiện trên $\mathcal{Y}$. Xét lược đồ MAC với khóa bí mật $K\xleftarrow{\$}\mathcal{K}$ với nhãn xác thực cho thông điệp $M\in\mathcal{X}$ được tính bởi:*

$$MAC(K,M)=\big\lfloor f\big(h_K(M)\big)\big\rfloor_t.$$

Gọi A là một kẻ tấn công thực hiện tối đa q_{mac} truy vấn Mac và tối đa q_{vrfy} truy vấn xác thực. Xác suất giả mạo thành công của A không vượt quá:

$$\epsilon\cdot q_{mac}^2+q_{vrfy}\cdot\max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\},$$

Chứng minh. Để chứng minh kết quả này ta xét A là một kẻ tấn công lên lược đồ Mac thực hiện tối đa q_{mac} truy vấn Mac và q_{vrfy} truy vấn xác thực. Gọi Coll là sự kiện có xảy ra va chạm giữa hai đầu ra T_1 và T_2 từ bộ tiên tri Mac của hai truy vấn M_1 và M_2 sao cho $T_1=T_2$ và $M_1\ne M_2$. Khi đó ta có:

$$\Pr\big[\mathrm{Exp}^{\mathrm{MAC}}(A)=1\big]$$
$$=\Pr\big[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\mathrm{Coll}\big]\cdot\Pr[\mathrm{Coll}]+\Pr\big[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\overline{\mathrm{Coll}}\big]\cdot\Pr\big[\overline{\mathrm{Coll}}\big]$$
$$\le\Pr[\mathrm{Coll}]+\Pr\big[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\overline{\mathrm{Coll}}\big].$$

Sau đây sẽ lần lượt đánh giá hai xác suất trên Ta có:

$$\Pr[\mathrm{Coll}]$$
$$=\Pr\Big[K\xleftarrow{\$}\mathcal{K};\exists i\ne j:h_K(M_i)=h_K\big(M_j\big)\wedge M_i\ne M_j\Big]$$
$$\le\textstyle\sum_{i\ne j}\Pr\Big[K\xleftarrow{\$}\mathcal{K};h_K(M_i)=h_K\big(M_j\big)\wedge M_i\ne M_j\Big]$$
$$\le q_{mac}^2\cdot\epsilon.\ \text{(theo Tính chất 1).}$$

Tiếp theo ta sẽ chứng minh rằng:

$$\Pr\left[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\overline{\mathrm{Coll}}\right] \le q_{vrfy}\cdot \max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\}.$$

Với mọi $j, 1 \le j \le q_{vrfy}$, đặt P_j là xác suất giả mạo thứ j của A là thành công mà không xảy ra va chạm trong q_{mac} truy vấn MAC. Khi đó ta có $\Pr\left[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\overline{\mathrm{Coll}}\right] \le \sum_1^{q_{vrfy}} P_j$. Ta sẽ chỉ ra rằng $P_j \le \max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\}$ theo phép quy nạp. Chú ý rằng kẻ tấn công đưa ra truy vấn xác thực phải khác những câu trả lời mà bộ tiên tri MAC đã đưa ra trước đó.

Trong trường hợp $j=1$. Nếu A chọn truy vấn xác thực (M_1', T_1') trong đó $T_1' = T_i$ với $1 \le i \le q_{mac}$ nào đó thì có hai trường hợp: $h_K(M_1') = h_K(M_i)$ hoặc $h_K(M_1') \ne h_K(M_i)$ nhưng $\left\lfloor f\left(h_K(M_1)\right)\right\rfloor_t = \left\lfloor f\left(h_K(M_2)\right)\right\rfloor_t$. Trường hợp đầu đồng nghĩa rằng A tìm được một va chạm, tuy nhiên xác suất thành công không vượt quá ϵ. Trong khi trường hợp thứ hai xảy ra với xác suất không quá $1/2^t$. Nếu A chọn truy vấn xác thực (M_1', T_1') với $T_1' \ne T_i$ với mọi $1 \le i \le q_{mac}$, gọi q'_{mac} là số các phần tử các nhãn T_i khác nhau thì ta có:

P_1

$= \Pr\left[f\left(h_K(M_1')\right) = T_1'|\overline{\mathrm{Coll}} \wedge \forall i, 1 \le i \le q_{mac}, : h_K(M_1') \ne h_K(M_i)\right]$

$= 1/(2^t - q'_{mac}) \le 1/(2^t - q_{mac})$.

Do đó $P_1 \le \max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\}$.

Giả sử đã chứng minh đến trường hợp $j-1$, ta sẽ chứng minh rằng $P_j \le \max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\}$. Nếu A chọn truy vấn xác thực $\left(M_j', T_j'\right)$ trong đó $T_j' = T_i$ với $1 \le i \le q_{mac}$ nào đó. Tương tự như trường hợp $j=1$, xác suất để A thành công không vượt quá $\max\left\{\epsilon,\frac{1}{2^t-q_{mac}}\right\}$. Nếu A chọn truy vấn xác thực $\left(M_j', T_j'\right)$ với $T_j' \ne T_i$ với mọi $1 \le i \le q_{mac}$. Ta xét hai trường hợp con. Trường hợp con thứ nhất, $M_j' \ne M_i'$ với $1 \le i \le j$ thì không gian của $T_j' \in \mathcal{Y} - \{T_i\}_{i=1..q_{mac}}$. Khi đó $\Pr\left[\mathrm{Exp}^{\mathrm{MAC}}(A)=1|\overline{\mathrm{Coll}}\right] = 1/(2^t - q'_{mac}) \le 1/(2^t - q_{mac})$. Trường hợp con thứ hai, $M_j' = M_i'$ với $1 \le i < j$, khi đó ta chỉ cần đánh giá xác suất thành công của A khi đưa ra nhãn $T_j' \ne T_i'$ (nếu ngược lại thì sẽ giống với trường hợp P_i). Giả sử rằng, A đã thực hiện r truy vấn xác thực có M_j' giống nhau, ta đánh dấu các lần đó là $j_1, \dots, j_r = j$.

Tương tự cách tính trong trường hợp $j=1$, ta có $P_{j_1} = 1/(2^t - q'_{mac}) \le 1/(2^t - q_{mac})$.

$P_{j_2} = \Pr\left[f\left(h_K\left(M'_{j_2}\right)\right) = T'_{j_2}|\overline{\mathrm{Coll}}\right] \cdot \Pr\left[f\left(h_K\left(M'_{j_2}\right)\right) \ne T'_{j_1}\right]$

$= \frac{1}{2^t - q'_{mac}-1} \cdot \left(1 - \frac{1}{2^t - q'_{mac}}\right) = \frac{1}{2^t - q'_{mac}} \le \frac{1}{2^t - q_{mac}}$.

$P_{j_r} =$

$= \Pr\left[f\left(h_K\left(M'_{j_r}\right)\right) = T'_{j_r}|\overline{\mathrm{Coll}}\right] \cdot \Pr\left[f\left(h_K\left(M'_{j_r}\right)\right) \notin \{T'_{j_1}, \dots, T'_{j_{r-1}}\}\right]$

$= \frac{1}{2^t - q'_{mac}-(r-1)} \cdot \left(1 - \frac{r-1}{2^t - q'_{mac}}\right) = \frac{1}{2^t - q'_{mac}} \le \frac{1}{2^t - q_{mac}}$. ■

Áp dụng Mệnh đề 2 và Mệnh đề 3, chúng tôi đưa ra hệ quả sau:

***Hệ quả 1**. Lợi thế giả mạo lên LightMAC của một kẻ tấn công bất kỳ chạy trong thời gian τ thực hiện tối đa q_{mac} truy vấn MAC và q_{vrfy} truy vấn xác thực với độ dài thông điệp tối đa là $2^s(n-s)$ bit, không vượt quá*

$$\frac{q_{mac}^2}{2^n}\cdot\left(1+\frac{2}{2^{n/2}-1}+\frac{1}{(2^{n/2}-1)^2}\right) + q_{vrfy}\cdot\max\left\{\frac{1}{2^n-2^{n/2+1}+1},\frac{1}{2^t-q_{mac}}\right\} + \mathrm{Adv}_E^{\mathrm{prp}}(q_{mac}\cdot(2^s-1),\tau_1) + \mathrm{Adv}_E^{\mathrm{prp}}(q_{mac},\tau_2) + \mathrm{Adv}_E^{\mathrm{prp}}(q_{vrfy}2^s,\tau_3),$$

trong đó n là kích cỡ khối, $\tau_1 \in \tau + O(q_{mac}\cdot(2^s-1)), \tau_2 \in \tau + O(q_{mac})$, và $\tau_3 \in \tau + O(q_{vrfy}2^s)$.

Chú ý. Trong trường hợp $t < n$ ta luôn có $\frac{1}{2^n-2^{n/2+1}+1} \le \frac{1}{2^t-q_{mac}}$, do đó để thu được kết quả như trong Định lý 1 ta cần phải đảm bảo điều kiện

$$\frac{2^t}{2^t-q_{mac}} \le \frac{2^n}{2^n-2^{n/2+1}+1}.$$

Điều này có nghĩa số lượng truy vấn q_{mac} lên bộ tiên tri **Mac** không được vượt quá $\frac{2^t(2^{n/2}+1)}{2^n}$. Tuy nhiên, việc đánh giá như Định lý 1 là không cần thiết bởi vì nó sẽ làm mất đi ý nghĩa của cận an toàn LightMAC trong trường hợp $t \le n/2$.

Thực tế độ an toàn của mã xác thực LightMAC vẫn phụ thuộc vào độ dài thông điệp vì khi đánh giá va chạm của hàm F vẫn xuất hiện biến độ dài theo khối l. Tuy nhiên, trong cận an toàn của LightMAC có thể biểu diễn thông qua giá trị l khoảng $\frac{q^2}{2^n - l}$, trong khi đối với các mã xác thực thông điệp trước đó là $\frac{q^2 l}{2^n}$. Hơn nữa, LightMAC sử dụng điều kiện số khối của thông điệp l không vượt quá 2^s và $s \leq n/2$ để làm mất đi sự phụ thuộc này. Khi đó, cận an toàn của LightMAC sẽ là $(1+\epsilon)\frac{q^2}{2^n}$ với $\epsilon \in \mathcal{O}\left(\frac{1}{2^{n/2}-1}\right)$, ở đây ta xét với số truy vấn xác thực $q_{vrfy} = 1$. Đối với các mã xác thực như CBC MAC, XOR MAC và PMAC, nếu ta cũng đặt giả thiết rằng số khối của thông điệp không vượt quá một hàm $f(n)$ nào đấy, khi đó cận an toàn của những mã xác thực này cũng không có biến độ dài thông điệp: $f(n) \cdot \frac{q^2}{2^n}$. Tuy nhiên, điều này không có ý nghĩa vì $f(n)$ là một số tương đối lớn và $f(n)$ cũng tượng trưng cho độ dài thông điệp.

V. KẾT LUẬN

Trong bài báo này, chúng tôi đã đánh giá lại cận an toàn cho mã xác thực LightMAC. Sau đó, chúng tôi so sánh sự phụ thuộc vào độ dài của LightMAC với các mã xác thực khác. Tuy nhiên, độ an toàn của LightMAC trong trường hợp sử dụng một khóa duy nhất (ví dụ như sử dụng một khóa K để dẫn xuất ra hai khóa K_1 và K_2) vẫn là câu hỏi mở cần phải nghiên cứu trong thời gian tiếp theo.

TÀI LIỆU THAM KHẢO

[1]. Luykx, A., et al. "A MAC mode for lightweight block ciphers". in International Conference on Fast Software Encryption, Springer, 2016.

[2]. Dodis, Y. and K. Pietrzak. "Improving the security of MACs via randomized message preprocessing". in International Workshop on Fast Software Encryption, Springer, 2007.

[3]. Bellare, M., K. Pietrzak, and P. Rogaway. "Improved security analyses for CBC MACs". in Annual International Cryptology Conference, Springer 2005.

[4]. Bellare, M. and P. Rogaway, "Introduction to modern cryptography". Ucsd Cse p. 207, 2005.

[5]. Katz, J. and Y. Lindell, "Introduction to modern cryptography". CRC press, 2014.

SƠ LƯỢC VỀ TÁC GIẢ

CN. Nguyễn Tuấn Anh

Email: tuananhnghixuan@gmail.com

Quá trình đào tạo: Nhận bằng cử nhân chuyên ngành Toán tài năng tại Đại học Khoa học tự nhiên, Đại học Quốc gia Hà Nội năm 2016.

Hướng nghiên cứu hiện nay: Mã hóa đối xứng.