Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ 28 (2012) 110-115

Nâng cao hiệu quả giấu tin trong ảnh nhị phân

Bùi Văn Tân*

Khoa Công nghệ Thông tin, trường Đại học Kinh tế - Kỹ thuật Công nghiệp,
353 Trần Hưng Đạo, Nam Định, Việt Nam

Nhận ngày 22 tháng 2 năm 2012

Tóm tắt. Giấu tin trong ảnh số là một lĩnh nghiên cứu có những ứng dụng thực tiễn quan trọng, đã có nhiều kỹ thuật giấu tin được đề xuất (Wu-Lee) [1], (CPT) [2]. Với mỗi khối điểm ảnh nhị phân $m \times n$, kỹ thuật giấu tin CPT có thể giấu được $\lfloor \log_2^{mn+1} \rfloor$ bit bằng cách thay đổi tối đa hai bit trên khối ảnh môi trường. Một số nghiên cứu đề xuất cải tiến cho kỹ thuật CPT đã được đưa ra: Nhóm tác giả Đào Thanh Tĩnh [3] đã chứng minh có thể giấu các giá trị nguyên trong miền $R_{mn} = \{0..mn-1\}$ vào mỗi khối $m \times n$; Một cải tiến khác được đề xuất bởi nhóm tác giả Phan Trung Huy [4], nâng hiệu quả giấu tin trên mỗi khối $m \times n$ lên $\lfloor \log_2^{mn} \rfloor + 1$ bit. Trong nghiên cứu này, tác giả đề xuất một cải tiến mới cho kỹ thuật CPT, với mỗi khối $m \times n$, kỹ thuật cải tiến có khả năng giấu được các giá trị nguyên không âm lên đến $2mn-1$, cũng với tối đa hai sự thay đổi trên khối ảnh môi trường.

1. Giới thiệu

Lịch sử loài người đã cho thấy, con người luôn có nhu cầu trao đổi, truyền những thông điệp, hình ảnh (Information) từ nơi này đến nơi khác một cách an toàn và bảo mật. Thuở sơ khai, con người đã biết sử dụng các đám khói hay chim bồ câu, chim nhạn để truyền đi những thông điệp giữa các cánh quân trong một trận chiến. Ngày nay, mạng Internet trở thành một môi trường truyền tải dữ liệu phục vụ nhu cầu trao đổi thông tin, thương mại điện tử, Chính phủ điện tử..., đem lại những đổi thay to lớn trong đời sống con người. Bên cạnh đó, cũng đặt ra những yêu cầu hết sức cấp thiết về an toàn và bảo mật thông tin trong thông tin liên lạc.

Các công nghệ và giải pháp để bảo vệ thông tin đã và đang được nghiên cứu, phát triển phù hợp với dạng lưu trữ của thông tin và cách thức truyền tin. Giải pháp bảo mật thông tin hiện đang được sử dụng phổ biến nhất là các hệ mật mã. Với giải pháp này, thông tin ban đầu (bản rõ) sẽ được mã hóa thành bản mật mã (bản mật) thường mang những giá trị "*vô nghĩa*". Chính điều này làm cho đối phương nghi ngờ và tìm cách thám mã, tấn công.

Một hướng tiếp cận khác là đem giấu thông tin quan trọng vào trong một đối tượng "*mang*", sao cho người ngoài khó có thể nhận biết được

* ĐT: 84-912611550.
E-mail: Tanbv.it@gmail.com

việc giấu này. Đối tượng mang hay môi trường để giấu tin có thể là bất kỳ đối tượng dữ liệu số nào nhưng phổ biến hơn cả là ảnh số [1]. Đối với môi trường giấu là ảnh màu đã có khá nhiều phương pháp giấu tin hiệu quả, tuy nhiên các phương pháp này lại khó có thể áp dụng cho ảnh đen trắng, vì đối với ảnh đen trắng, mỗi điểm ảnh được biểu diễn bởi một bit nên khi thay đổi các bit sẽ rất dễ bị phát hiện. Kỹ thuật giấu tin do ba tác giả Y.Y. Chen, H. Pan, Y. Tseng (CPT) đề xuất [2], được đánh giá đạt hiệu quả tốt đối với ảnh đen trắng.

2. Kỹ thuật giấu tin CPT

Trong báo cáo "*A secure data hiding scheme for two-color images*" công bố tại hội nghị chuyên đề về máy tính và truyền thông năm 2000 [2], Y.Y. Chen, H. Pan, Y. Tseng đề xuất một kỹ thuật giấu thông tin trong ảnh nhị phân, ảnh được chia thành nhiều khối có cùng kích thước $m\times n$ pixel ($F_{m\times n}$). Với mỗi khối dữ liệu ảnh có thể giấu được tối đa *r* bit ($r=\left\lfloor \log_2^{(mn+1)} \right\rfloor$) thông tin, bằng cách thay đổi không quá hai bit trong khối dữ liệu ảnh. Kỹ thuật này sử dụng ma trận khóa nhị phân $K_{m\times n}$ có kích thước bằng khối ảnh; Ma trận trọng số $W_{m\times n}$ có các phần tử thoả mãn điều kiện $\{w_{ij} \mid 1\le i\le m, 1\le j\le n\}=\{1,2,3,...,2^r-1\}$. Gọi *b* là giá trị nguyên không âm của dãy *r* bit nhị phân cần giấu b_1, b_2,..., b_r. Cho các ma trận nhị phân *A*, *B*, *C* cùng cấp, toán tử $\oplus,\otimes$ được định nghĩa như sau:

$$C=A\oplus B \quad \Leftrightarrow \quad C_{ij}=\begin{cases}1 & if \quad A_{ij}\neq B_{ij}\\ 0 & if \quad A_{ij}=B_{ij}\end{cases} \quad (1)$$

$$C=A\otimes B \quad \Leftrightarrow \quad C_{ij}=A_{ij}\times B_{ij} \quad (2)$$

Kỹ thuật CPT sẽ thực hiện đảo giá trị bit tối đa tại 2 vị trí trên ma trận điểm ảnh $F_{m\times n}$, để được ma trận $G_{m\times n}$ thoả mãn bất biến: $b=SUM((G\oplus K)\otimes W)\bmod 2$. Thuật toán CPT được mô tả gồm 4 bước:

Bước 1: Tính $T=F\oplus K$

Bước 2:

Tính $S=SUM(T\otimes W)\bmod 2^r$

$$\alpha=(b-S)\bmod 2^r$$

Bước 3:

Xây dựng tập

$$Z_\alpha=\{(i,j)\mid((W_{ij}=\alpha)\ and\ (T_{ij}=0))\ or\ ((W_{ij}=2^r-\alpha)\ and\ (T_{ij}=1))\}$$

Bước 4: Dựa vào tập Z_α, đảo giá trị bit tại tối đa 2 vị trí trong ma trận *F* để được ma trận *G*. Sao cho với

$S'=SUM(G\oplus K)\otimes W)\bmod 2^r$ thì ta luôn có $S'=b\bmod 2^r$.

Với hai ma trận $K_{m\times n}$, $W_{m\times n}$ dùng làm khoá để giải tin, độ bảo mật của thuật toán khá cao, có thể lựa chọn *K* và *W* từ: $2^{mn}C_{2^r-1}^{mn}\times(2^r-1)!\times(2^r-1)^{mn-(2^r-1)}$ khả năng.

Một số nghiên cứu đã đề xuất những cải tiến cho kỹ thuật này. Nhóm tác giả Đào Thanh Tĩnh, Tống Minh Đức [3], chứng minh có thể giấu các số nguyên có giá trị tối đa bằng $mn-1$ vào mỗi khối $m\times n$; Một cải tiến khác được đề xuất bởi nhóm tác giả Phan Trung Huy [4], kỹ thuật mới giấu được nhiều hơn kỹ thuật gốc 1 bit trên mỗi khối $m\times n$. Trong nghiên cứu này, tác giả đề xuất một cải tiến mới cho kỹ thuật CPT. Với mỗi khối điểm ảnh nhị phân $m\times n$, thuật toán cải tiến có khả năng giấu được các giá trị nguyên tối đa là $2mn-1$ và

cũng chỉ thay đổi nhiều nhất hai bit trên khối ảnh môi trường.

3. Kỹ thuật giấu tin CPT cải tiến

Kỹ thuật cải tiến thực hiện giấu các giá trị nguyên vào các khối điểm ảnh có kích thước $m \times n$.

Input:

Ma trận ảnh nhị phân $F = (f_{ij})_{m\times n}$

Ma trận khóa nhị phân $K = (k_{ij})_{m\times n}$

p là hằng số, $p = mn$

q là hằng số, $q = 2mn$

Ma trận trọng số

$W = \{w_{ij} \mid 1 \le i \le m, 1 \le j \le n\} = \{1,2,3,..., p\}$

$0 \le b \le q-1$ là giá trị cần giấu.

Output:

Ma trận nhị phân $G = (g_{ij})_{m\times n}$ đã được giấu giá trị b.

3.1. Nội dung thuật toán cải tiến

Thuật toán giấu b ($0 \le b \le q-1$) trong một khối điểm ảnh nhị phân F gồm 5 bước:

Bước 1. {Khởi tạo}

$T = F \oplus K$.

$S = Sum[T \otimes W] \bmod q, \quad 0 \le S \le q-1$.

Đặt $\alpha = (b - S) \bmod q$.

If $\alpha = 0$ then

Begin

G:=F; goto Bước 5;

End;

Bước 2. {Xây dựng tập S_α}

$$S_\alpha = \begin{cases} \{(F_{ij}) \mid (T_{ij} = 0) \text{ and } (W_{ij} = \alpha)\}, \text{if } 1 \le \alpha < p \\ \{(F_{ij}) \mid (T_{ij} = 1) \text{ and } (W_{ij} = q - \alpha)\}, \text{if } p < \alpha \le q-1 \\ \{(F_{ij}) \mid W_{ij} = \alpha\}, \text{if } \alpha = p \end{cases}$$

Bước 3. {Trường hợp thay đổi một bit}

If $S_\alpha \ne \phi$ then

Begin

Đảo giá trị bit tại vị trí F_{ij} với $F_{ij} \in S_\alpha$;

G:= F;

Goto Bước 5;

End;

Bước 4. {Trường hợp thay đổi hai bit}

Tìm số tự nhiên $h>1$ nhỏ nhất sao cho $S_{h\alpha} \ne \phi$ và $S_{\alpha - h\alpha} \ne \phi$

Đảo giá trị bit tại vị trí $F_{ij} \in S_\alpha$

Đảo giá trị bit tại vị trí $F_{uv} \in S_{\alpha - h\alpha}$

G:= F;

Bước 5. {Kết thúc}

Return G;

3.2. Chứng minh tính đúng đắn của thuật toán cải tiến.

Cũng giống như đối với thuật toán CPT, có nhiều cách để chứng minh tính đúng đắn của thuật toán cải tiến, có thể chứng minh trực tiếp hoặc sử dụng định lý Diophantus... Ở đây tác giả tham khảo cách chứng minh được trình bày trong [4]. Việc chứng minh tính đúng đắn của thuật toán cải tiến, được quy về chứng minh định lý 1 sau đây.

Định lý 1.

Cho F, K là ma trận bit cỡ $m \times n$ và W là ma trận các số tự nhiên cùng cỡ thỏa mãn: $\{w_{ij} \mid 1 \le i \le m, 1 \le j \le n\} = \{1,2,3,..., p\}$ với $p = mn$.

Cho $0 \le b \le q-1$ là giá trị cần giấu. Có thể đảo giá trị nhiều nhất tại hai vị trí trên F để được bất biến $b = Sum[T \otimes W] \bmod q$.

Để phục vụ cho việc chứng mình định lý 1, ta sẽ chứng minh các bổ đề sau:

Bổ đề 1: **Cho $\alpha \neq 0$, $S_\alpha = \phi$ thì** $S_{q-\alpha} \neq \phi$.

Ta có: $S = SUM[T \otimes W] \bmod q$, $\alpha \in \{1,2,...,q-1\}$, giả sử ta có $W_{ij} \in W$ thì $W_{ij} = \alpha$ (nếu $1 \le \alpha \le p$) hoặc $W_{ij} = q - \alpha$ (nếu $p \le \alpha \le q-1$). Theo giả thiết $S_\alpha = \phi$ nên mỗi cặp (*i,j*) có hai trường hợp sau:

- $W_{ij} = \alpha$ thì $T_{ij} = 1$
- $W_{ij} = q - \alpha$ thì $T_{ij} = 0$

Ta thấy $W_{ij} = q - \alpha$ thì $T_{ij} = 0$ nên $S_{q-\alpha} \neq \phi$.

Bổ đề 2: $S_p \neq \phi$.

Giả sử $S_p = \phi$ theo bổ đề 1 ta có $S_{q-p} \neq \phi$, mà $q = 2mn = 2p$ vậy suy ra $S_p \neq \phi$, điều mẫu thuẫn với giả thiết. Do đó $S_p \neq \phi$.

Bổ đề 3: Cho $\alpha \neq 0 \bmod q$ và $S_\alpha = \phi$ Tồn tại *h>1* sao cho $S_{ha} \neq \phi$. Nếu *h* là số tự nhiên nhỏ nhất thỏa mãn $S_{ha} \neq \phi$ và *h>1*, thì $S_{\alpha-ha} \neq \phi$.

Ta viết $\alpha = 2^l \times s$, với *s* là số lẻ và $l \ge 0$, $1 \le 2^l \times s \le q-1$, $r = \log_2 p$, xét hai trường hợp sau:

Trường hợp 1, *S=1* ta có:

$$\left.\begin{array}{l} \alpha = 2^l \\ S_\alpha = \phi \\ S_{2^r} \neq \phi \end{array}\right\} \Rightarrow 2^l < 2^r \text{ vậy } l < r.$$

Đặt $h = 2^{r-l} \Rightarrow ha = 2^{r-l} \times 2^l = 2^r = p$

Vậy suy ra $S_{ha} = S_p \neq \phi$

Trường hợp 2, *s>1* ta có:

$$\begin{cases} s = 2k+1 \\ 2^l < 2^r \end{cases} \Rightarrow l < r \text{ suy ra}$$

$\alpha = 2^l(2k+1) = 2^{l+1} \times k + 2^l$. Chọn $h = 2^{r-l}$, chúng ta có $h > 1$ và $h\alpha = (2^{l+1} \times k + 2^l) \times 2^{r-l} = 2^{r+1} \times k + 2^r = 2^r \bmod 2^{r+1} = p \bmod q$ điều này chứng tỏ $S_{ha} = S_p \neq \Phi$

Do đó trong trong mọi trường hợp ta luôn chọn được số tự nhiên *h>1* sao cho $S_{ha} \neq \phi$, với *h* là số tự nhiên nhỏ nhất nên $S_{(h-1)a} = \phi$, theo bổ đề 2.5 ta có $S_{q-(h-1)a} \neq \phi$ hay $S_{a-ha} \neq \phi$

Chứng minh định lý 1:

Chúng ta sẽ xem xét ba trường hợp:

+ Trường hợp $S = b \bmod q$ hay $\alpha = 0$, ta không cần thay đổi ma trận *F*

+ Trường hợp $\alpha \neq 0$ *and* $S_\alpha \neq \phi$, đảo giá trị bit đúng một vị trí F_{ij} ($F_{ij} \in S_\alpha$) trên ma trận *F* ta thu được: $S(new) = (S + \alpha) \bmod q = b$

Trường hợp $\alpha \neq 0$ *and* $S_\alpha = \phi$, theo bổ đề 3, luôn tồn tại một số tự nhiên nhỏ nhất *h >1* sao cho $S_{h\alpha} \neq \phi$ *and* $S_{\alpha-h\alpha} \neq \phi$, đảo giá trị bit tại $F_{ij} \in S_{h\alpha}$ và $F_{uv} \in S_{\alpha-h\alpha}$ ta thu được: $S(new) = S + ha + (\alpha - h\alpha) = b \bmod q = b$.

3.3. Đánh giá hiệu quả kỹ thuật giấu tin cải tiến

Hiệu quả của một kỹ thuật giấu tin, thường dựa trên sự đánh giá các tiêu chí: Dung lượng tin được giấu; tính "*vô hình*" của tin giấu; độ bền vững của tin giấu; độ phức tạp của thuật toán giấu và thám tin. Để so sánh hiệu quả của kỹ thuật giấu tin cải tiến với CPT, ta đánh giá hai kỹ thuật này thông qua bốn tiêu chí nêu trên.

a) Dung lượng tin giấu

Những đề xuất cải tiến trình bày ở trên, đã mang lại cho kỹ thuật mới hiệu quả giấu tin cao hơn hẳn so với kỹ thuật CPT. Trên mỗi khối điểm ảnh, khả năng giấu tin của kỹ thuật cải tiến luôn nhiều hơn CPT ít nhất một bit, trong một số trường hợp nó có thể giấu được nhiều hơn CPT đến hai bit. Với một ảnh được chia thành K khối để giấu tin, kỹ thuật giấu tin cải tiến có thể giấu nhiều hơn CPT ít nhất là *K* bit, nhiều nhất là *2K* bit. Bảng 1 so sánh hiệu quả giấu tin của hai kỹ thuật này trên một số kích thước khối điểm ảnh.

Kích Thước Khối $(m\times n)$	Thuật toán CPT gốc		Thuật toán cải tiến	
	$b_{\max}=(2^{\log_2(mn+1)}-1)$	số bit tối đa	$b_{\max}=(2mn-1)$	số bit tối đa
4×4	15_{10} hay 1111_2	4	31_{10} hay 11111_2	5
5×7	31_{10} hay 11111_2	5	69_{10} hay 1000101_2	7
8×8	63_{10} hay 111111_2	6	127_{10} hay 1111111_2	7
8×10	63_{10} hay 111111_2	6	159_{10} hay 10011111_2	8
17×10	127_{10} hay 1111111_2	7	339_{10} hay 101010011_2	9

(Bảng 1 – So sánh dung lượng tin giấu của hai thuật toán trên một số khối)

Ví dụ, xét ảnh một nhị phân có kích thước 1024×800, ảnh được chia thành các khối 5×7 để giấu tin. Khi đó, kỹ thuật giấu tin cải tiến sẽ giấu nhiều hơn CPT xấp xỉ 5851 Byte hay một chuỗi dài 5851 ký tự mã ASCII, chuỗi ký tự này tương đương với một bức thư dài bốn trang giấy.

b) Tính "vô hình" của tin giấu

Giấu tin trong ảnh ít nhiều cũng gây ra những thay đổi trên dữ liệu ảnh môi trường. Tính "*vô hình*" của kỹ thuật giấu tin thể hiện mức độ biến đổi ảnh môi trường. Để giấu tin vào một khối điểm ảnh môi trường, cả kỹ thuật giấu tin cải tiến và CPT đều thay đổi số lượng điểm ảnh như nhau. Do đó, hai kỹ thuật có mức độ ảnh hưởng đến ảnh mang là hoàn toàn như nhau.

c) Tính bền vững của tin giấu

Tính bền vững của kỹ thuật giấu tin là khả năng chống lại các tấn công có chủ định hoặc không có chủ định lên ảnh có giấu tin. Việc tấn công được thực hiện thông qua các phép biến đổi ảnh khác nhau như lọc, thêm nhiễu, quay, nén mất mát thông tin,…Kỹ thuật giấu tin cải tiến và CPT đều thực hiện giấu tin trên miền không gian ảnh vì vậy thông tin mật có cùng độ bền vững.

d) Độ phức tạp của thuật toán

Độ phức tạp của thuật toán được Dễ nhận thấy, để giấu tin trên một khối điểm ảnh nhị phân $F_{m\times n}$, thuật toán giấu tin của hai kỹ thuật đều có độ phức tạp về thời gian trong trường hợp xấu nhất là $O(mn)$. Đồng thời, hai kỹ thuật sử dụng ma trận $W_{m\times n}$, $K_{m\times n}$ làm khoá để giải tin. Đối với kẻ thám tin, muốn "*tách*" tin mật được giấu bởi kỹ thuật cải tiến hay CPT,

đều phải tìm được ma trận khóa nhị phân $K_{m\times n}$ và ma trận trọng số $W_{m\times n}$. Do đó, độ an toàn tin giấu của hai kỹ thuật này là tương đương nhau.

4. Kết luận

Từ kết quả khảo sát ở bảng 1 và thực nghiệm giấu tin trên một số mẫu ảnh. Có thể thấy rằng, trên cùng một tệp ảnh, kỹ thuật giấu tin cải tiến có khả năng giấu một lượng thông tin lớn hơn so với kỹ thuật CPT. Điều đáng chú ý là hai kỹ thuật tương đương nhau về tính "*vô hình*", tính bền vững và độ an toàn của thông tin giấu.

Tài liệu tham khảo

[1] M. Wu, J. Lee (1998), *A novel data embedding method for two-color fascimile images*. In Proceedings of international symposium on multimedia information processing. Chung-Li, Taiwan, R.O.C.

[2] Y. Chen, H. Pan, Y. Tseng (2000). *A secure data hiding scheme for two- color images*. In IEEE symposium on computers and communications.

[3] Đào Thanh Tĩnh, Tống Minh Đức (2008), *Một cải tiến thuật toán giấu tin trong ảnh nhị phân*, tạp chí Công nghệ Thông tin và Truyền thông, số 20, tháng 10 năm 2008.

[4] Phan Trung Huy, Nguyễn Mạnh Thắng, Trương Đức Mạnh, Vũ Phương Bắc, Vũ Tiến Đức, Nguyễn Tuấn Nam, *A new CPT extension scheme for high data embedding ratio in binary images*, KSE '09 Proceedings of the 2009 International Conference on Knowledge and Systems Engineering.

Improve efficiency for Hiding Information in Binary Image

Bui Van Tan

University of Economic and Technical Industries, 353 Tran Hung Dao, Nam Dinh, Vietnam

The hidden information of digital images is a important research area that exhibited practical applications and there are many hidden techniques have been proposed (Wu-Lee) [3], (CPT) [4]. Within a binary pixel block $m\times n$, hidden technique can hide $\lfloor \log_2^{mn+1} \rfloor$ of bit through the maximum changing of two bits of environmental pixel block. Some ideals of improvement of CPT technique were shown: Dao Thanh Tinh group, demonstrated the possibility of hidden range of integer values R_{mn}={0..mn-1} in a block $m\times n$; another improvement was introduced by Phan Trung Huy group [2], the performance of hidden information of a block $m\times n$ was increased to $\lfloor \log_2^{mn} \rfloor$ + 1 bits. In this research, the authors supported a novel development of CPT technique, for each block $m\times n$, new developed technique could hide non-negative integer values through $2mn-1$ for the maximum changing of two bits of environmental pixel block, as well.