

MỘT THUẬT TOÁN GIẤU TIN TRONG ẢNH CÓ BẢNG MÀU VÀ ỨNG DỤNG GIẤU TIN MẬT TRONG ẢNH GIF

Trần Quang Sơn, Nguyễn Văn Tào (*Khoa Công nghệ thông tin - ĐH Thái Nguyên*)

Tóm tắt

Ngày nay, nhu cầu trao đổi thông tin trên mạng là rất lớn. Theo đó, vấn đề bảo đảm an toàn cho những thông tin mật cũng trở nên cấp thiết. Có nhiều giải pháp nhằm đảm bảo an toàn cho thông tin trao đổi, giấu tin trong ảnh là một giải pháp được nhiều nhà khoa học quan tâm nghiên cứu. Bài báo này đề xuất một thuật toán giấu tin trong ảnh có bảng màu. Giới thiệu việc áp dụng thuật toán cho giấu tin trong định dạng ảnh GIF.

Từ khoá: Giấu thông tin, an toàn thông tin, ảnh số, bảng màu.

I. TỔNG QUAN

Trong môi trường phân phối điện tử rất phát triển như hiện nay, việc bảo vệ cho các thông tin quan trọng trong quá trình trao đổi trở nên cấp thiết. Theo phương pháp truyền thống, thông tin mật trước khi truyền đi sẽ được mã hóa, như vậy trong quá trình truyền, những người ngoài cuộc quan sát bản tin đã mã hóa sẽ biết được tầm quan trọng của bản tin trao đổi, điều đó làm tăng sự tò mò muốn khám phá để tìm ra nội dung thực của bản tin.

Gần đây, một phương pháp mới được nhiều nhà khoa học quan tâm nghiên cứu đó là nhúng các thông tin mật vào các đối tượng dữ liệu khác (phương tiện chứa) như ảnh, audio, video,... rồi sử dụng chính các phương tiện chứa đã bao gồm thông tin mật để trao đổi.

Bài báo này đề xuất một thuật toán giấu tin mật trong ảnh có bảng màu với một sự thay đổi cảm nhận về ảnh là ít nhất. Từ thuật toán này, chúng tôi xây dựng lược đồ giấu tin trong ảnh áp dụng với định dạng ảnh GIF.

II. MỘT SỐ LƯỢC ĐỒ GIẤU TIN TRONG ẢNH CÓ BẢNG MÀU

1. Cấu trúc bảng màu

Mỗi màu sắc trong máy tính sử dụng 24 bit để biểu diễn: 8 bit cho màu đỏ (*Red*), 8 bit cho màu lục (*Green*), 8 bit cho màu lam (*Blue*). Một bức ảnh định dạng 24 bit (*sử dụng màu sắc thực*) chứa tối đa $2^{24} = 16777216$ màu sắc khác nhau. Như vậy, để biểu diễn thông tin một điểm ảnh (*pixel*) cần phải sử dụng 3 byte: 1 byte Red, 1 byte Green, 1 byte Blue. Tuy nhiên trong thực tế không phải bức ảnh nào cũng chứa tối đa 16777216 màu sắc.

Có rất nhiều màu sắc giống nhau xuất hiện trong một bức ảnh, thay cho sự sao chép lặp lại các thông tin về màu sắc người ta sử dụng cấu trúc bảng màu. Với cấu trúc bảng màu, người ta sử dụng 1 byte lưu trữ cho một điểm ảnh là giá trị tương ứng của một màu trong bảng màu. Một bảng màu thực chất là một vector có số phần tử tối đa là 256 phần tử khác nhau và mỗi một phần tử này bao gồm thông tin của ba thành phần Red, Green, Blue (*RGB*). Như vậy, có thể thấy một bức ảnh có bảng màu chứa tối đa là 256 màu sắc khác nhau.

2. Giấu tin trong bảng màu đơn giản (BS)

Một cấu trúc ảnh có bảng màu thông thường gồm ba phần chính: Phần thông tin mô tả ban đầu (*Header*), phần thông tin bảng màu (*Palette/Color Table*), phần dữ liệu ảnh (*Image Data*). Ý tưởng cơ bản của kỹ thuật này là tiến hành giấu tin vào vị trí các bit ít quan trọng (*Least Significant Bit - LSB*) đối với mỗi phần tử trong bảng màu.

Quá trình giấu tin:

① Với một ảnh gốc *I*, lấy ra phần thông tin bảng màu *P*. Bảng màu bao gồm các bộ ba thành phần *RGB* tương ứng với tất cả các màu trong bức ảnh.

② Thông điệp mật dưới dạng nhị phân là một chuỗi *k* bit $b = b_1, b_2, \dots, b_k$.

③ Mỗi một bit b_i sẽ được giấu vào vị trí *LSB* của một bộ ba thành phần *RGB* tương ứng trong bảng màu theo một thuật toán giấu tin nào đó.

④ Sau quá trình giấu tin sẽ thu được một bảng màu *P'* đã chứa thông tin cần giấu.

Quá trình tách tin:

Khi nhận được ảnh đã giấu tin, việc giải mã tin sẽ được thực hiện theo các bước:

① Đọc ảnh chứa tin, lấy ra được phần thông tin bảng màu P' .

② Mỗi một bit b_i sẽ được trích ra từ vị trí *LSB* của một bộ ba thành phần *RGB* tương ứng trong bảng màu.

Như vậy, sau khi xét hết các thành phần trong bảng màu ta thu được chuỗi các bit, chuỗi này là thông tin nhị phân đã giấu cần phải lấy ra.

a. Lược đồ giấu tin của Romana Machado (EZStego)

Kỹ thuật giấu tin trong bảng màu đơn giản BS để thực hiện. Tuy nhiên, kích thước tin giấu luôn là cố định và không phụ thuộc vào dữ liệu ảnh.

Lược đồ giấu tin của Romana Machado [1] dựa trên ý tưởng cải tiến là sắp xếp lại toàn bộ bảng màu theo thứ tự về độ chói. Từ đó, đề xuất một phương pháp giấu tin vào trong khối dữ liệu ảnh. Mỗi một giá trị điểm ảnh (*pixel*) mang một bit thông tin giấu. Dưới đây là lược đồ mô tả quá trình giấu một bit thông tin b vào trong một điểm ảnh k :

Quá trình giấu tin:

① Sắp xếp lại bảng màu gốc P theo thứ tự về độ chói được bảng màu P_s .

➤ Độ chói $Y = 0.299 * R + 0.587 * G + 0.114 * B$

② Tìm phần tử tương ứng I_s của điểm ảnh k trong bảng màu P_s .

③ Thay thế bit thông tin giấu b vào LSB của I_s thu được I_s' .

④ Tìm phần tử I trong bảng màu gốc P tương ứng với I_s' trong bảng màu P_s .

⑤ Thay thế giá trị của điểm ảnh k bằng I .

Quá trình trích tin:

① Sắp xếp lại bảng màu gốc P theo thứ tự về độ chói được bảng màu P_s .

② Tìm phần tử tương ứng I_s của điểm ảnh k trong bảng màu P_s .

③ Trích ra LSB trong I_s thu được bit thông tin b .

b. Lược đồ giấu tin của Jessica Fridrich và Du Rui (FR)

Kỹ thuật giấu tin trong ảnh có bảng màu của Romana Machado đã khắc phục được nhược điểm của lược đồ BS. Tuy nhiên, trong lược đồ của EZ các bước thực hiện còn khá phức tạp và một thực tế cho thấy là hai thành phần màu có thể cùng độ

chói nhưng lại rất khác nhau về nhận biết. Để khắc phục nhược điểm tính toán phức tạp và những ngoại lệ có thể xảy ra J.Fridrich và D.Rui [2] đã đề xuất ra một kỹ thuật cải tiến giảm đi quá trình tính toán và tăng chất lượng ảnh sau khi giấu tin.

Lược đồ giấu tin FR về cơ bản vẫn dựa trên tư tưởng giấu một bit thông tin vào trong điểm ảnh nhưng đã khắc phục được những vấn đề trên bằng ý tưởng khoảng cách giữa hai màu $C_1=(R_1, G_1, B_1)$ và $C_2=(R_2, G_2, B_2)$ tìm ra cặp màu gần nhau nhất. Dưới đây là lược đồ mô tả quá trình giấu một bit thông tin b vào trong một điểm ảnh k :

Quá trình giấu tin:

① Với mỗi bit thông tin b chọn ngẫu nhiên một điểm ảnh k có màu $C_1(R_1, G_1, B_1)$. Tìm trong bảng màu P màu gần nhất $C_2(R_2, G_2, B_2)$ với C_1 thỏa mãn khoảng cách $D_{1,2}$ giữa hai màu là nhỏ nhất đồng thời thỏa mãn điều kiện chẵn lẻ (*parity*) bit:

$$D_{1,2}^2 = (R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2 \quad (D_{1,2}^2 > 0) \text{ và } b = (R_2 + G_2 + B_2) \text{ Mod } 2$$

② Thay thế giá trị màu C_1 của điểm ảnh k bằng một giá trị màu mới C_2 .

Quá trình trích tin:

① Đọc giá trị màu của điểm ảnh k là $C_1(R_1, G_1, B_1)$ tiến hành trích ra bit thông tin b theo biểu thức:

$$b = (R_1 + G_1 + B_1) \text{ Mod } 2$$

III. ĐỀ XUẤT LƯỢC ĐỒ GIẤU TIN TRONG ẢNH CÓ BẢNG MÀU

1. Ý tưởng

Một bảng màu chứa tối đa là 256 màu khác nhau. Nhưng thực tế qua quá trình kiểm nghiệm chúng tôi nhận thấy rằng số lượng thành phần màu trong một bảng màu thường là nhỏ hơn 256. Đối với một số định dạng ảnh thông dụng như ảnh GIF, qua quá trình kiểm tra bảng màu thì số lượng thành phần màu thường là khá nhỏ (khoảng 128 màu). Với nhiều ảnh có số lượng thành phần màu thấp thì khoảng cách giữa các màu lại cao, sự phân biệt giữa các màu là rõ rệt. Như vậy có thể thấy rằng nếu áp dụng lược đồ giấu tin FR với những trường hợp này sẽ có thể gây ra những khác biệt về cảm nhận lớn. Dựa trên những nhận xét như vậy, chúng tôi đề xuất một kỹ thuật cải tiến quá

trình biến đổi bảng màu đảm bảo được sự thay đổi về cảm nhận là nhỏ nhất có thể.

2.Quá trình phân tích bảng màu (thuật toán ST)

Với đề xuất trên, chúng tôi đưa ra một lược đồ phân tích bảng màu với mục tiêu giảm thiểu nhất sự thay đổi về cảm nhận sau khi giấu tin.

① Phân tích ảnh, thống kê xác suất xuất hiện của mỗi màu trong bảng màu. Sắp xếp lại bảng thống kê theo thứ tự giảm dần về tần số xuất hiện.

② Thay thế những màu có tần số xuất hiện quá ít (nhỏ hơn 10) bằng những màu gần nhất với những màu có tần số xuất hiện nhiều. Màu gần nhất của $C_1 = (R_1, G_1, B_1)$ là C_2 được tính như sau:

$$C_2 = (R_1, G_1, B_1 \pm 1)$$

③ Làm mịn dần quá trình biến đổi bảng màu:

➤ Tính $D_i (i = 0..255), D_i = j$ với $Min |C_i - C_j|_{j = 0..255, j \neq i}$ trong đó khoảng cách $|C_i - C_j| = (R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2$

➤ Sắp xếp các cặp $i, D_i (i = 0..255)$ theo thứ tự tăng dần về khoảng cách

➤ $C = \emptyset$, lặp lại quá trình sau cho tới khi C chứa tất cả các màu trong bảng màu: Nếu tồn tại cặp i, D_i với $i \notin C, D_i \notin C$ thỏa mãn $(R_i + G_i + B_i) \ll (R_{D_i} + G_{D_i} + B_{D_i}) \text{ Mod } 2$ thì kết nạp vào tập C cặp giá trị i, D_i . Ngược lại, với $i \notin C, D_i \notin C$ thỏa mãn $(R_i + G_i + B_i) = (R_{D_i} + G_{D_i} + B_{D_i}) \text{ Mod } 2$ thì tiến hành biến đổi về $(R_i + G_i + B_i) \ll (R_{D_i} + G_{D_i} + B_{D_i}) \text{ Mod } 2$ rồi kết nạp vào tập C cặp giá trị i, D_i

a. Quá trình giấu tin

Tiến hành nhúng chuỗi k bit $b = (b_1, b_2, \dots, b_k)$ vào chuỗi n điểm ảnh $p=(p_1, p_2, \dots, p_n)$ với $p_i = R_i, G_i, B_i$ thu được chuỗi p' theo các bước [3]:

① Đọc bảng màu và biến đổi bảng màu theo thuật toán ST.

② Tính $f(x) = \bigoplus_{i=1}^n x_i \times db(i)$, trong đó $db(i)$ là biểu diễn nhị phân của $i, x_i = (R_i + G_i + B_i) \text{ Mod } 2$.

③ Tính $s = b \oplus f(x)$.

④ Nếu $s = 0$ thì lấy $p' = p$, ngược lại thay thế p_s bởi p_s' thỏa mãn

➤ $D_{ss'}^2 = (R_s - R_{s'})^2 + (G_s - G_{s'})^2 + (B_s - B_{s'})^2$ nhỏ nhất.

➤ $R_s + G_s + B_s \ll R_{s'} + G_{s'} + B_{s'}$.

⑤ Sắp xếp lại bảng màu thu được sau khi biến đổi theo thứ tự về độ chói hoặc khoảng cách màu.

b.Quá trình trích tin

Quá trình lọc tìm lại b từ chuỗi p' với $p'_i = R'_i, G'_i, B'_i$ được thực hiện theo công thức:

$$b = f(x') \text{ với } x'_i = (R'_i + G'_i + B'_i) \text{ Mod } 2.$$

IV.ÁP DỤNG THUẬT TOÁN ST ĐỂ GIẤU TIN TRONG ẢNH GIF

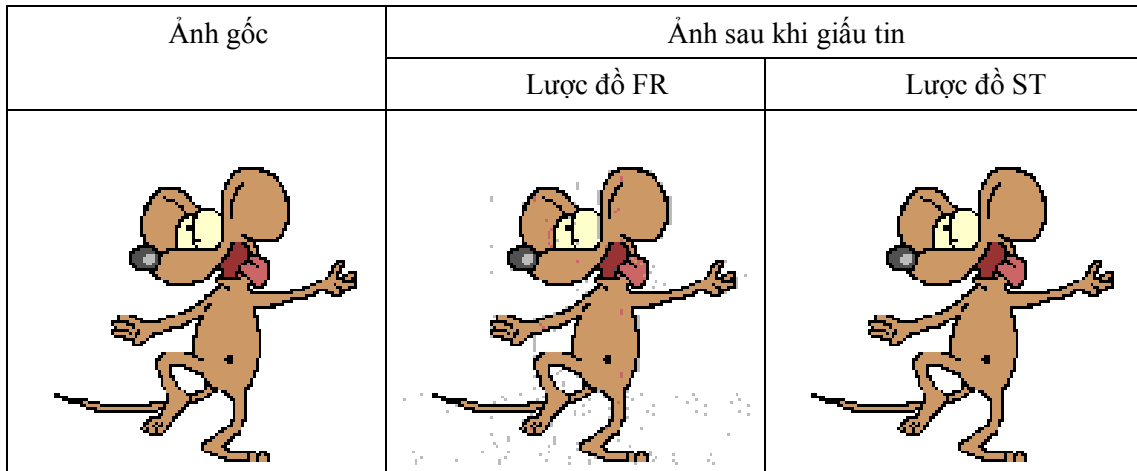
GIF (*Graphics Interchange Format*) ra đời vào năm 1980 bởi CompuServe như là một phương tiện hiệu quả để trao đổi hình ảnh trên mạng.

Vào đầu những năm 1990, GIF đã xuất hiện trên World Wide Web và nhanh chóng trở nên quen thuộc với người sử dụng. Gần như phần lớn các hình ảnh trên Web hiện nay là GIF, và hầu như mọi trình duyệt Web đều hỗ trợ định dạng GIF. GIF file là một hệ thống nén để kích thước file là nhỏ nhất, và nó có giới hạn là 8 bit màu (256 màu khác nhau). Một vài loại khác của GIF hỗ trợ màu trong suốt (*transparents*) và kết hợp nhiều hình ảnh trong một file tạo ảnh động (*animation*).

Chính vì những lý do thực tế nêu trên, chúng tôi đã lựa chọn thực hiện thuật toán giấu tin trên môi trường định dạng ảnh GIF. GIF là một định dạng ảnh có bảng màu nên hoàn toàn có thể áp dụng mọi ý tưởng được đề xuất ở trên cho quá trình giấu tin và lọc tin. Hệ thống nén của GIF là nén bảo toàn – LZW [4] (*Lempel Zev Welch*) cho nên có thể áp dụng kỹ thuật nén và giải nén kết hợp với quá trình thực hiện giấu tin và lọc tin. Dưới đây là một số kết quả thử nghiệm thu được qua đánh giá tỉ lệ PSNR [5] (*Peak Signal To Noise Ratio – thể hiện sự tương quan sai khác giữa ảnh gốc và ảnh chứa tin giấu, PSNR càng lớn thể hiện sự sai khác càng nhỏ*) áp dụng với hai thuật toán FR và ST trên cùng một ảnh gốc kích thước 120x120:

Kích thước dữ liệu ảnh	Kích thước tin giấu (byte)	Giá trị PSNR	
		Lược đồ FR	Lược đồ ST
120x120	100	46.17	71.02
120x120	128	45.55	70.25
120x120	256	40.98	67.41
120x120	400	30.46	65.47

Bảng 1. So sánh chất lượng ảnh trước và sau khi giấu tin

Hình 1. So sánh ảnh trước và sau khi giấu tin

V.KẾT LUẬN

Trong bài báo này, chúng tôi đề xuất một thuật toán giấu tin mật ST trên ảnh có bảng màu trên ý tưởng biến đổi bảng màu cho thỏa mãn tối ưu về cảm nhận sự thay đổi của ảnh.

Với thuật toán cơ sở nêu trên, chúng tôi xây dựng một lược đồ giấu tin mật trong ảnh áp dụng trong trao đổi thông tin mật. Hoàn toàn có thể áp dụng tốt thuật toán trên với ảnh đa cấp xám và các định dạng ảnh sử dụng bảng màu nói chung.

Chúng tôi tiến hành thử nghiệm thuật toán trên môi trường định dạng ảnh GIF phù hợp với xu hướng thực tế trong trao đổi thông tin. Thuật toán cho kết quả thử nghiệm rất tốt với các mẫu ảnh thu được từ trên mạng Internet. Sự thay đổi của ảnh sau khi giấu tin gần như không cảm nhận được bởi hệ thống thị giác của con người, cho thấy thuật toán có thể áp dụng trong thực tế.

Chúng tôi cũng đã nghiên cứu sử dụng hệ thống khóa bí mật, kết hợp với các phương pháp kiểm soát chất lượng ảnh khi giấu tin nhằm xây dựng lược đồ giấu tin mật trong ảnh đảm bảo tính mật, tính an toàn cao đối với tin giấu, đồng thời ảnh chứa tin giấu không có những thay đổi bất thường so với ảnh gốc. Vì khuôn khổ bài viết nên các nội dung về sử dụng kết hợp hệ thống khóa, đánh giá độ an toàn của tin giấu sẽ được trình bày trong các bài viết tiếp theo.

TÀI LIỆU THAM KHẢO

- [1].EZStego,<http://www.fqa.com/romana/>;
<http://www.stego.com/>
- [2].Fridrich J. A New Steganographic Method for Palette-Based Images. Proc. of the IS&T PICS conference, April 1998, Savannah, Georgia (1998) pp. 285–289.
- [3]. Nguyễn Văn Tào, *Một số thuật toán giấu tin và áp dụng giấu tin mật trong ảnh*. Tạp chí Khoa học và Công nghệ, Đại học Thái Nguyên, 4(2), tr. 25-32.
- [4].LZW, <http://en.wikipedia.org/wiki/Lempel-Ziv-Welch>
- [5].PSNR, http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio

Summary**A COLOR IMAGE DATA HIDING ALGORITHM APPLIED
IN THE GIF IMAGE DATA HIDING**

The applications of information communication are now fast increased. Therefore, information security is essential in many communication applications. There are many information security solutions were proposed, image data hiding is a solution attracting many scientists. This paper proposes a color image data hiding and introduces its application in GIF image.

Keyword: *Data Hidding, Security, Digital Image*