Bài 5.3 Thủy vân số và một số kỹ thuật

1.Một số ứng dụng cụ thể

- Bảo vệ bản quyền tác giả (copyright protection)

Một thông tin nào đó mang ý nghĩa xác định quyền sở hữu của tác giả (người ta gọi nó là thuỷ vân) sẽ được nhúng vào các sản phẩm dữ liệu đa phương tiện và chỉ duy nhất người chủ sở hữu hợp pháp các sản phẩm đó có thủy vân và được dùng làm minh chứng cho bản quyền sản phẩm. Giả sư, có một thành

phẩm dữ liệu dạng đa phương tiện như ảnh, âm thanh, video cần được lưu trên mạng. Việc bảo vệ các sản phẩm chống lại các hành vi lấy cắp hoặc làm nhái cần phải có một kỹ thuật để dán tem bản quyền vào sản phẩm này. Việc dán tem chính là việc "nhúng" thuỷ vân, cần phải đảm bảo không để lại một ảnh hưởng lớn nào đến việc cảm nhận sản phẩm. Yêu cầu kỹ thuật đối với ứng dụng này là thuỷ vân phải tồn tại bền vững cùng với sản phẩm, muốn hủy

bỏ thuỷ vân này mà không được phép của người chủ sở hữu thì chỉ có cách là phá huỷ sản phẩm.

- Xác thực thông tin hay phát hiện xuyên tạc thông tin (authentication and tamper detection)

Một tập các thông tin sẽ được giấu trong phương tiện chứa sau đó được sử dụng để nhận biết xem dữ liệu trên phương tiện gốc đó có bị thay đổi hay không. Các thuỷ vân nên được "ẩn" để tránh sự tò mò của đối

phương. Hơn nữa, việc làm giả các thuỷ vân hợp lệ hay xuyên tạc thông tin nguồn cũng cần được xem xét. Trong các ứng dụng thực tế, người ta mong muốn tìm được vị trí bị xuyên tạc cũng như phân biệt được các thay đổi (ví như phân biệt xem một đối tượng đa phương tiện chứa giấu thông tin đã bị thay đổi, xuyên tạc nội dung hay chỉ là bị nén mất dữ liệu). Yêu cầu chung đối với dữ liệu này là khả năng giấu được nhiều thông tin và thuỷ vân không cần bền

vững trước các phép xử lvý trên các đối tượng đã được giấu tin.

- Giấu vân tay hay dán nhãn (fingeprinting and labeling)

Thuỷ vân được sử dụng để nhận diện người gửi hay người nhận của một thông tin nào đó trong ứng dụng phân phối sản phẩm. Thuỷ vân trong trường hợp này cũng tương tự như số serial của sản phẩm phần mềm. Mỗi một sản phẩm sẽ mang một thuỷ vân riêng. Ví dụ như các thủy vân khác nhau sẽ được nhúng vào các bản copy khác nhau của

thông tin gốc trước khi chuyển cho người nhận. Với những ứng dụng này thì yêu cầu đảm bảo độ an toàn cao cho các thuỷ vân tránh sự xoá dấu vết trong khi phân phối.

- Kiểm soát sao chép (copy control)

Điều mong muốn đối với các hệ thống phân phối dữ liệu đa phương tiện là tồn tại một kỹ thuật chống sao chép trái phép dữ liệu. Có thể sử dụng thuỷ vân để chỉ trạng thái sao chép của dữ liệu. Các thuỷ vân trong

trường hợp này được sử dụng để kiểm soát sao chép đối với các thông tin. Các thiết bị phát hiện ra thuỷ vân thường được gắn sẵn vào trong các hệ thống đọc - ghi.

Ví dụ như hệ thống quản lý sao chép DVD đã được ứng dụng ở Nhật. Thuỷ vân mang các giá trị chỉ trạng thái cho phép sao chép dữ liệu như "không được sao chép" (copy never) hay "chỉ được sao chép một lần" (copy once), sau khi copy xong, bộ đọc - ghi thuỷ

vân sẽ ghi thuỷ vân mới chỉ trạng thái mới lên DVD. Các ứng dụng loại này cũng yêu cầu thuỷ vân phải được đảm bảo an toàn và cũng sử dụng được phương pháp phát hiện thuỷ vân đã giấu mà không cần thông tin gốc.

- Giấu tin mật (steganography) Các thông tin có thể giấu được trong những trường hợp này càng nhiều càng tốt sao cho vẫn đảm bảo yêu cầu là không thể phát hiện được. Việc giải mã để lấy được thông tin cũng không

cần phương tiện mang gốc ban đầu. Các yêu cầu về chống tấn công của đối phương không cần cao lắm, thay vào đó là yêu cầu thông tin giấu phải được bảo mật.

Kỹ thuật giấu thông tin đang được áp dụng cho nhiều loại đối tượng chứ không riêng gì dữ liệu đa phương tiện như ảnh, audio, video. Gần đây, đã có một số nghiên cứu giấu tin trong cơ sở dữ liệu quan hệ. Trong bài báo tiếp theo chúng tôi sẽ trình bày về kỹ thuật phát

hiện ảnh có chứa thông tin ẩn hay không...

Data hiding: Là thuật ngữ chỉ kỹ thuật giấu thông tin nói chung bao gồm cả giấu tin mật và thuỷ vân số.

Steganography: Chỉ kỹ thuật giấu thông tin mật trong một đối tượng.

Watermarking: Thuỷ vân số, chỉ những kỹ thuật giấu thông tin dùng để bảo vệ vật chứa thông tin giấu.

Phương tiện chứa (host signal): Là phương tiện gốc

được dùng để nhúng thông tin. Giấu thông tin trong ảnh thì nó mang tên ảnh chứa, còn trong audio là audio chứa… và ta cũng gọi phương tiện chứa là môi trường.

Thông tin giấu (embeded data): Là lượng thông tin được nhúng vào trong phương tiện chứa. Trong giấu thông tin mật Steganography, thông tin giấu được gọi là thông điệp giấu (message) còn trong kỹ thuật thuỷ vân số thì thông

tin giấu được gọi là thuỷ vân (watermark).

5. Các thành phần của Hệ "Ẩn – Giấu tin".

Các thành phần chính của một hệ **"Ẩn – Giấu tin"** trong ảnh gồm có:

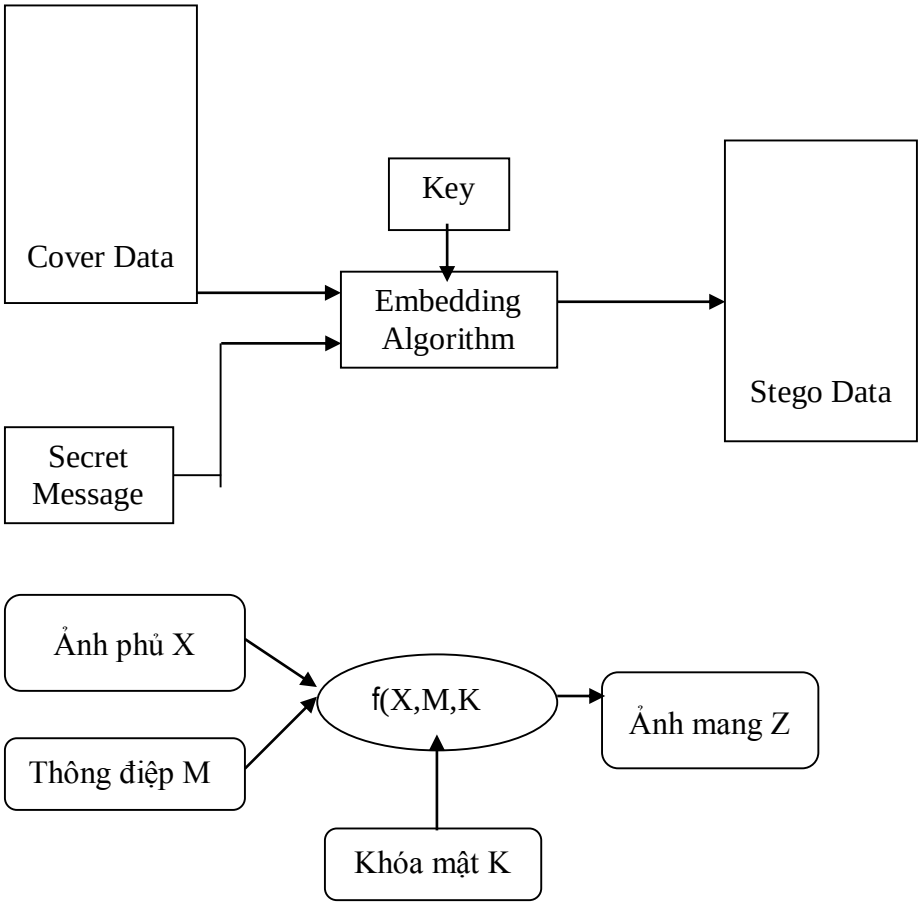
+ Bản tin mật: có thể là văn bản, hình ảnh hay tệp tin tùy ý (audio,

video,…), vì trong quá trình giấu tin, chúng đều được chuyển thành chuỗi các bít.

+ Môi trường sẽ chứa tin mật: Thường là ảnh, nên gọi là Ảnh phủ hay Ảnh gốc.

+ Khóa K: Khóa viết mật, tham gia vào quá trình giấu tin để tăng tính bảo mật.

+ Môi trường đã chứa tin mật: thường là ảnh, nên gọi là Ảnh mang, là ảnh sau khi đã nhúng tin mật vào.

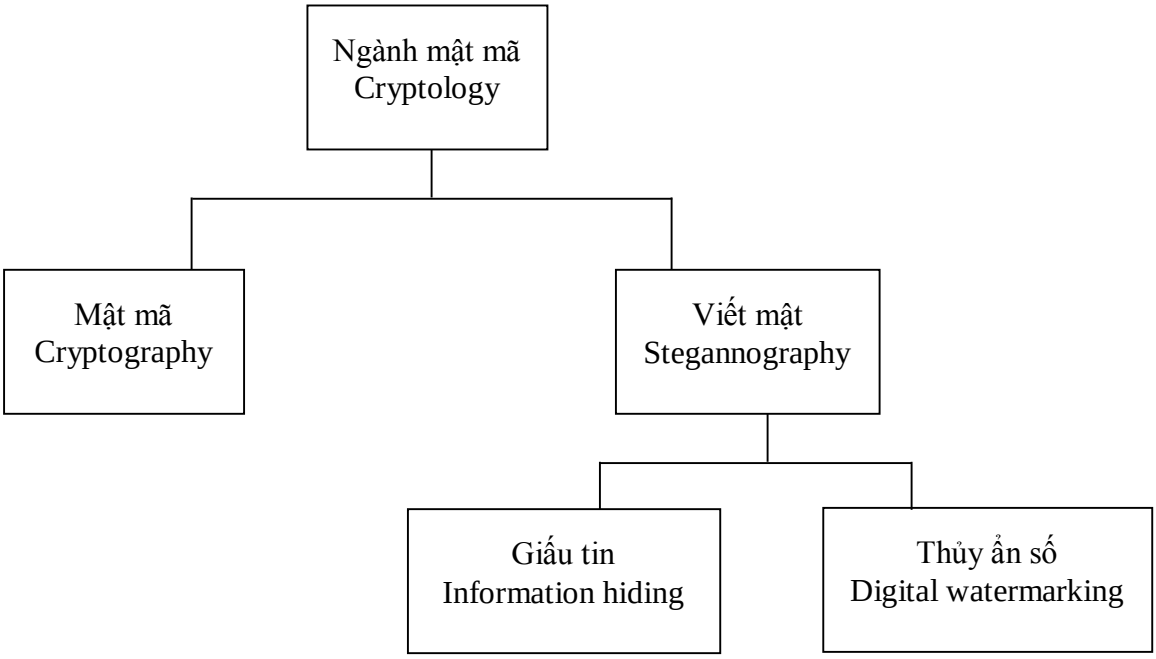


Hình 1: ***Sơ đồ "Ẩn – giấu tin" trong ảnh***

6. Ẩn – Giấu tin và Mật mã.

Có thể xem "Ẩn – giấn tin" là một ngành mật mã với mục

tiêu là nghiên cứu các phương pháp “che giấu” thông tin mật.



Hình 2: ***Các lĩnh vực nghiên cứu của Mật mã học***

"Ẩn – giấu tin" và **"Mã hóa"** tuy cùng có mục đích là để đối phương **"khó"** phát hiện ra tin cần giấu, tuy nhiên nó khác với mã hóa ở chỗ:

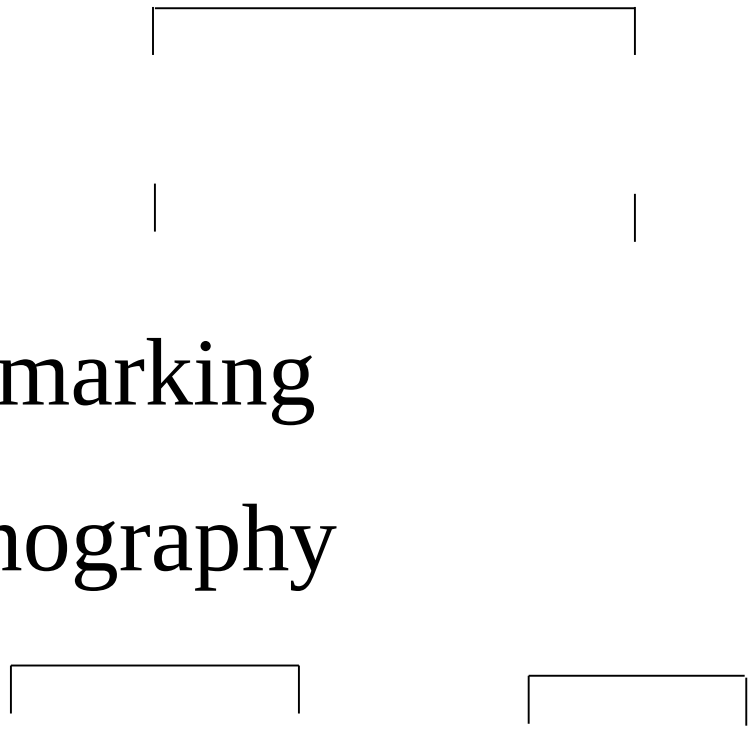
+ "Mã hóa" là giấu đi **"ý nghĩa"** của thông tin.

+ Ẩn – Giấu tin" là giấu đi **"sự hiện diện"** của thông tin.

Về bản chất, **"Ẩn – giấu tin"** gần với "Nén tin" hơn.

7. Phân loại "Ẩn – Giấu tin"

Trong lĩnh vực bảo mật thông tin, **"Ẩn – giấu tin"** bao gồm các vấn đề sau:



Watermarking

Fingerprinting Intrinsic

Pure

Hình 3: ***Các vấn đề nghiên cứu trong "Ẩn – giấu tin"***

1) Giấu tin (Steganography) là kỹ thuật nhúng **"mẩu tin mật"** (Mẩu tin cần giữ bí mật) vào **"môi trường giấu tin"** (môi trường phủ),

+ **“Giấu tin có xử lý”** (Intrinsic Steganogaraphy) là một dạng “Giấu tin”, trong đó để tăng bảo mật, có thể phải dùng khóa viết mật. Để giải mã, người ta cũng phải có khóa viết mật đó.

Khóa viết mật không phải dùng để mã hóa mẩu tin, nó có thể là khóa dùng để sinh ra “hàm

băm", phục vụ "rải tin mật" vào môi trường giấu tin.

+ "Giấu tin đơn thuần" (Pure Steganography) là một dạng **"Giấu tin",** trong đó không dùng khóa viết mật để giấu tin, tức là chỉ giấu tin đơn thuần vào môi trường giấu tin.

2). **"Thủy ẩn số"** (Watermarking) là kỹ thuật

nhúng **“dấu ẩn số”** (tin giấu) vào một “tài liệu số” (“sản phẩm số”), nhằm chứng thực (đánh dấu, xác thực) nguồn gốc hay chủ sở hữu của “tài liệu số” này.

Ví dụ: **“dấu ẩn số”** dùng để xác nhận bản quyền một **“tác phẩm số”.** Tạm gọi **“Thủy ẩn số”** là **“Ẩn tin”,** để phân biệt với **“Giấu tin”.**

+ “Dấu vân tay” (Fingerprinting) là một dạng **“Thủy ẩn số”,** trong đó **“dấu ấn”** (tin giấu) là một định danh duy nhất (ví dụ định danh người dùng).

3) So sánh “Ẩn tin” (Watermarking) và “Giấu tin” (Steganography).

Về mặt hình thức, **“Ẩn tin”** giống “Giấu tin” ở chỗ đều tìm

cách nhúng thông tin vào một môi trường.

Về mặt nội dung, **“Ẩn tin”** (thủy ẩn” có một số điểm khác so với **“Giấu tin”.**

* Về mục tiêu:

+ Mục tiêu của **“Ẩn tin”** là nhúng “mẩu tin” thường là biểu tượng, chữ ký, dấu nhỏ đặc trưng vào môi trường phủ, nhằm phục

vụ việc chứng thực bản quyền tài liệu.

Như vậy "mẩu tin" cần nhúng (để làm biểu tượng xác thực) không nhất thiết phải là bí mật, nhiều khi cần lộ ra cho mọi người biết để mà **"dè chừng" !**

+ **"Ẩn tin"** có thể vô hình hoặc hữu trên vật mang tin.

“Ẩn tin” tìm cách biến **“tin giấu”** thành một thuộc tính của vật mang tin.

Mục đích của **“Ẩn tin”** là bảo vệ môi trường giấu tin.

+ Mục tiêu của **“Giấu tin”,** là nhúng “mẩu tin” thường là bí mật, vào môi trường phủ, sau đó có thể lấy ra (tách lại) tin mật từ môi trường phủ.

+ “Giấu tin” không cho phép nhìn thấy (bằng mắt) **“tin giấu”** trên vật mang tin.

Mục đích của “Giấu tin” là bảo vệ tin được giấu.

* Về đánh giá hiệu quả: Theo tiêu chí hay chỉ tiêu nào?

+ Chỉ tiêu quan trọng nhất của **“Ẩn tin”** là tính bền vững của tin được giấu.

+ Chỉ tiêu quan trọng nhất của **“Giấu tin”** là độ bí mật của tin được giấu.

8. Các tính chất của **“Ẩn – Giấu tin”** trong Ảnh.

Hiện nay có nhiều phương pháp **“Ẩn – Giấu tin”** trong ảnh được nghiên cứu.

Để đánh giá chất lượng của một phương pháp **“Ẩn – Giấu tin”,** người ta dựa vào một số tiêu chí sau:

1). Bảo đảm tính “vô hình”.

“Ẩn – giấu tin” trong ảnh sẽ làm biến đổi ảnh mang tin. Tính **“vô hình”** thể hiện mức độ biến đổi ảnh mang. Phương pháp **“Ẩn – giấu tin”** tốt, sẽ làm cho thông

tin mật trở nên **“vô hình”** trên ảnh mang, người dùng khó thể nhận ra trong ảnh có ẩn chứa thông tin mật.

Chú ý rằng với **“Ẩn tin”** thì trong thực tế không phải khi nào cũng cố gắng để đạt được tính vô hình cao nhất, ví dụ trong truyền hình, người ta gắn hình ảnh mờ

gọi là **“thủy ẩn”** để bảo vệ bản quyền bản tin.

2). Khả năng chống giả mạo.

Mục đích của **“Giấu tin”** là để truyền đi thông tin mật. Nếu không thể do thám tin mật, thì kẻ địch cũng cố tìm cách làm sai lạc tin mật, làm giả mạo tin mật để gây bất lợi cho đối phương. Phương pháp **“Giấu tin”** tốt

phải đảm bảo tin mật không bị tấn công một cách chủ động trên cơ sở những hiểu biết về thuật toán nhúng tin và có ảnh mang (nhưng không biết khóa "**Giấu tin**").

Đối với "Ẩn tin" thì khả năng chống giả mạo là yêu cầu vô cùng quan trọng, vì có như vậy mới bảo vệ được bản quyền,

minh chứng tính pháp lý của sản phẩm.

3). Dung lượng giấu.

Dung lượng giấu được tính bằng tỷ lệ của lượng tin cần giấu so với kích thước ảnh mang tin. Các phương pháp đều cố gắng giấu được nhiều tin trong ảnh nhưng vẫn giữ được bí mật. Tuy nhiên trong thực tế người ta luôn

phải cân nhắc giữa dung lượng và các chỉ tiêu khác như vô hình, tính bền vững (ổn định).

Hình 4: ***Cân nhắc giữa chất lượng, dung lượng và tính bền vững***

4). Tính bền vững.

Sau khi **"Ẩn – giấu tin"** vào ảnh mang, bản thân ảnh có thể phải qua các biến đổi khác nhau

nhu lọc (tuyến tính, phi tuyến), thêm nhiễu, làm sắc nét, mờ nhạt, quay, nén mất dữ liệu,… Tính bền vững là thước đo **"sự nguyên vẹn"** của tin mật sau những biến đổi như vậy.

5). Độ phức tạp tính toán.

"Độ phức tạp" của thuật toán **"Ẩn – giấu tin"** và **"Giải tin"** (tách tin) cũng là chỉ tiêu quan

trọng để đánh giá một phương pháp **"Ẩn – giấu tin"** trong ảnh. Chỉ tiêu này cho chúng ta biết "tài nguyên" (thời gian và bộ nhớ) tốn bao nhiêu dùng cho một phương pháp **"Ẩn – giấu tin".**

Với chủ nhân "Ẩn – giấu tin" thì thời gian thực hiện phải "nhanh", nhưng với kẻ thám tin thì **"Tách tin"** phải là bài toán

“khó”. Ví dụ bài toán **“Tách tin”** từ “Thủy ẩn” để đánh dấu bản quyền cần phải là bài toán “khó”, thì mới chịu được sự tấn công của tin tặc nhằm phá hủy **“Thủy ẩn”.**

9. Vấn đề tấn công Hệ thống “Ẩn – giấu tin”.

Tấn công một hệ **“Ẩn – giấu tin”** được gọi là “Steganalsis).

Đó là các phương pháp để phát hiện, phá hủy, trích rút hay sửa đổi tin mật. Nghiên cứu các biện pháp của kẻ tấn công, sẽ hữu ích cho việc thiết kế một hệ **"Ẩn – giấu tin"** tốt.

Việc tấn công được coi là thành công hay không tùy theo ứng dụng. Đối với liên lạc bí mật, việc phát hiện và chứng

minh được một ảnh có chứa tin mật được coi là thành công. Đối với bảo vệ bản quyền hay chống giả mạo, thì việc tấn công được coi là thành công nếu không chỉ phát hiện ra **“thủy ẩn”,** mà còn phá hủy hay sửa đổi nó, nhưng không làm giảm chất lượng của ảnh mang.

Có điểm giống nhau giữa **"mã hóa"** và **"giấu tin"** là người ta giả thiết thám tin biết trước phương pháp mã hóa hay giấu tin. Như vậy việc thám tin theo một phương pháp cụ thể (mã hóa hay giấu tin) phụ thuộc vào **"khóa",** chứ không phải phụ thuộc vào độ phức tạp của

phương pháp này (Nguyên ký Kerkhoff: [4]).

Tương tự như thám mã trong mã hóa, các kỹ thuật thám tin trong giấu tin cũng được chia thành làm năm nhóm:

- Biết ảnh mang tin.
- Biết ảnh gốc và ảnh mang tin.

- Biết có tin giấu trong ảnh mang tin.
- Biết thuật toán giấu tin.
- Biết thuật toán trích (tách) tin mật.

Thám tin phát hiện **"thủy ẩn"** hay tin mật có thể thực hiện bằng cách phân tích vùng nhiễu quá mức trên ảnh. Tin tặc kinh nghiệm có thể nhận thấy các

vùng nhiễu này bằng mắt thường. Nếu biết được ảnh gốc thì việc thám tin còn đơn giản hơn nữa, vì khi đó có thể so sánh ảnh mang tin với ảnh gốc để tách nhiễu.

Nếu thám tin biết được có tin ẩn giấu, người ta có thể tạo ra các cặp ảnh gốc và ảnh mang để phân tích và xét xem liệu ảnh

đang tìm hiểu có mang dấu ấn của chữ ký hay tin mật không.

Việc phá tin mật có thể đơn giản hay phức tạp tùy thuộc vào phương pháp **"Ẩn – giấu tin".** Đối với phương pháp nhúng tin vào **"bit có trọng số thấp",** thì việc phá tin mật chỉ đơn thuần là thay đổi lại các bit này, như vậy

ảnh mang tin trở về trạng thái ban đầu.

Phá tin mật đối với phương pháp **“Ẩn tin”,** mà vẫn giữ nguyên ảnh mang là một việc khó. Vì mục tiêu của **“thủy ẩn”** là phải đạt được độ bền vững sao cho nếu có ai đó phá **“thủy ấn”,** thì cũng làm hỏng ngay cả ảnh gốc.

Thông thường người ta tìm cách áp dụng nhiều phép biến đổi ảnh với hy vọng rằng: tuy từng phép biến đổi không có tác dụng, nhưng tổ hợp của chúng có thể giúp cho việc phá hủy **"thủy ấn"** mà vẫn giữ được ảnh mang.

Nếu biết tin mật và ảnh mang tin, thì cơ hội phá tin mật sẽ cao hơn.

Nếu biết thuật toán "**Ấn giấu tin**", kẻ thám tin có thể dùng nó thử **"Ẩn giấu tin"** lên nhiều ảnh khác nhau, qua đó dùng phương pháp thống kê để tìm ra các quy luật gây nhiễu, cũng như dùng nó để kiểm thử xem một ảnh có mang tin mật hay không.

Việc thám tin khó nhất đó là sửa đổi tin trong ảnh mang và

suy ra được **“khóa viết mật”** (Stego-key) dùng để nhúng tin. Nếu biết khóa viết mật, kẻ thám tin có thể làm giả các tin khác giống như nó được gửi đi từ chính chủ.

Phương pháp thám tin để biết thuật toán **“Ẩn giấu tin”** và thuật toán “tách tin” hay được dùng trong các hệ thám tin.

Nhiều kỹ thuật thám tin trong **"Ẩn giấu tin"** được chuyển sang từ kỹ thuật thám tin (trong mã hóa).

10. Các ứng dụng của "Ấn – Giấu tin".

10.1. Liên lạc bí mật.

Bản mã của tin mật có thể gây ra sự chú ý của tin tặc, nhưng tin mật được giấu vào trong môi

trường nào đó, rồi gửi đi trên mạng máy tính, thì ít gây ra sự chú ý của tin tặc. Đó là một ứng dụng của "Giấu tin".

Hiện nay người ta phối hợp đồng thời nhiều giải pháp để truyền tin mật trên mạng công khai: Đầu tiên tin mật được nén tin, sau đó mã hóa bản tin nén,

cuối cùng giấu bản mã vào trong môi trường nào đó.

10.2. Bảo vệ bản quyền.

1). “Thủy ấn” (Watermark):

+ Một biểu tượng bí mật gọi là **“Thủy ấn”** (Watermark) được **“nhúng”** vào trong một tài liệu (hình ảnh, âm thanh,…) để xác nhận quyền sở hữu về tài liệu.

+ **“Thủy ấn”** được đính lên tranh ảnh khi bán hoặc phân phối, thêm vào đó có thể gán một **“nhãn thời gian”** (Time stamp) để chống giả mạo.

+ **“Thủy ấn”** cũng được dùng để phát hiện xem các ảnh có bị sửa đổi hay không. Việc phát hiện **“Thủy ấn”** được thực hiện bằng thống kê, so sánh độ tương quan

hoặc bằng cách đo đạc xác định chất lượng của "Thủy ấn" trong ảnh mang. [5].

2). "Điểm chỉ số": Điểm chỉ số tương tự như số Seri của phần mềm [6].

"Điểm chỉ số" dùng để chuyển thông tin về người nhận **"sản phẩm số"** (không phải chủ

sở hữu), nhằm chứng thực bản sao duy nhất của sản phẩm.

3). Gán nhãn":

Tiêu đề, chú giải, nhãn thời gian… có thể được **"nhúng"** vào **"sản phẩm số"**. Gắn tên người lên ảnh của họ, gắn tên địa phương lên bản đồ. Khi đó nếu sao chép ảnh thì cũng sẽ sao chép cả thông tin đã "nhúng"

vào nó. Chủ sở hữu của sản phẩm, người có **"khóa viết mật"** (Stego-Key) có thể tách ra và xem các chú giải.

Trong một cơ sở dữ liệu ảnh, người ta có thể **"nhúng"** các từ khóa, đề các động cơ tìm kiếm có thể tìm nhanh một bức ảnh nào đó. Nếu là một khung ảnh cho cả một đoạn phim, người ta

có thể gán cả thời điểm diễn ra sự kiện (timing) đề đồng bộ hình ảnh với âm thanh. Người ta cũng có thể gán số lần mà hình ảnh được xem, để tính tiền thanh toán.

11. Một số chương trình "Ẩn – Giấu tin".

1). Chương trình Hide and Seek v4.1

Chương trình này của Colin Maroney, chạy dưới hệ điều hành DOS, để giấu tin vào các ảnh GIF. Nó thực hiện giấu tin vào ảnh mang một cách ngẫu nhiên, do đó nếu lượng tin cần giấu nhỏ thì tin sẽ được rải đều khắp ảnh mang. Nếu lượng tin

nhiều, thì các vùng thay đổi dày hơn, vì vậy dễ bị phát hiện.

2). Chương trình StegoDos

Chương trình chạy dưới hệ điều hành DOS, sử dụng ảnh mang 320 x 200 điểm ảnh và 256 mầu.

3). Chương trình White Noise Storm

Chương trình này của Ray (Arsen) Archelian, dễ dùng hơn và nhúng được nhiều tin hơn các chương trình trước. Ảnh mang không cần có kích thước cố định, tính vô hình cao.

4). Chương trình S-Tools for Windows

Một chương trình giấu ảnh tốt. Có thể giấu tin trong ảnh BMP,

GIF, tệp âm thanh WAV, các vùng chưa dùng đến của đĩa mềm. Giao diện đồ họa kéo thả. Để giấu tin chỉ cần kéo biểu tượng tệp tin cần giấu và thả lên ảnh.

Một yếu tố khác mà các hệ giấu tin nhắm tới và khai thác, đó là những điểm yếu trong hệ thống thị giác con người. Một

trong những phương pháp giấu tin là tạo ra các mặt nạ giác quan để đánh lừa mắt người. Vậy nên các nghiên cứu về phương pháp giấu tin trong ảnh có liên quan mật thiết với lĩnh vực xử lý ảnh, lý thuyết mật mã và các kiến thức về hệ thống thị giác.

Bài 2. PHƯƠNG PHÁP GIẤU TIN TRONG ẢNH

1. Giấu tin trong ảnh đen trắng.

Ảnh đen trắng số được thể hiện là một ma trận điểm ảnh gồm số 0 hay 1.

Giấu tin trong ảnh đen trắng là việc khó khăn, vì dễ bị nhận biết mắt thường. Số lượng tin giấu là hạn chế. Tuy nhiên ảnh đen trắng ngày càng ít được dùng, do đó việc nghiên cứu giấu tin trong loại ảnh này là ít thực tế.

1.1. Thuật toán giấu tin sử dụng tính chẵn lẻ của tổng số bit 1.

1). Ý tưởng:

Chia ảnh mang thành các khối nhỏ. Mỗi khối nhỏ sẽ được **"gài"** 1 bit của tin cần giấu. Dựa vào tính **"chẵn lẻ"** của tổng số các bit 1 trong khối để quy định giấu bit 1 hay bit 0. Cụ thể là sau khi

giấu, thì tổng số bit 1 trong khối và bit cần giấu sẽ có cùng tính **“chẵn lẻ”.**

2). Thuật toán “giấu tin”:

Input: FF là File ảnh Bitmap đen trắng (sẽ mang tin giấu): Fb là File tin cần giấu; K là khóa bí mật, đó là kích thước của khối nhỏ sẽ được tách từ FF.

Output: FF là File ảnh đã được giấu File tin mật Fb.

Bước 1: Tiền xử lý.

+ Chuyển File tin cần giấu Fb sang xâu nhị phân b.

+ Đọc Header của ảnh, đọc bảng màu, để lấy thông tin về ảnh.

Đọc phần dữ liệu ảnh vào mảng 2 chiều (ma trận) F.

Bước 2: Giấu tin.

Input: F là ma trận ảnh mang; b là dãy bit bí mật cần giấu;

K là khóa bí mật, đó là kích thước của khối nhỏ (được xác định trước).

Output: F là ma trận ảnh đã được giấu dãy bit bí mật b.

B1: Chia ảnh mang F thành các khối nhỏ với kích thước K.

B2: Theo một thứ tự xác định trước, xét từng khối nhỏ;

+ Nếu muốn giấu bit 1 vào một khối thì phải thỏa mãn điều kiện:

(L): Tổng số các bit 1 trong khối đó là số **"lẻ"** (tức là cùng tính **"lẻ"** của 1).

+ Nếu muốn giấu bit 1 vào một khối, nhưng không thỏa

mãn điều kiện (L), thì trong khối đó chọn ngẫu nhiên một bit và thay đổi giá trị của nó (từ 0 đổi sang 1 hay từ 1 đổi sang 0). Bằng cách đó, khối mang tin sẽ thỏa mãn điều kiện (L).

+ Nếu muốn giấu bit 0 vào một khối thì phải thỏa mãn điều kiện:

(C): Tổng số các bit 1 trong khối đó là số **"chẵn"** (tức là cùng tính **"chẵn"** của 0).

+ Nếu muốn giấu bít 0 vào một khối, nhưng không thỏa mãn điều kiện (C), thì trong khối đó chọn ngẫu nhiên một bit và thay đổi giá trị của nó (từ 0 đổi sang 1 hay từ 1 đổi sang 0). Bằng cách

đó, khối mang tin sẽ thỏa mãn điều kiện (C).

+ Mã hóa: Dùng khóa K và ảnh mang đã được xử lý để mã hóa theo theo phép XOR 2 ma trận.

3). Thuật toán tách "tin giấu":

Input: F là ảnh được giấu dãy bit bí mật b.

K là khóa bí mật, đó kích thước của khối nhỏ (được xác định trước).

Output: F là ảnh mang (ảnh trước khi giấu tin mật), b là dãy bit bí mật đã được giấu.

B0: Đọc Header của ảnh, đọc bảng màu, để lấy thông tin về ảnh.

Đọc phần dữ liệu ảnh vào mảng 2 chiều (ma trận) F.

B1: Chia ảnh F mang thành các khối nhỏ với kích thước K.

B2: Theo một thứ tự xác định trước, xét từng khối nhỏ:

Dùng khóa K để giải ra ảnh mang chưa được mã hóa, sau đó ta tách ảnh mang và dãy bit K theo cách:

Nếu tổng số bit 1 là "lẻ" thì ta thu được bit giấu là 1.

Nếu tổng số bit 1 là "chẵn" thì ta thu được bit giấu là 0.

Chú ý:

Độ an toàn của thuật toán không cao, vì chỉ cần biết khóa K, tức là kích thước các khối giấu tin là có thể dễ dàng tách được tin mật.

3) Ví dụ: Giấu bit 1 vào khối V sau:

Vì V có 6 bit 1, nên giấu bit 1 vào V, ta phải chọn ngẫu nhiên 1 bit và đổi giá trị.

1	0	1	1
0	0	1	0
1	1	0	0
0	0	0	0

Kết quả ta có thể giấu tin V

1	0	**0**	1
0	0	1	0
1	1	0	0
0	0	0	0

Bài tập

Giả sử ta có các file ảnh như sau:

1	1	0	0	1	0	1	1	1

0	1	0	1	1	1	0	1	1
1	0	1	0	0	1	1	1	0
0	0	0	1	1	0	1	0	1

Khóa K – cỡ của khối 2 chiều.

b=101

.2. Thuật toán giấu tin M.Y.Wu – J.H.Lee

1). Ý tưởng:

Thuật toán giấu tin kinh điển trong ảnh đen trắng của M.Y.Wu – J.H.Lee [6]

Với mục tiêu là giấu được càng nhiều tin vào trong ảnh càng tốt. Ý tưởng chính của thuật toán là chia ảnh ra thành các khối bằng nhau, tìm khối nào ít bị phát hiện nhất, (tức là vùng thứ yếu trên ảnh), giấu 1 bit thông tin vào khối đó.

2). Thuật toán:

Input: F là ảnh mang; b là bit bí mật cần giấu: K là khóa bí mật (ma trận m x n).

Output: F là ảnh đã được giấu bit b bí mật.

Ký hiệu SUM(F) là số các số 1 có trong ma trận F.

B1: Chia F thành các khối nhỏ F_1 có kích thước m x n (như ma trận K).

B2: Với mỗi khối nhỏ F_1 , kiểm tra điều kiện: $0 < SUM(F_1 \wedge K) < SUM(K)$.

Nếu đúng thì sang B3, để giấu b vào F_1 ; Nếu không đúng thì giữ nguyên F_1 .

B3: Giả sử bit cần giấu vào khối F_1 là b. Thay đổi F_1 như sau:

If $(SUM(F_1 \wedge K) \bmod 2 = b)$ then giữ nguyên F_1

Else if (SUM(F_1 ^ K) ≠b) then

Chọn ngẫu nhiên 1 bit thỏa mãn $[F_1]_{jk} = 1$, lật $[F1]_{jk}$ thành 0;

Else chọn ngẫu nhiên 1 bit thỏa mãn $[K]_{jk} = 1$,

Lật bit $[F_1]_{jk}$ từ 0 thành 1 hay từ 1 thành 0;

Chú ý:

Nếu m và n đủ lớn thì sự thay đổi trên ảnh mang không dễ gì bị

phát hiện ra bằng mắt thường. Có một số hướng cải tiến cho các thuật toán trên nhằm mục đích giấu được nhiều bit hơn vào khối ảnh.

3. Giấu tin trong ảnh màu.

Có nhiều phương pháp giấu tin trong ảnh màu hơn ảnh đen trắng. Tin giấu trong ảnh mầu

khó bị phát hiện hơn trong ảnh đen trắng.

Ảnh mầu "số" là một mảng các số thể hiện cường độ sáng tại mỗi điểm (pixel). Một ảnh 640 x 480 pixels, sử dụng 256 mầu (8 bits cho một điểm ảnh) là phổ biến. Một ảnh như vậy có thể chứa chừng 300 kilobits dữ liệu.

2.1. Các yêu cầu kỹ thuật.

Các kỹ thuật giấu tin trong ảnh phải đáp ứng các yêu cầu sau:

1/. Chất lượng ảnh mang vẫn đảm bảo, tin giấu không nhìn được bằng mắt thường.

2/. Tin giấu phải được mã hóa trực tiếp vào ảnh mang, chứ không vào phần khác, như vậy

mới giữ được cho nhiều dạng tệp ảnh khác nhau.

3/. Tin giấu phải bền vững với các sửa đổi và tấn công từ bên ngoài. Ví dụ, nhiễu trên đường truyền, lọc, lấy mẫu, cắt xén, mã hóa, nén dữ liệu, in, quét, biến đổi số sang "tương tự" và ngược lại, tác động đến tin giấu là ít nhất.

4/. Đảm bảo toàn vẹn dữ liệu, vì điều khó tránh khỏi là tin giấu cũng sẽ bị thay đổi, nếu biến đổi ảnh mang.

5/. Chú ý các phương pháp giấu tin cho phép phục hồi tin giấu, không cần ảnh gốc.

10.2.2.2. Phương pháp giấu tin trong ảnh mầu.

a). Phân theo nhóm phương pháp giấu tin theo "kỹ thuật".

Theo phương pháp này, các phương pháp giấu tin trong ảnh hiện nay đều thuộc một trong ba nhóm sau:

1/. Giấu tin mật vào các "bit có trọng số thấp" (LSB: Least Significant Bit)

Nhóm phương pháp nhúng tin giấu vào các “bit có trọng số thấp” của ảnh hay được áp dụng trên các ảnh Bitmap không nén và các ảnh dùng bảng mầu (như GIF, TIF). Ý tưởng chính của phương pháp này là lấy từng bit của mầu tin mật, rải nó lên ảnh mang, gài vào các bit có trọng số thấp.

2/. Giấu tin dựa vào kỹ thuật “biến đổi ảnh”.

Nhóm phương pháp dựa vào kỹ thuật **“biến đổi ảnh”,** lợi dụng việc biến đổi ảnh từ miền biểu diễn này sang miền biểu diễn khác, để giấu bit tin mật. Ví dụ biến đổi miền không gian sang miền tần số.

Một ví dụ về hệ thống dùng phương pháp này là "Jpeg-Jsteg" [4]. Hệ thống này nhúng tin bằng cách điều chế các hệ số của phép biến đổi "Cosin rời rạc", theo các bit tin cần giấu và sự làm tròn lỗi khi lượng hóa.

Một số phương pháp khác thuộc nhóm này sử dụng ảnh như mô hình vật lý với các dải phổ

thể hiện mức năng lượng. Khi đó giấu tin giống như việc điều chế một tín hiệu dài hẹp vào một dải tần rộng (ảnh phủ, ảnh mang).

3/. Giấu tin sử dụng "mặt nạ" giác quan.

Nhóm phương pháp dùng **"mặt nạ"** giác quan, dựa trên nguyên lý **"đánh lửa"** hệ thống giác quan con người. Một số

điểm yếu của hệ thống giác quan là:

+ Hiệu ứng **“mặt nạ”** của các cạnh.

+ Sự nhậy cảm đối với độ tương phản là một hàm của miền tần số.

+ Khả năng nhạy cảm kém đối với các thay đổi nhỏ trong độ

chói trên các mảng ảnh có cấu tạo ngẫu nhiên.

+ Sự nhạy cảm kém đối với các tần số miền không gian thấp, ví dụ như sự thay đổi liên tục của độ sáng trên ảnh.

“Mặt nạ” ở đây ám chỉ hiện tượng mắt người “không cảm nhận” được một tín hiệu, nếu nó

ở bên cạnh một tín hiệu nhất định nào đó.

b). Phân nhóm phương pháp giấu tin theo **"định dạng ảnh".**

1/. Nhóm phương pháp "phụ thuộc định dạng ảnh".

Hạn chế của nhóm phương pháp này là thông tin giấu dễ bị "tổn thương" bởi các phép biến đổi ảnh.

Trong nhóm này lại chia ra theo dạng ảnh, có các phương pháp cho:

Ảnh dựa vào bảng mầu. Ảnh JPEG.

2/. Nhóm phương pháp "độc lập với định dạng ảnh".

Đặc trưng của nhóm phương pháp này là lợi dụng vào việc biến đổi ảnh để giấu tin vào

trong đó, ví dụ giấu vào các hệ số biến đổi. Như vậy có bao nhiêu phép biến đổi ảnh thì cũng có thể có bấy nhiêu phương pháp giấu ảnh.

Một số phép "biến đổi ảnh":

+ Phương pháp biến đổi theo miền không gian (Spacial domain).

+ Phương pháp biến đổi theo miền tần số (DCT, DFT, Wavelet).

+ Phương pháp biến đổi hình học.

Phương pháp nhóm thứ hai có nhiều ưu điểm hơn về tính bền vững, tuy nhiên lượng thông tin giấu được sẽ ít hơn và cài đặt cũng sẽ phức tạp hơn.

c). Phân nhóm phương pháp giấu tin theo "đặc điểm kỹ thuật".

1/. Phương pháp thay thế.

+ Thay thế các bit dữ liệu trong bản đồ bit (bit plane).

+ Thay thế bảng mầu (palette).

2/. Phương pháp xử lý tín hiệu.

+ Các phương pháp biến đổi ảnh (Transform).

+ Các kỹ thuật điều chế dải phổ.

3/. Phương pháp mã hóa (coding).

+ Lượng hóa, dithering.

+ Mã hóa sửa lỗi.

4/. Phương pháp thống kê – kiểm thử giả thuyết.

5/. Phương pháp sinh "mặt nạ" – Fractal.

Giấu tin sử dụng tính chẵn lẻ của tổng số bit 1.

Đối với ảnh mầu hay ảnh đa cấp xám, cũng có thể áp dụng thuật toán "giấu tin sử dụng tính chẵn lẻ của tổng số bit 1" cho ảnh đen trắng (Xem 5.2.1).

Với các loại ảnh này, mỗi điểm ảnh được biểu diễn bằng nhiều bit, trong đó có những bit ít quan trọng (LSB: Least Significant bit). Từ mỗi điểm

ảnh, ta chọn ra một bit LSB, lưu nó vào ma trận 2 chiều F gồm các bit 0, 1.

Trên ma trận F ta áp dụng thuật toán "Giấu tin sử dụng tính chẵn lẻ của tổng số bit 1" cho ảnh đen trắng.

4. Giấu tin vào các bit có trọng số thấp (LSB).

10.2.4.1. Cơ sở kỹ thuật.

Ý tưởng chính của phương pháp nhúng tin giấu vào các bit có trọng số thấp (LSB). Khi chuyển ảnh tương tự sang ảnh số, người ta chọn 3 cách thể hiện mầu:

* 24-bit mầu:

Mỗi điểm ảnh có thể nhận một trong 2 ^ 24 mầu, mỗi mầu được tạo từ 3 mầu cơ bản:

Red ®, green (G), blue (B), mỗi mầu nhận một giá trị từ 0 đến 255 (8 bit).

* 8-bit mầu:

Mỗi điểm ảnh có thể nhận một trong 256 mầu, chọn từ một bảng mầu (Palette).

* 8-bit dải xám:

Mỗi điểm ảnh có thể nhận một trong 256 (2^8) sắc thái xám.

Phương pháp LSB sửa bit hay các bit có trọng số thấp nhất (ít quan trọng nhất để tạo nên mầu điểm ảnh), gài các thông tin mật vào đó. Các thông tin được giấu sẽ lẫn vào đâu đó giống như nhiễu ảnh.

Áp dụng kỹ thuật LSB, một điểm ảnh 24-bit có thể giấu được ba bit thông tin (vì mỗi điểm

được thể hiện bằng ba byte). Mọi sự thay đổi trên điểm ảnh có trọng số thấp đều không gây nên sự chú ý của mắt người.

Hình 5: ***Giấu tin vào các bit ít quan trọng của điểm ảnh***

4.2. Ví dụ về phương pháp LSB

Chữ cái A có mã ASCII là 65 hệ thập phân (0100 0001 hệ nhị phân).

Ví dụ các điểm ảnh trước khi giấu là:

00100111 11101001 11001000

00100111 11001000 11101001

11001000 00100111 11101001

Để giấu chữ A cần ba điểm ảnh liên tiếp.

Chèn giá trị nhị phân của chữ A vào ba điểm ảnh trên bắt đầu từ byte trên cùng bên trái, sẽ cho kết quả:

00100110 11101000 11001001

00100110 11001000 11101000

11001000 00100110 11101001

Các bit được gạch chân là các bit bị lật. Có thể dùng hai bit có trọng số thấp để giấu tin mà chất lượng không thay đổi mấy đối với mắt thường.

Từ ví dụ trên ta có thể suy ra rằng nếu dùng 1 LSB thì xác suất phải lật bit là 50%, vậy nên lượng nhiễu gây ra cho ảnh là rất ít.

Đối với ảnh mầu 24 bit, đôi khi chúng ta có thể dùng đến 2 hoặc thậm chí 3 bit thấp mà vẫn không để lộ thông tin mật.

Đối với ảnh 8 bit thì điều này là không thể, và người ta chỉ dùng 1 bit thấp nhất để giấu tin.

4.3. Dung lượng tin giấu.

Phương pháp LSB giấu được nhiều thông tin.

Với ảnh 24 bit / điểm ảnh, dùng một bit có trọng số thấp có thể giấu được:

3 bit ẩn / 1 điểm ảnh (24 bit dữ liệu) = 1/8 bit ẩn / bit dữ liệu

Nếu dùng 2 bit có trọng số thấp

6 bit ẩn / 1 điểm ảnh (24 bit dữ liệu) = 1/4 bit ẩn / bit dữ liệu

Trong các ảnh sặc sỡ chúng ta có thể dùng thậm chí 3 bit LSB, khi đó thu được tỷ lệ bit ẩn / bit dữ liệu là 3/8.

Đôi khi người ta hỏi ngược lại là cần bao nhiêu byte ảnh để có thể giấu 1 byte tin mật. Nếu chỉ dùng 1 bit thấp ta cần 8 byte, nếu dùng đến 2 bit, ta chỉ cần 4

byte dữ liệu là đã giấu được 1 byte thông tin.

Nếu áp dụng kỹ thuật LSB lên ảnh 8-bit, cần phải chú ý hơn vì ảnh 8-bit không dễ chấp nhận thay đổi như ảnh 24-bit. Nên tránh dùng các ảnh vẽ phức tạp (như Mona Lisa). Các ảnh đơn giản như ảnh động vật, ví dụ chó, mèo, mèo phù hợp hơn. Khi sửa

bit trọng số thấp trong ảnh 8-bit, các con trỏ chỉ đến bảng mầu cũng bị thay đổi theo. Chú ý rằng đôi khi chỉ cần thay đổi 1 bit có thể dẫn đến sự khác biệt về dài Red và dài Blue. Các thay đổi như vậy sẽ bị nhận ra ngay. Vì vậy các chuyên gia về giấu tin trong ảnh khuyên nên dùng bảng

mầu xám vì sự khác biệt giữa các cấp màu không dễ thấy.

4.4. Tính bền vững.

Phương pháp LSB rất dễ bị **"tổn thương"** bởi một loạt các phép biến đổi ảnh, ngay cả phép biến đổi ảnh đơn giản và thông dụng nhất.

Nén ảnh mất dữ liệu như JPEG rất dễ dàng phá hủy tin

mật. Vấn đề là ở chỗ, những **"lỗ hổng"** trong hệ thống thị giác con người – ít nhạy cảm với các nhiễu bổ sung – mà phương pháp chèn bit LSB khai thác lại cũng chính là yếu tố mà phương pháp nén mất dữ liệu dựa lên đó để giảm mức dữ liệu của một ảnh.

Các phép biến đổi hình học như dịch chuyển hay xoay cũng

dễ làm mất dữ liệu mật vì khi đó vị trí của các bit giấu sẽ bị thay đổi. Chỉ có một phép dịch chuyển đơn giản thì mới có thể phục hồi lại dữ liệu mật.

Các phép xử lý ảnh khác như làm mờ ảnh cũng sẽ làm mất dữ liệu hoàn toàn.

Tóm lại phương pháp LSB là phương pháp có tính ổn định kém nhất.

BÀI TẬP CHƯƠNG X. ẨN GIẤU TIN.

Để hiểu cách thức giấu tin và tách tin đã giấu từ môi trường giấu tin với từng phương pháp

cụ thể, bài tập chương VIII tập trung vào việc lập chương trình giấu tin và tách tin đã giấu.

Bài tập

Viết chương trình thực hiện Giấu tin sau:

1/. Thuật toán giấu tin sử dụng tính chẵn lẻ của tổng số bit 1.

2/. Thuật toán giấu tin M.Y.Wu – J.H.Lee

3/. Giấu tin vào các bit có trọng số thấp (LSB).

Mẫu chương trình

* Mỗi chương trình giấu tin thực hiện các công việc theo thực đơn sau:

Thực đơn chính.

G. Giấu tin.

T. Tách tin.

K. Kết thúc.

Bài tập nhanh

Hãy giấu bản tin x=11010 theo tính chẵn lẻ
Vào 5 khối sau:

1	0	1	1		1	0	1	1		1	0	1	1		
0	0	1	0		0	0	1	0		0	0	1	0		
1	1	0	0		1	1	1	0		1	1	0	1		
0	1	0	0		0	0	1	1		0	0	0	1		
1	0	1	1		1	0	1	1							
1	1	1	0		0	0	1	0							
1	1	0	0		1	1	1	0							
0	0	0	1		0	1	0	0							

Ma trận K

1	1	1	1		1	0	1	1		1	0	1	1		
0	0	1	0		0	0	1	0		0	0	1	0		
1	1	0	0		1	1	1	0		1	1	0	1		
0	1	0	0		0	1	1	0		1	0	1	1		
1	0	1	1		1	0	1	1							
1	1	1	1		1	0	1	0							
1	1	0	0		1	1	1	0							
0	0	1	1		0	1	0	0							

Khối ảnh F_1				Khóa K_1			
1	0	1	1	1	1	1	1
0	0	1	0	0	0	1	0

1 1 0 0 0 1 0 0	1 1 0 0 0 1 0 0
$F_1 \wedge K_1$	
1 0 1 1 0 0 1 0 1 1 0 0 0 1 0 0	$SUM(F_1 \wedge K_1)=7$ $<SUM(K_1)=8$