3

Session

Steganography & Steganalysis

Ẩn giấu thông tin
và phương pháp điều tra

Tài liệu Thực hành

Pháp chứng Kỹ thuật số

GVTH: ThS. Phan Thế Duy

Học kỳ II – Năm học 2018-2019

Tp. HCM, 3.2019

Lưu hành nội bộ

A. TỔNG QUAN

1. Mục tiêu

Bài thực hành này giúp sinh viên được làm quen, sử dụng, tăng cường kiến thức về các kỹ năng điều tra kỹ thuật số liên quan đến việc phân tích đĩa cứng máy tính, ẩn giấu dữ liệu (steganography).

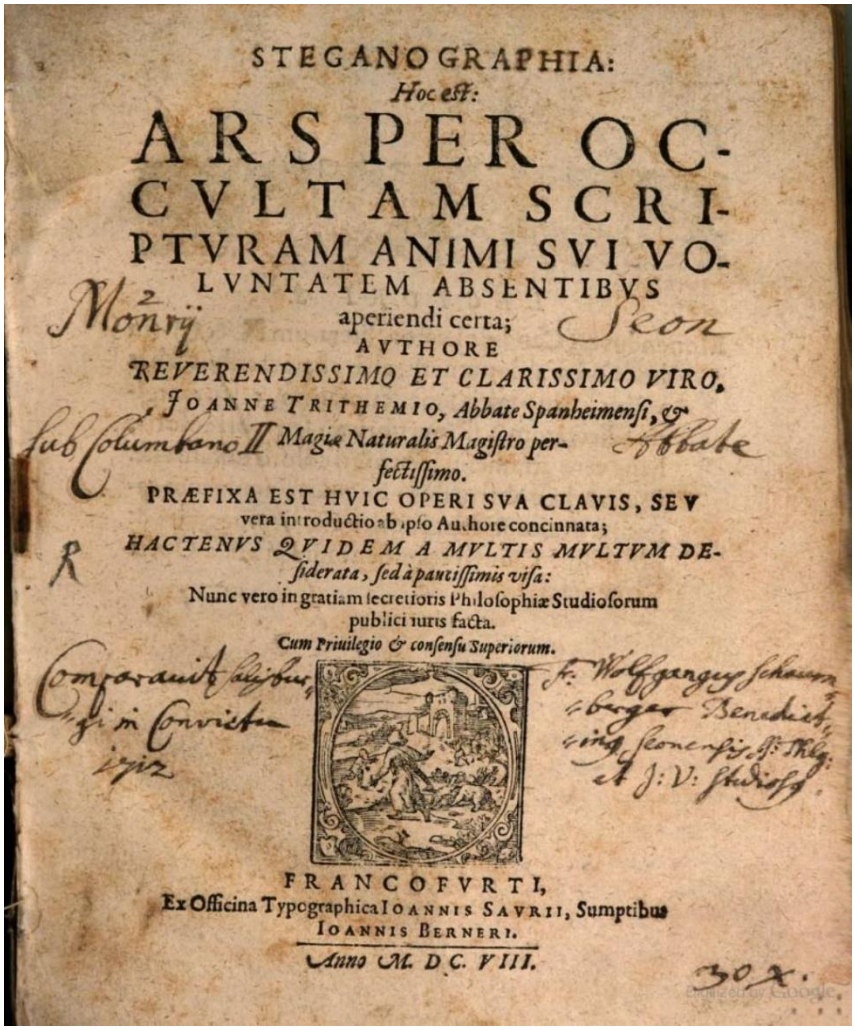
2. Giới thiệu kỹ thuật ẩn giấu thông tin

Ẩn giấu thông tin (Steganography)

Kỹ thuật ẩn giấu thông tin (tiếng Anh: steganography) là một hình thức che dấu sự tồn tại của thông điệp. Khác với Cryptography thì Steganography ẩn đi các thông tin cần giữ bí mật trong các dữ liệu vô hại để đối phương không thể biết được sự hiện diện của thông điệp.

Kỹ thuật giấu tin là nghệ thuật và khoa học về việc viết và chuyển tải các thông điệp một cách bí mật, sao cho ngoại trừ người gửi và người nhận, không ai biết đến sự tồn tại của thông điệp, là một dạng của bảo mật bằng cách che giấu. Từ steganography có gốc Hy Lạp có nghĩa là "giấu tin" kết hợp từ hai từ steganos (στεγανός) nghĩa là "ẩn nấp để bảo vệ" và graphein (γράφειν) nghĩa là "viết". Thuật ngữ steganography được Johannes Trithemius sử dụng lần đầu năm 1499 trong cuốn sách Steganographia, một luận văn về mật mã và kỹ thuật giấu tin được ngụy trang dưới dạng một cuốn sách ma thuật. Trong kỹ thuật giấu tin, thông thường thông điệp xuất hiện dưới một dạng khác trong quá trình truyền tải: hình ảnh, bài báo, danh sách mua hàng, bì thư, hoặc thông điệp ẩn có thể được viết bằng mực vô hình giữa các khoảng trống trong một lá thư bình thường.

Thông thường, Steganography được thực hiện bằng cách thay các bit vô ích hoặc không sử dụng trong tập tin bằng các bit khác (chính là các bit của dữ liệu cần che dấu). Thông tin ẩn dấu có thể là plaintext (dữ liệu gốc chưa mã hóa), ciphertext (dữ liệu đã mã hóa). Steganography thường được sử dụng khi việc mã hóa dữ liệu bị cấm. Hoặc, phổ biến hơn, người ta dùng steganography để bổ sung cho việc mã hóa.


STEGANOGRAPHIA:
Hoc eſt:
ARS PER OC-
CVLTAM SCRI-
PTVRAM ANIMI SVI VO-
LVNTATEM ABSENTIBVS
aperiendi certa;
AVTHORE
REVERENDISSIMO ET CLARISSIMO VIRO,
IOANNE TRITHEMIO, Abbate Spanheimenſi, &
Magiæ Naturalis Magiſtro per-
fectiſſimo.
PRÆFIXA EST HVIC OPERI SVA CLAVIS, SEV
vera introductio ab ipſo Authore concinnata;
HACTENVS QVIDEM A MVLTIS MVLTVM DE-
ſiderata, ſed à paucissimis viſa:
Nunc vero in gratiam ſecretioris Philoſophiæ Studioſorum
publici iuris facta.
Cum Priuilegio & conſenſu Superiorum.
FRANCOFVRTI,
Ex Officina Typographica IOANNIS SAVRII, Sumptibus
IOANNIS BERNERI.
Anno M. D C. VIII.

Hình 1. Quyển sách đầu tiên viết về Steganography – 1499

Technique	Purpose	Comments
Steganography	Hiding existence of digital content from outsiders	Content generally of limited time value. Needs carrier file
Cryptography	Rendering the digital content inaccessible to outsiders	Content generally of limited time value. No need for carrier file
Watermarking	Protection of digital content of carrier	May or may not be readily detectable. Durability is essential

Hình 2. Sự khác biệt giữa Steganography và Cryptography, Watermarking

Ẩn giấu thông tin trong lịch sử

Thực chất Steganography đã được con người biết tới từ rất lâu rồi. Năm 440 trước công nguyên, người Herodotus (Người Hy Lạp xa xưa) đã cạo trọc đầu các nô lệ tin cậy rồi

xăm lên đó các thông điệp và chờ tóc mọc lại. Mục đích của việc này là nhằm gửi tin đi trong cuộc chiến tranh giữa người Herodotus và Persians (Người Ba Tư).

Trong cuộc cách mạng của Mỹ, mực không màu cũng được sử dụng để trao đổi thông điệp giữa người Mỹ và người Anh.

Trong các cuộc chiến tranh thế giới, Steganography cũng được sử dụng. Người Đức đã sử dụng mực không màu để viết các dấu chấm nhỏ lên phía trên và dưới các chữ cái bằng cách thay đổi chiều cao các chữ trong đoạn văn bản. Trong chiến tranh thế giới thứ nhất, các tù nhân cũng sử dụng mã Morse để gửi thư về nhà bằng cách viết các dấu chấm và gạch ngang lên các chữ cái i,j,t,f.

Trong chiến tranh thế giới thứ 2, các gián điệp của Đức đã sử dụng đoạn văn bản sau:

"Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects for pretext embargo on by-products, ejecting suets and vegetable oils."

Bằng cách đọc những chữ thứ 2 của các từ ta sẽ được đoạn thông điệp sau:

"Pershing sails for NY June 1".

Ngày nay trên mạng có rất nhiều đoạn nhạc hay các bức hình tưởng chừng như bình thường nhưng nó lại chứa các thông tin hoàn toàn bí mật. Điều này càng làm cho Steganography càng trở nên phổ biến.

Ẩn giấu thông tin trong việc chống tiền giả

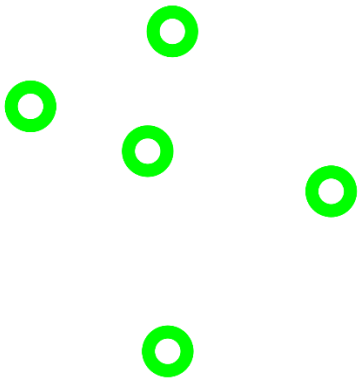
Có những chòm sao bí ẩn, "quyền lực" trên tờ tiền khiến máy photo, scan có thể sập nguồn khi được nhận dạng. Bạn đã nghe chưa?

Làm và lưu hành tiền giả là hành động vi phạm pháp luật, nhưng đã bao giờ bạn thử dùng máy scan hoặc photocopy để scan hoặc sao chép một tờ tiền? Nếu thực hiện điều này với những tờ tiền như đô-la Mỹ, bảng Anh hay đồng Euro... thì bạn sẽ nhận được một kết quả bất ngờ...

Mọi tờ tiền đều có những cơ chế tinh vi để chống làm giả, và thậm chí, ít người biết được rằng cơ chế chống làm giả này còn ngăn chặn việc sử dụng máy photocopy hoặc scan để sao chép hoặc scan một tờ tiền.

Trên phần lớn tờ tiền giấy hiện nay được trang bị một ma trận các điểm ẩn, được gọi là EURion Constellation (tạm dịch là "Chòm sao EURion"), được thêm vào trên các tờ tiền, mà khi máy photocopy hoặc scan gặp phải những "chòm sao" này sẽ ngừng hoạt động hoặc không thể sao chép được, điều này sẽ giúp ngăn chặn việc sao chép hoặc scan tờ

tiền. Khi đó, máy photocopy hoặc scan sẽ không thể sao chép hoặc scan toàn bộ tờ tiền, mà sẽ bỏ qua những khu vực có sự xuất hiện của "chòm sao EURion".

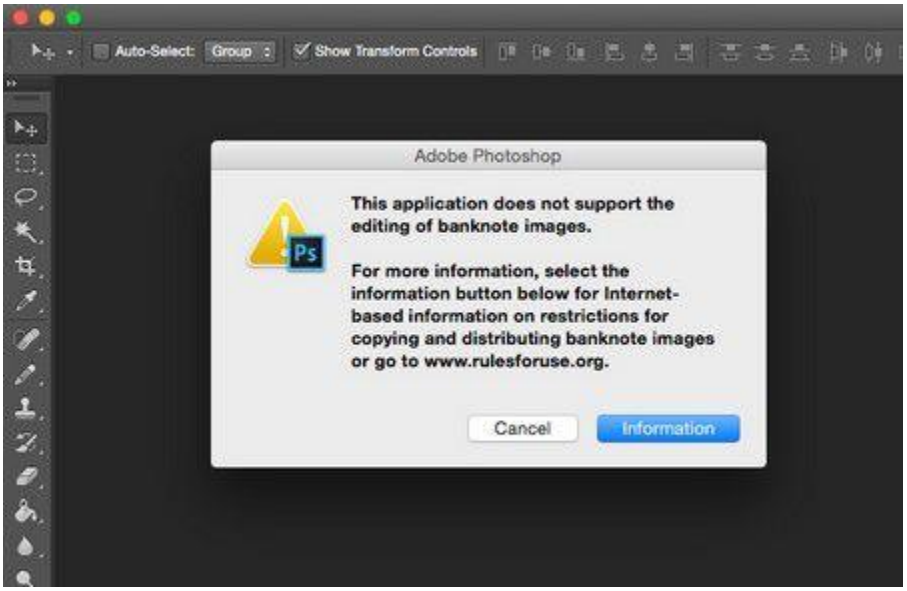


Hình 3. "Chòm sao" EURion

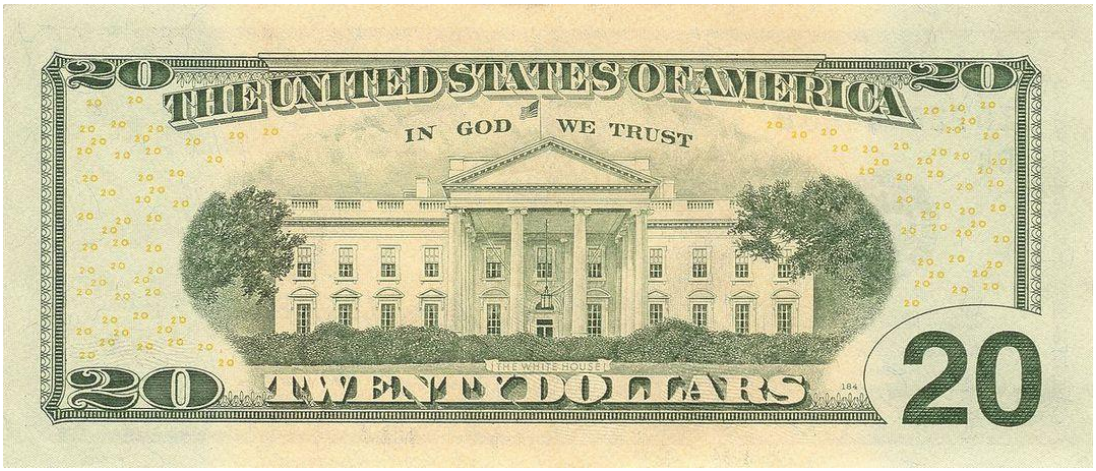
Chòm sao EURion là một mẫu ảnh 5 chấm tròn màu vàng, xanh hay cam mô phỏng chòm sao Orion. Những chấm tròn này chứa những thông tin dùng để ngăn chặn việc sao chép tạo tiền giả từ các máy in ngày nay. Tuy nhiên, công nghệ bảo mật chống làm giả EURion Constellation chỉ được áp dụng rộng rãi từ năm 1996, do vậy với một số tờ tiền được xuất hiện trước thời gian đó thì vẫn có thể sử dụng máy photocopy hoặc scan để sao chép bình thường. Ngoài ra, theo trang Business Insider, EURion Constellation không phát huy tác dụng với một số máy photocopy hoặc scan thế hệ cũ.

Hiện có khoảng 53 loại tiền tệ đang được áp dụng kỹ thuật chống tiền giả EURion Constellation này. Ngoài ra, theo trang công nghệ Gizmodo, các loại tiền mới được phát hành còn được trang bị một hệ thống chống làm giả khác có tên gọi Counterfeit Deterrence System (CDS), tuy nhiên cách thức hoạt động thực sự của hệ thống CDS chưa bao giờ được tiết lộ. CDS được phát triển bởi Hiệp hội Ngăn chặn tiền giả của Ngân hàng Trung ương, là một nhóm gồm 27 ngân hàng Trung ương của 27 quốc gia khác nhau, được Chủ trì bởi Chủ tịch Ngân hàng Quốc gia Thụy Sĩ.

Không giống như công nghệ chống làm giả "chòm sao EURion", CDS không chỉ giúp ngăn chặn việc các tờ tiền bị sao chép hoặc scan bằng máy photocopy và máy scan, CDS còn ngăn chặn việc hình ảnh các tờ tiền bị chỉnh sửa trên các phần mềm đồ họa, chẳng hạn như Photoshop. Với kỹ thuật chống tiền giả này, khi bạn mở một hình ảnh tờ tiền trong phần mềm Photoshop để chỉnh sửa nó, lập tức sẽ nhận được thông báo "phần mềm không hỗ trợ việc chỉnh sửa hình ảnh tờ giấy bạc".



Hình 4. Chòm sao EURion trên mẫu các nốt nhạc ở một tờ tiền của Anh

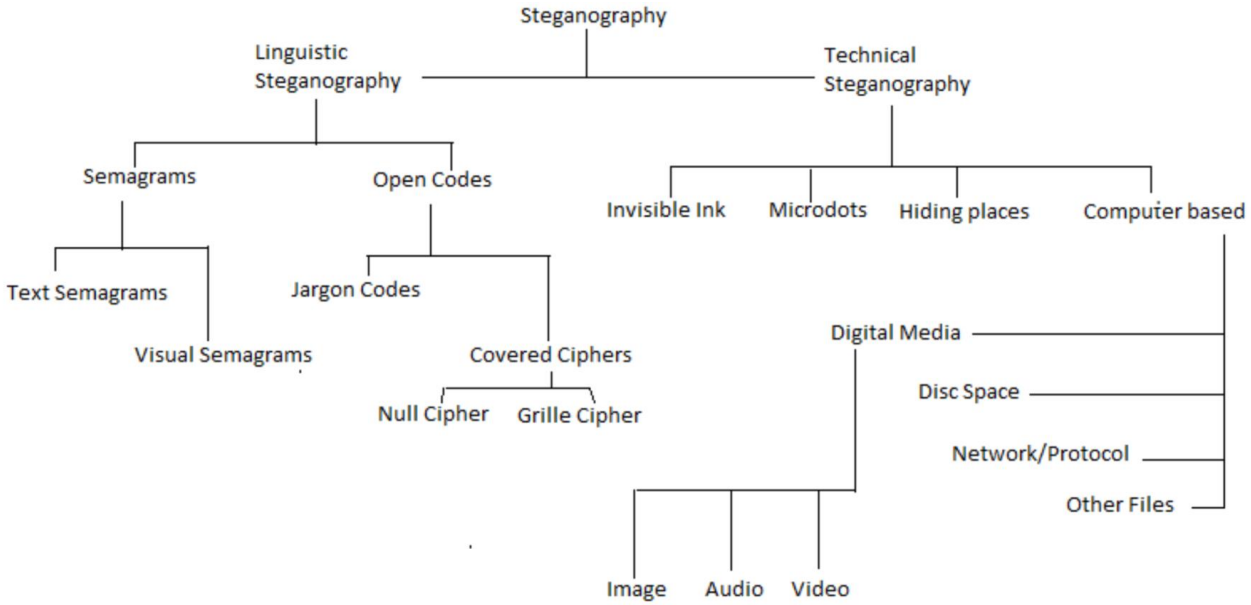


Hình 5. Các ký số 0 ở 2 phía của tờ 20 đô la Mỹ được dùng để phân biệt tiền thật

Hình 6. Các số 0 tạo thành mẫu chòm sao Orion

Ví dụ một số chòm sao Eurion trên các tờ tiền ở một số nước:

Các trường phái ẩn giấu thông tin



Một số phương pháp dùng trong ẩn giấu tin có thể dùng với các trường phái trên:

Technique	Method	Effect on carrier file	Comments
Injection techniques	Using built-in information recording tools or "open" file space	No change of content	Very limited hiding capacity
Substitution techniques	Part of digital content of carrier file changed to reflect stego message	Some degradation of content quality	Increased risk of detection with increasing volume of stego content
File creation	Stego message hidden in larger amount of new, irrelevant digital content	None- new carrier file created	Inefficient, detection risk highly dependent on context of message
Stego encryption	Stego content is encrypted as it is included in carrier file	Nome beyond pure steganography	Key exchange increases risk detection

Hiện nay, có hơn 100 chương trình steganography khác nhau, từ dạng tải về miễn phí đến hình thức thương mại. Sau đây, một ví dụ về ẩn giấu một bức ảnh bản đồ sân bay Burlington, Vermont (định dạng GIF) trong các tập tin GIF, JPEG và WAV.

Hình 7. Hình ảnh bản đồ sân bay Burlington

Chương trình **gIf-It-Up** của Nelsonsoft sẽ ẩn giấu bức ảnh trong Hình 2 trong một file ảnh GIF khác bằng cách sử dụng các bit thấp cùng với tùy chọn mã hóa. Kết quả như Hình 3.

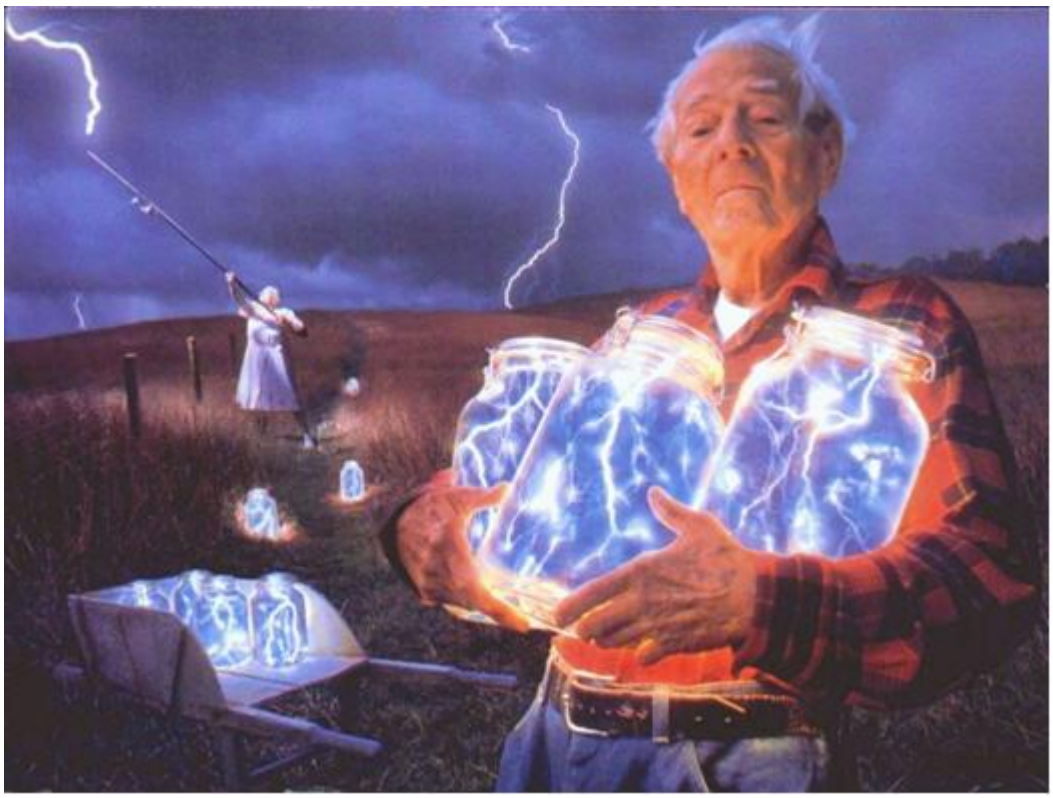
Hình 8. Bức ảnh Washington DC sau khi chèn bức ảnh bản đồ Burlington

Hình 9. Thông tin ảnh Washington DC rước khi chèn thông tin

Hình 10. Thông tin bức ảnh Washington sau khi chèn thông tin

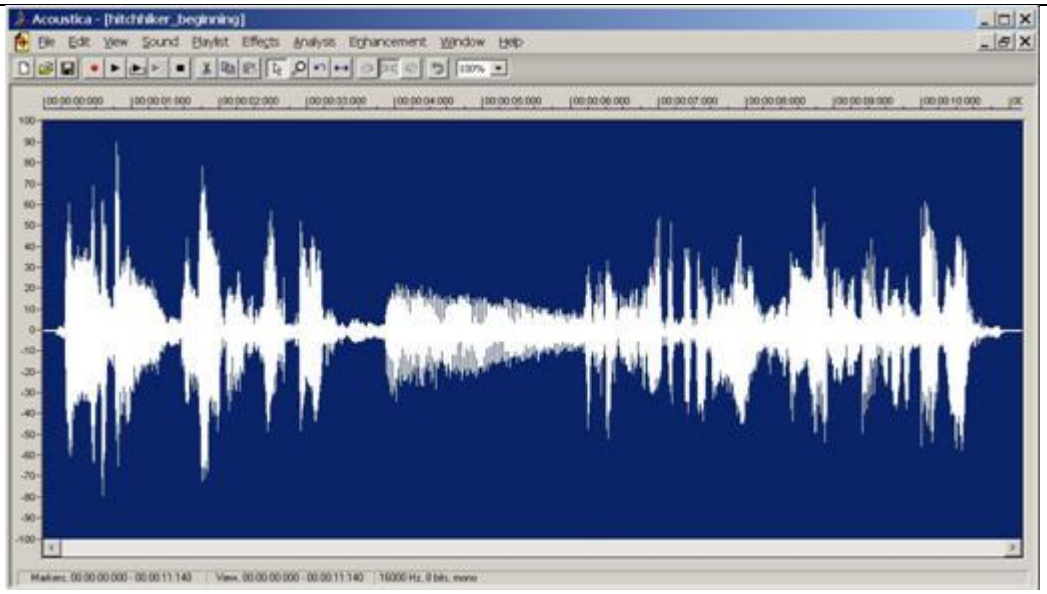
Cụ thể, trước chi chèn bức ảnh bản đồ vào, ảnh cảnh đêm Washington DC có dung lượng 632,778 bytes, với 249 màu. Sau khi được chèn thông tin vào, nó có kích thước 677,733 bytes, sử dụng 256 màu. Rõ ràng, kích thước file ảnh sau khi được chèn dữ liệu vào có sự thay đổi bởi vì lựa chọn điều chỉnh màu sắc (color extension) để giảm thiểu sự phá vỡ, ảnh hưởng tới bức ảnh gốc.

Tương tự, chương trình **JP Hide-&-Seek (JPHS)** được viết bởi Allan Latham dùng để ẩn giấu thông tin trong các ảnh JPEG ở các bit thấp (least significant bit) với thuật toán mã hóa Blowfish.

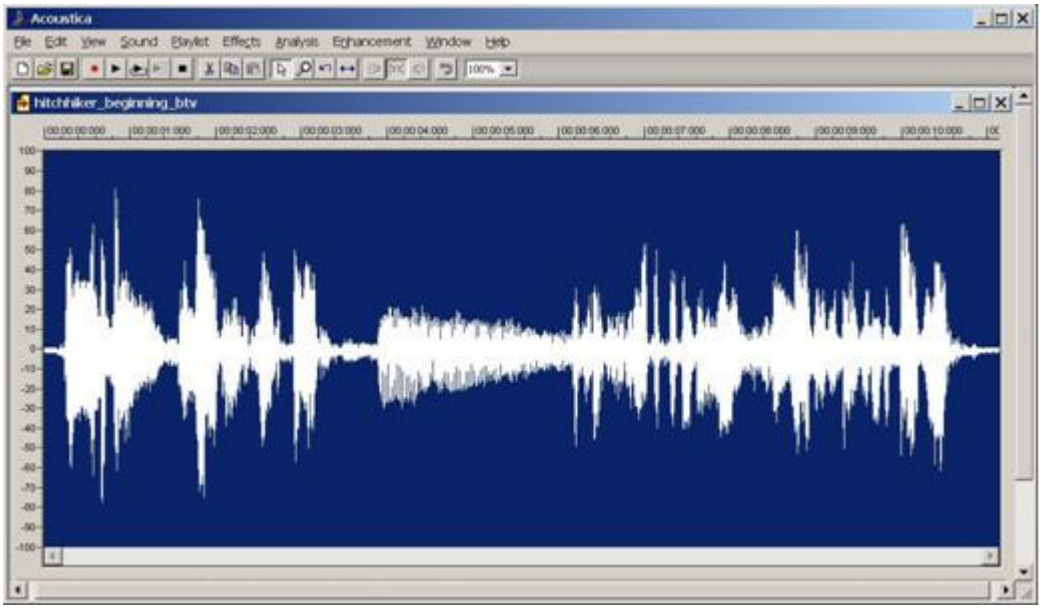


Hình 11. Bức ảnh chứa bản đồ sân bay sau khi sử dụng JPHS

Cuối cùng, trong Hình 7 và Hình 8, bức ảnh sân bay Burlington được chèn vào một tệp tin âm thanh bằng ***công cụ S-Tools*** được viết bởi Andy Brown. Phần mềm này có thể giấu thông tin vào bên trong các tập tin ảnh GIF, BMP và âm thanh WAV. Nó cũng sử dụng bit thấp (least significant) và nén không mất mát (lossless compression).



Hình 12. Thông tin tập tin âm thanh ban đầu với kích thước 178,544 bytes

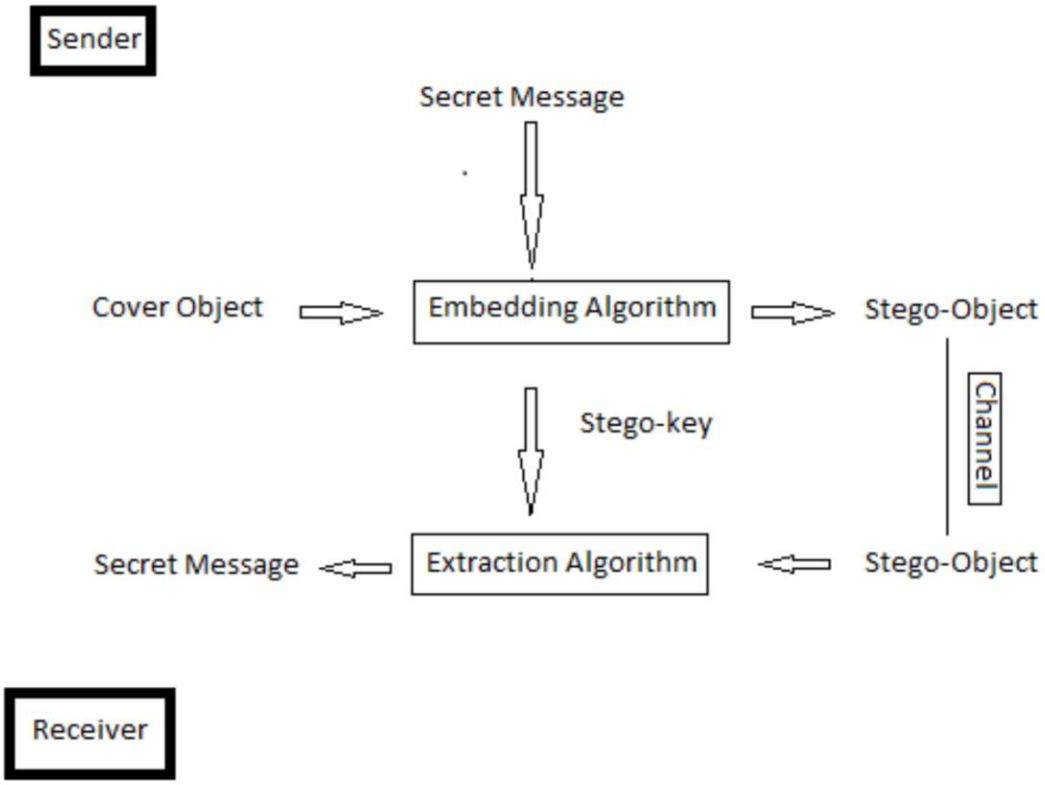


Hình 13. Tập tin âm thanh sau khi chèn thông tin với kích thước chỉ 178,298 bytes (nhỏ hơn kích thước file gốc ban đầu, mặc dù đã được chèn bức ảnh vào)

Quá trình mã hóa và giải mã của Steganography

Đầu tiên, thông điệp và đối tượng ban đầu (cover object – đối tượng được chọn để chứa thông tin cần ẩn giấu, ví dụ như bức ảnh, file âm thanh,…) sẽ được đưa vào trong bộ mã hóa (encoder). Trong bộ mã hóa, một bộ giao thức sẽ được thực thi để nhúng thông điệp vào trong bức ảnh. Kiểu của giao thức sẽ phụ thuộc vào kiểu của thông điệp

và cách ta nhúng. Ví dụ hình dưới đây sẽ cho ta thấy giao thức để ẩn thông điệp trong các đối tượng được chọn. Trong nhiều trường hợp, khóa là cần thiết trong quá trình giấu thông tin. Người gửi có thể dùng khóa chính để mã hóa và người nhận có thể dùng khóa công khai để giải mã. Điều này có thể giảm thiểu khả năng bên thứ 3 biết được thông điệp và giải mã.



Hình 14. Qui trình ẩn và giải mã thông tin trong Steganography

Kỹ thuật giấu tin trong ảnh (Image Steganography)

Giấu tin trong ảnh là việc thực hiện giấu thông tin với môi trường chứa là các file ảnh. Hiện nay, giấu tin trong ảnh chiếm tỉ lệ lớn trong các ứng dụng giấu tin trong dữ liệu đa phương tiện bởi vì lượng thông tin được trao đổi bằng hình ảnh là rất lớn. Giấu tin trong ảnh có nhiều ứng dụng trong thực tế, ví dụ như trong việc xác định bản quyền sở hữu, chống xuyên tạc thông tin hay truyền dữ liệu một cách an toàn,…

Các khái niệm thường được dùng trong giấu tin trong ảnh:

- Ảnh môi trường: Là ảnh gốc được dùng để nhúng thông tin.
- Thông tin nhúng: Là các thông tin mật cần gửi.
- Ảnh đã nhúng: Là ảnh gốc sau khi đã được nhúng thông tin mật.

- Khóa mật: Là khóa tham gia vào quá trình nhúng, được trao đổi giữa người gửi và người nhận.

Các yêu cầu trong giấu tin trong ảnh:

- Tính bền vững: Thể hiện khả năng ít bị thay đổi (về nội dung, hình dạng) trước những tấn công từ bên ngoài. Hiện nay, chưa có kỹ thuật giấu tin nào đảm bảo được yêu cầu này một cách tuyệt đối.
- Khả năng không bị phát hiện: Thể hiện ở việc khó xác định được đối tượng có chứa thông tin mật hay không. Các kỹ thuật giấu tin hiện nay cố gắng đảm bảo yêu cầu này dựa vào hệ thống thị giác của con người.
- Khả năng lưu trữ: Thể hiện ở lượng thông tin được lưu trữ. Do còn phải đảm bảo "khả năng không bị phát hiện" nên với những thông tin mật lớn, ta thường chia nhỏ nó ra, nhúng nhiều lần và vào các đối tượng khác nhau.

Cơ chế ẩn giấu thông điệp trong ảnh bằng cơ chế LSB (Least Significant Bit)

Đối với ảnh 8 bit chúng ta chỉ có thể ẩn 1 bit dữ liệu cho mỗi pixel còn đối với ảnh 24 bit ta có thể ẩn 3 bit dữ liệu cho mỗi pixel. Quá trình mã hóa để giấu thông điệp vào bức ảnh:

- Trước hết chúng ta cần chuyển thông điệp sang dạng nhi phân. Nếu ta muốn giấu chữ A thì ta cần giấu đoạn mã sau: 10000001
- Đọc các pixel trong b ức ảnh (Original Image), để tìm ra các giá trị R-G-B của từng pixel một.

Giả sử ta chọn ra 3 pixel để ẩn 3bit của byte thấp của kí tự A như sau:

```
00100111 11101001 11001000

00100111 11001000 1110100 1

11001000 00100111 1110100 1
```

- Ta sẽ thay từng bit của mã nhị phân của kí tự A vào từng pixel trên. Lúc này ta có pixel sau:

```
00100111 11101001 1100100 0

00100111 11001000 1110100 0

11001000 00100111 1110100 1
```

Như vậy chỉ có 3 bit thay đổi trong 3 pixel nên chúng ta rất khó có thể phát hiện bằng mắt thường. Quá trình giải mã để lấy thông điệp ra khỏi bức ảnh:

- Đọc từng pixel của bức ảnh nhận được.
- Đọc giá trị R-G-B của từng pixel, đọc các LSB của từng giá trị R-G-B, ta được một chuỗi nhị phân ban đầu của thông điệp gốc. Sau đó ta chuyển chuỗi nhị phân về dạng ASCII là hoàn thành.

Steganalysis (Phát hiện giấu tin)

Giống như trong Mật mã (Cryptography), thì Thám mã (Cryptanalysis) là kỹ thuật đối lập nhưng song song tồn tại và phát triển cùng với sự phát triển của kỹ thuật mật mã, nhằm giải mã các—bản mã‖ thu được để hiểu rõ nội dung ban đầu của bản mã, thì phát hiện giấu tin (Steganalysis) là kỹ thuật đối lập với Steganography nhằm dò tìm ảnh số nào đó có giấu thông tin hay không.

Việc nghiên cứu Steganalysis ngoài ý nghĩa khoa học còn có hai ý nghĩa thực tiễn, đó là: Thứ nhất, nhằm phục vụ đắc lực cho lĩnh vực an toàn thông tin; Thứ hai, nhằm nâng cấp và thúc đẩy sự phát triển của kỹ thuật giấu tin trong ảnh. Với hai mục đích nêu trên dẫn đến hai hướng nghiên cứu khác nhau. Hướng thứ nhất, cố gắng xây dựng thuật toán phát hiện mù (blind steganalysis) cho ảnh có giấu tin sử dụng kỹ thuật giấu bất kỳ. Hướng thứ hai, dựa vào kỹ thuật giấu tin nào đó đã biết, có thể xây dựng được thuật toán phát hiện phù hợp (phát hiện có ràng buộc – constraint steganalysis).

Đã có nhiều công trình nghiên cứu công bố trên thế giới thành công theo hai hướng này. Tuy nhiên, các kỹ thuật giấu tin ra đời sau ngày càng tinh xảo hơn đòi hỏi các nhà phát hiện ảnh giấu tin không ngừng tìm ra phương pháp phát hiện phù hợp bắt kịp với xu hướng phát triển của kỹ thuật giấu tin. Đặc biệt với tốc độ phát triển nhanh chóng của Internet thì nhu cầu trao đổi thông tin bằng ảnh ngày càng lớn mạnh, do đó để đảm bảo an toàn an ninh, quốc phòng hay nhằm hỗ trợ nâng cấp, cải tiến kỹ thuật giấu nào đó an toàn hơn đang là bài toán cấp thiết đặt ra cho các nhà nghiên cứu trong lĩnh vực an toàn thông tin hiện nay.

3. Môi trường & cấu hình

- Sử dụng các thiết bị và tài liệu, khuyến cáo được cung cấp bởi GVTH, yêu cầu tác phong nghiêm túc trong quá trình thực hiện.
- Công cụ gợi ý: **Irfanview, Stegdetect, stegobreak, JSteg, S-Tools, Our secret ...**
- Tài liệu nên đọc: Sách ***"Handbook of Research on Secure Multimedia Distribution"*** (tác giả: Shiguo Lian), Sách "***Steganography: A New Technique To Hide Information Within Image File***" (tác giả: Ram Kumar Singh, Amit Asthana).

B. THỰC HÀNH

Sinh viên thực hiện điều tra theo yêu cầu của GVHD, làm theo nhóm thực hành đã đăng ký trên lớp trong buổi thực hành.

B1. Phân tích ảnh bằng công cụ Irfanview

Giúp sinh viên nắm bắt và hiểu rõ các tính năng của công cụ phần mềm **Irfanview** khi tiến hành điều tra và tìm kiếm thông tin trong một file ảnh.

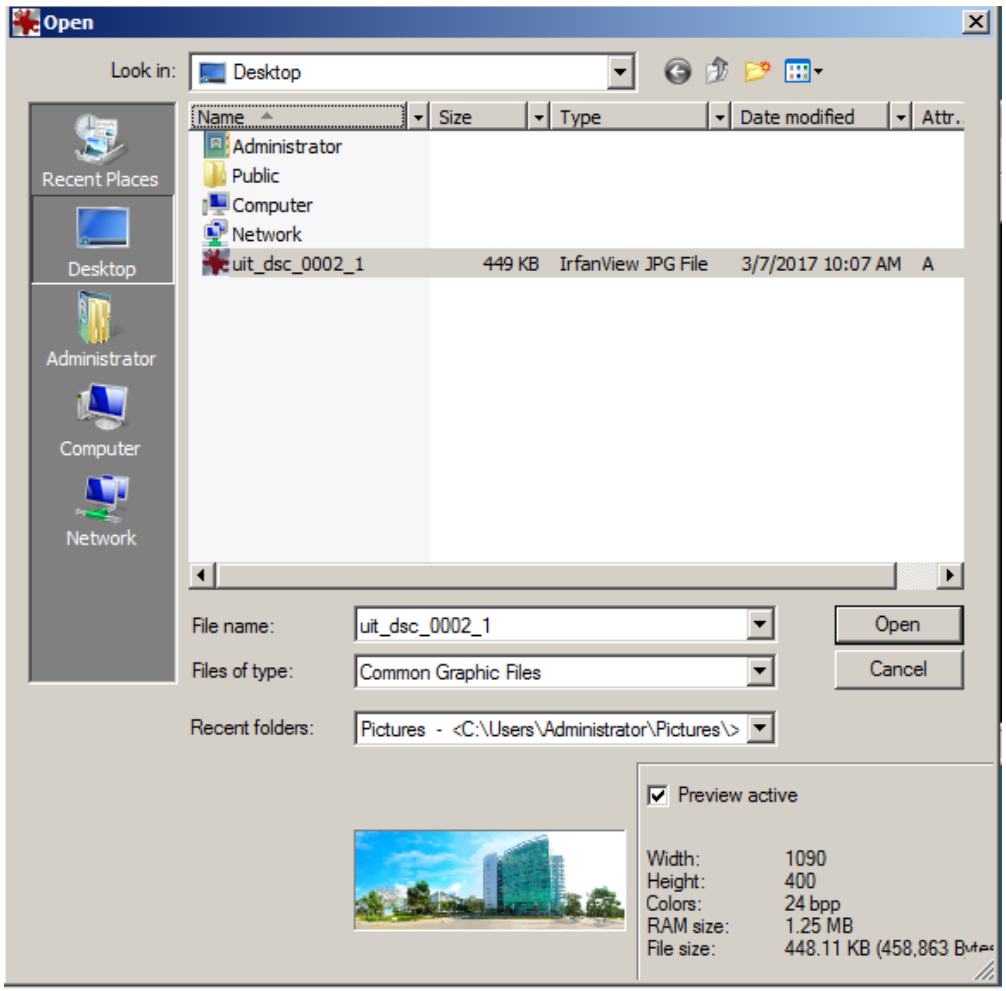
– Thực hiện cài đặt **Irfanview.**

Liên kết tải: https://www.irfanview.net/

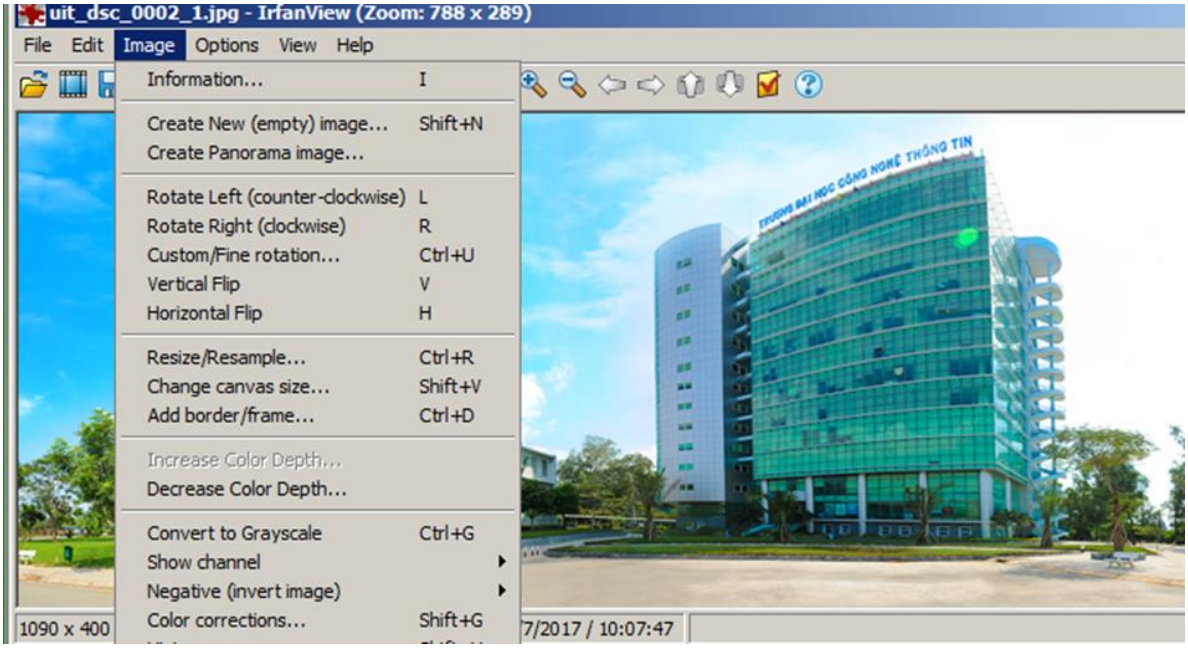
– Tải về tập tin ảnh **uit_dsc_0002_1.jpg** tại liên kết:

http://www.uit.edu.vn/sites/vi/files/slider/uit_dsc_0002_1.jpg

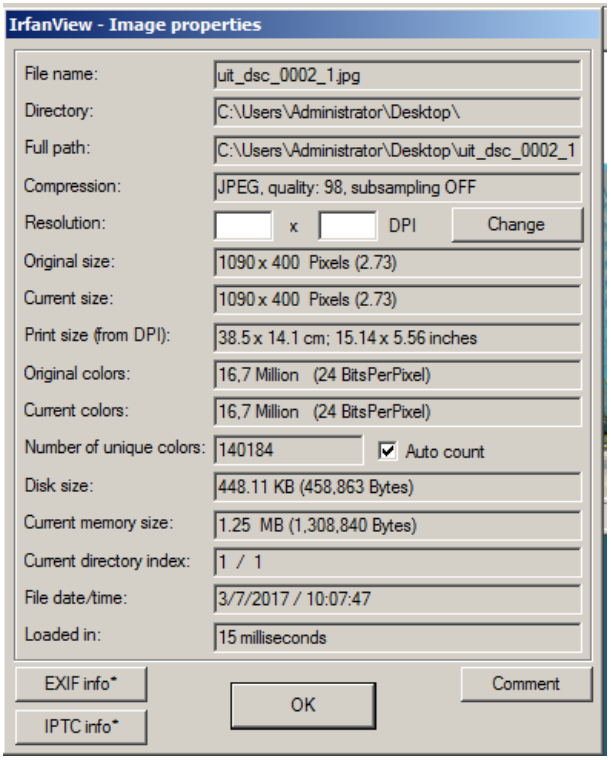
– Mở file ảnh vừa tải bằng Irfanview.



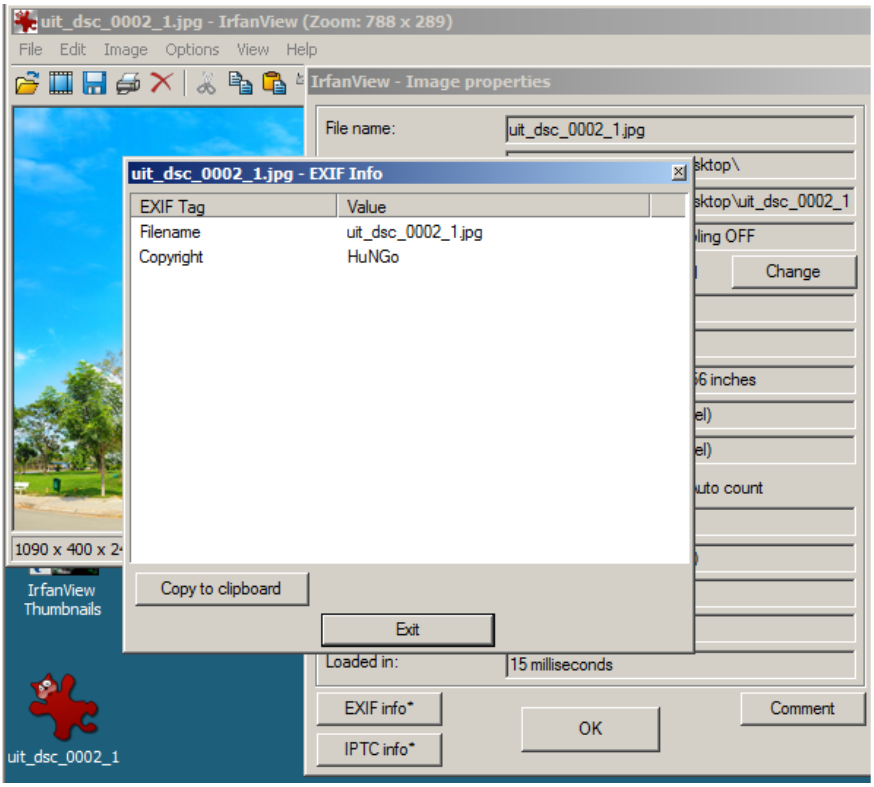
– Xem thêm thông tin của ảnh. Ở menu, chọn Image -> Information



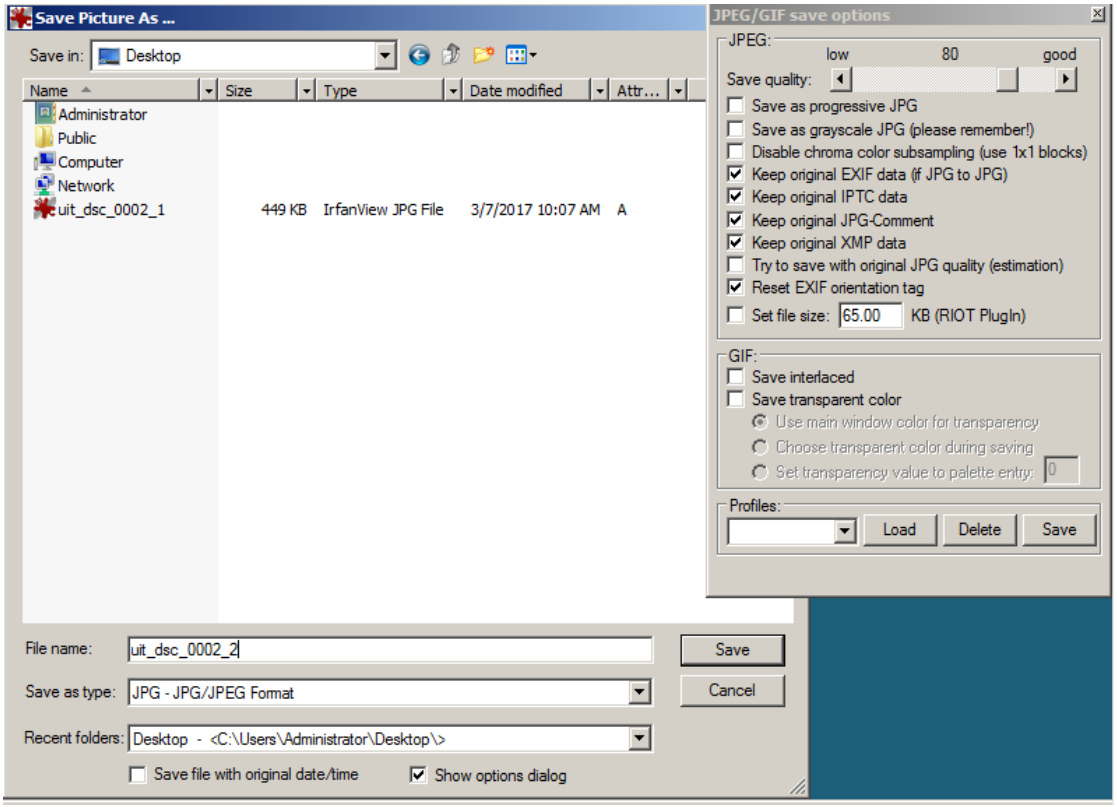
– Thông tin của ảnh:



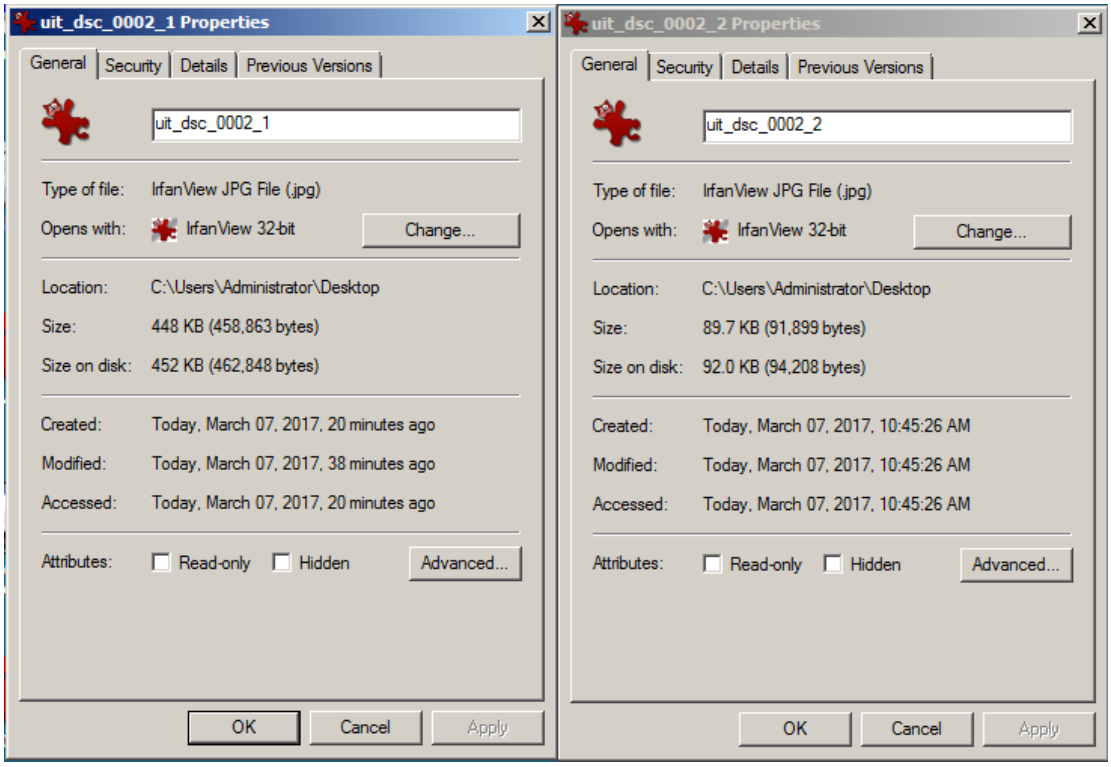
Click vào EXIF info, xem thông tin EXIF của ảnh.



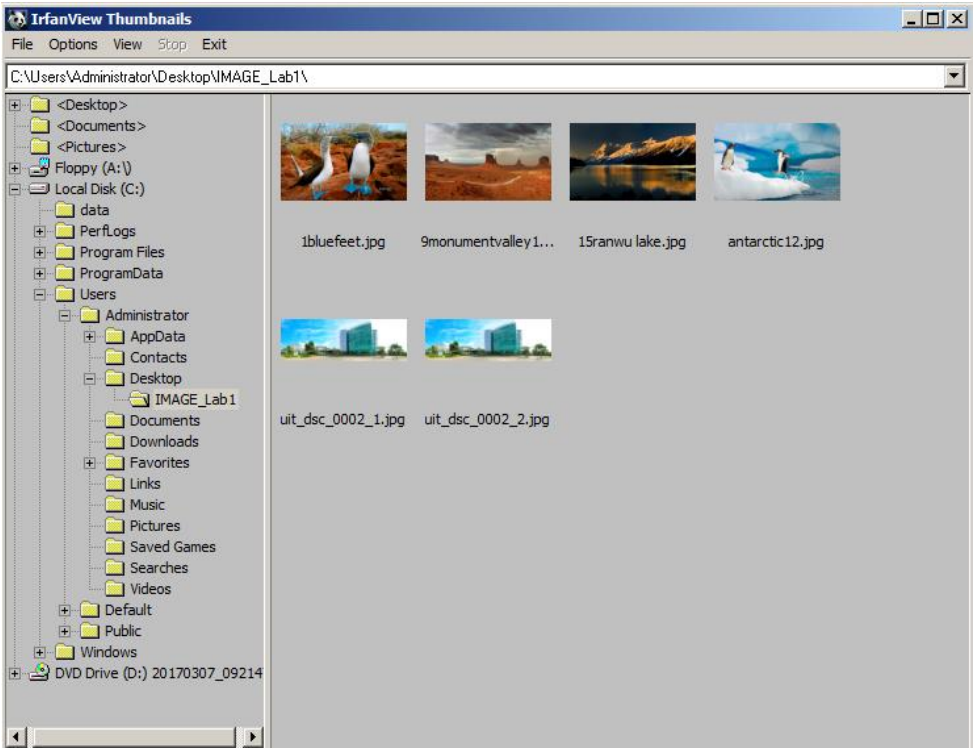
– Điều chỉnh kích thước, chất lượng ảnh:



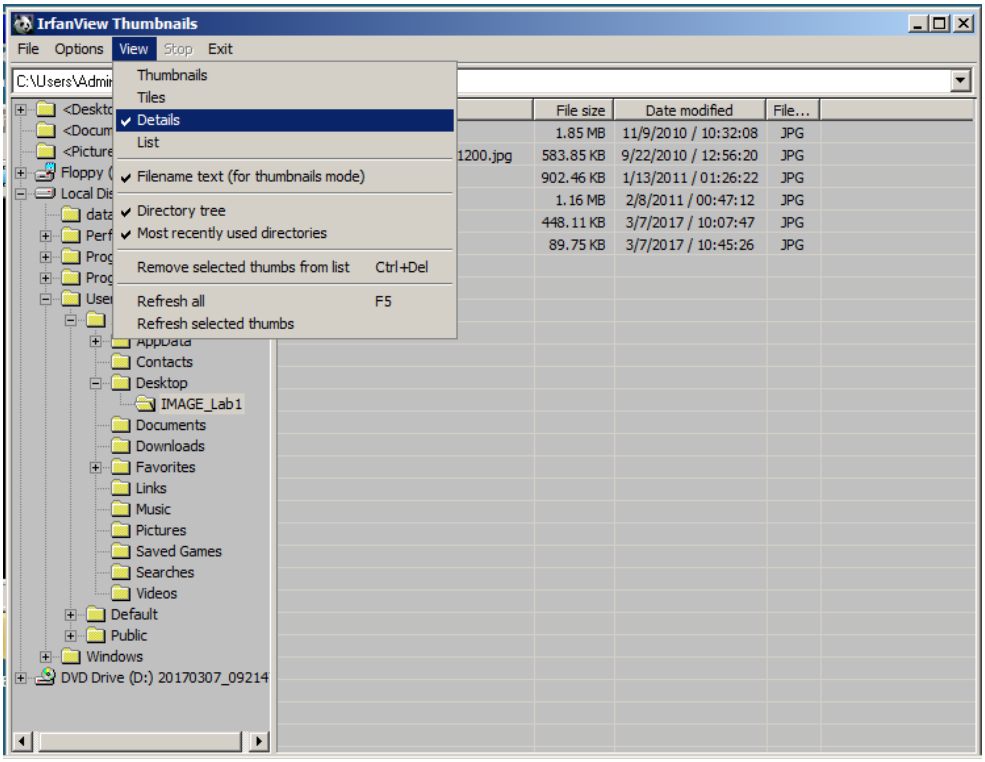
– Kết quả sau khi lưu ảnh đã chỉnh sửa:



– Sắp xếp ảnh trong Irfanview bằng ngày chỉnh sửa cuối. Chọn File -> Thumbnails (T)



– Vào mục View -> Details để hiển thị nhiều thông tin về ảnh hơn.



Kịch bản 01-a. Thực hiện phân tích thông tin tập tin ảnh

- Tài nguyên thực hiện, nằm trong thư mục kb-01-a
- Yêu cầu: Cung cấp các thông tin chi tiết liên quan tới các bức ảnh trên bằng phần mềm IrfanView

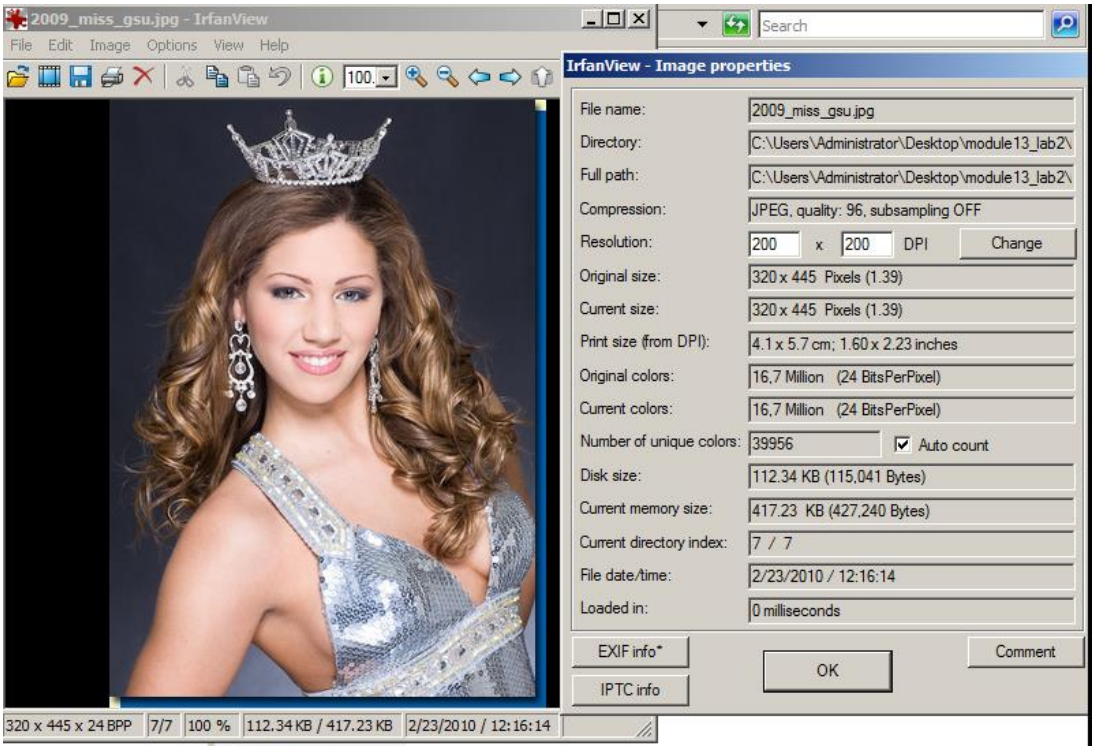
Đáp án:

B2. Thực hiện giấu tin và giải mã thông tin trong ảnh

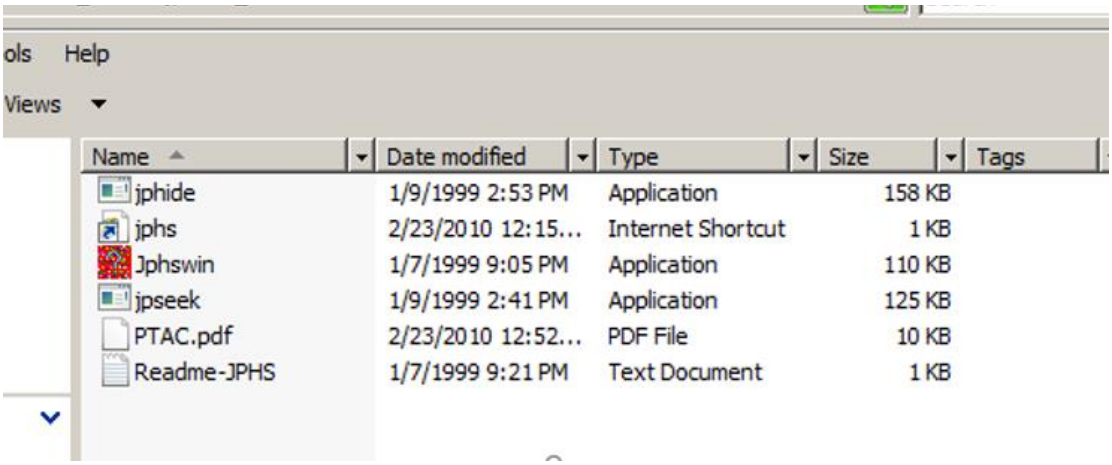
Phần này sẽ giới thiệu các chức năng của phần mềm JP Hide & Seek.

a. Kịch bản 01-b: Giấu tin và giải mã thông tin trong ảnh

- Tài nguyên ảnh trong file nén images_session03.zip (Ảnh: 2009_miss_gsu.jpg)



- Công cụ của kịch bản, dữ liệu cần giấu nằm trong file nén jphs05_session03.zip. Giải nén ra:

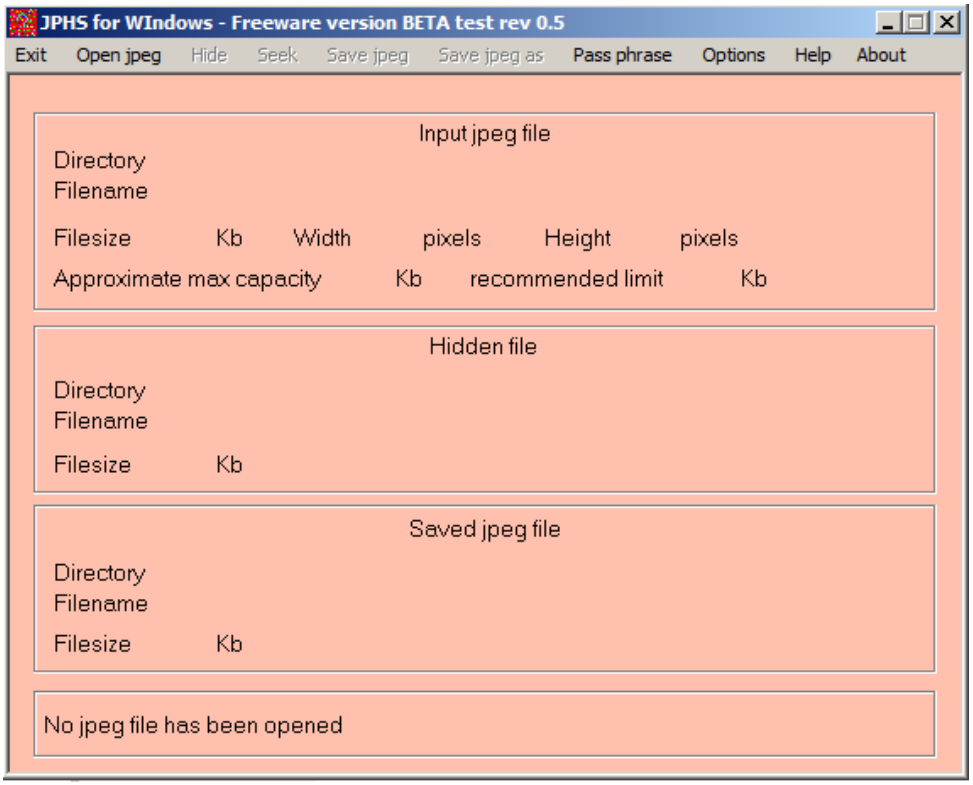


- Nội dung file PTAC.pdf

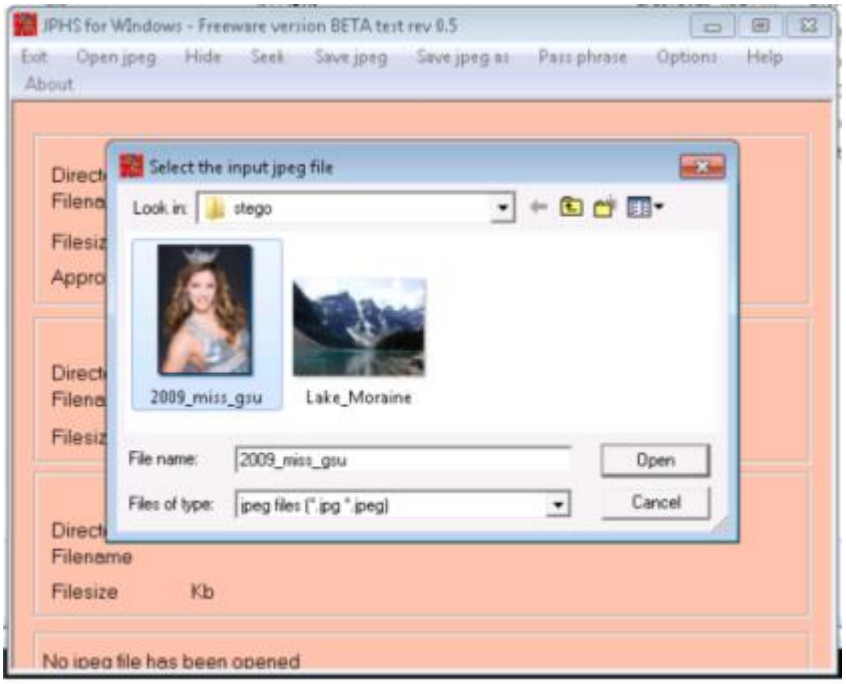
PTAC.pdf

Done	Name	User Name	E-mail	Password	External Login Info.
	Ann Rawlings	ARawlings	ARawlings@peoriaud.k12.az.us	your current password	pusd11\user name
	Chris Kuczka	CKuczka	CKuczka@peoriaud.k12.az.us	your current password	pusd11\user name
	Cindy Callaway	CCallaway	CCallaway@peoriaud.k12.az.us	your current password	pusd11\user name
	Dave Pearson	DPearson	DPEarson@peoriaud.k12.az.us	your current password	pusd11\user name
	David Snyder	Dsnyder	DSnyder@peoriaud.k12.az.us	your current password	pusd11\user name
	Jo Little	JLittle	JLITTLE@peoriaud.k12.az.us	your current password	pusd11\user name
	Julia Erickson	JErickson	JErickson@peoriaud.k12.az.us	your current password	pusd11\user name
	Larry Buchanan	Lbuchanan	LBuchanan@peoriaud.k12.az.us	your current password	pusd11\user name
	Lissa Cuellar	LCuellar	LCuellar@peoriaud.k12.az.us	your current password	pusd11\user name
	Maggie Olney	MOlney	MOlney@peoriaud.k12.az.us	your current password	pusd11\user name
	Nan Gillispie-DAC	NGillisp	NGillisp@peoriaud.k12.az.us	your current password	pusd11\user name
	Nathan Bowler	NBowler	NBowler@peoriaud.k12.az.us	your current password	pusd11\user name
	Patti Beltram	PBeltram	PBeltram@peoriaud.k12.az.us	your current password	pusd11\user name
	Phil Valentine	PValentine	PValentine@peoriaud.k12.az.us	your current password	pusd11\user name
	Robert Keagle	RKeagle	RKeagle@peoriaud.k12.az.us	your current password	pusd11\user name
	Rosemary Martin-Moore	RMMoore	RMMoore@peoriaud.k12.az.us	your current password	pusd11\user name
	Samantha Middagh	Smiddagh	SMiddagh@peoriaud.k12.az.us	your current password	pusd11\user name
	Sarah Balder	SBalder	SBalder@peoriaud.k12.az.us	your current password	pusd11\user name
	Shona Miranda	SMiranda	SMiranda@peoriaud.k12.az.us	your current password	pusd11\user name
	Steve Savoy	SSavoy	SSavoy@peoriaud.k12.az.us	your current password	pusd11\user name
	Teri Nevarez	TNevarez	TNevarez@peoriaud.k12.az.us	your current password	pusd11\user name
	Terrie Rust	TRust	TRust@peoriaud.k12.az.us	your current password	pusd11\user name
	Valerie Naish	VNaish	VNaish@peoriaud.k12.az.us	your current password	pusd11\user name
				your current password	pusd11\user name
	Bill Copeland	Bill.Copeland	GSMOM@COX.NET	BcNeT45	pusd11ext\Bill.Copeland
	Tammara Edgin	Tammara.Edgin	tammarae@microsoft.com	TeCom23	pusd11ext\Tammara.Edgin
	Diane Douglas	Diane.Douglas	dmdouglas@cox.net	DdNet21	pusd11ext\Diane.Douglas
	Mary Crespino	Mary.Crespino	crespy@cox.net		

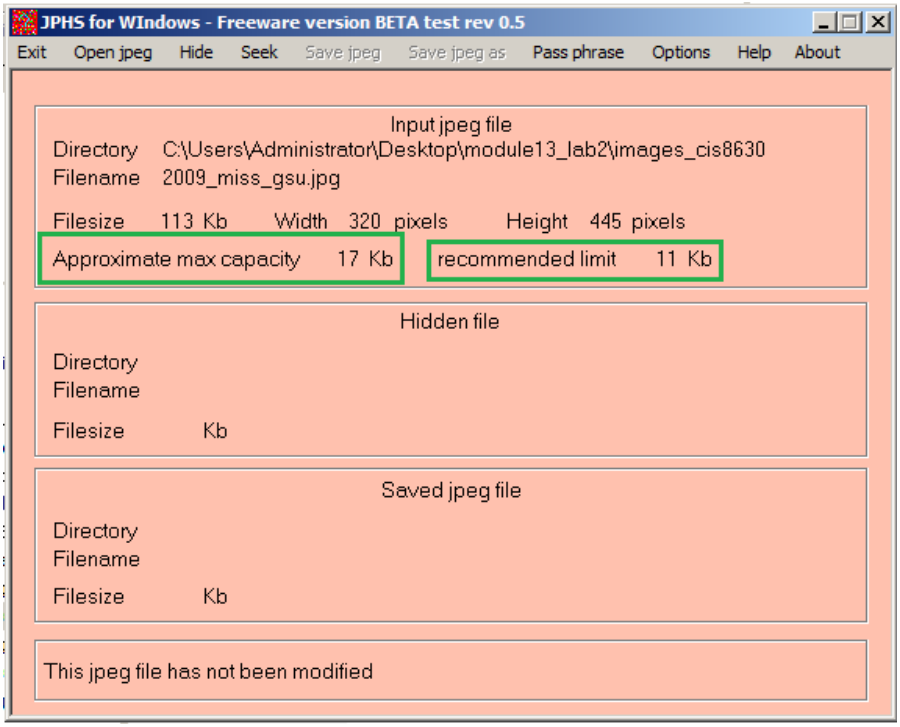
- Chạy file Jphswin.exe:



- Chọn "Open jpeg" để mở ảnh muốn giấu tin (2009_miss_gsu.jpg)



- Sau khi bấm chọn thêm ảnh vào, phần mềm sẽ hiển thị các thông tin gợi ý liên quan đến kích thước dữ liệu được hỗ trợ khi chèn vào ảnh:

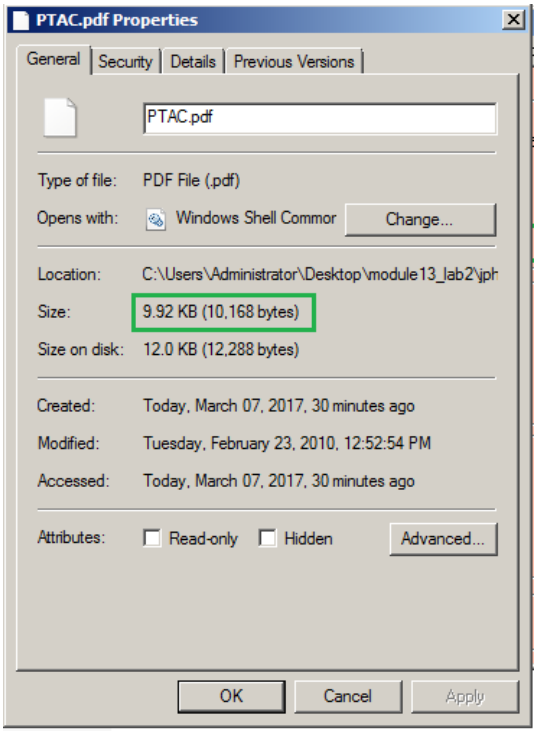


Dung lượng ảnh: 113Kb

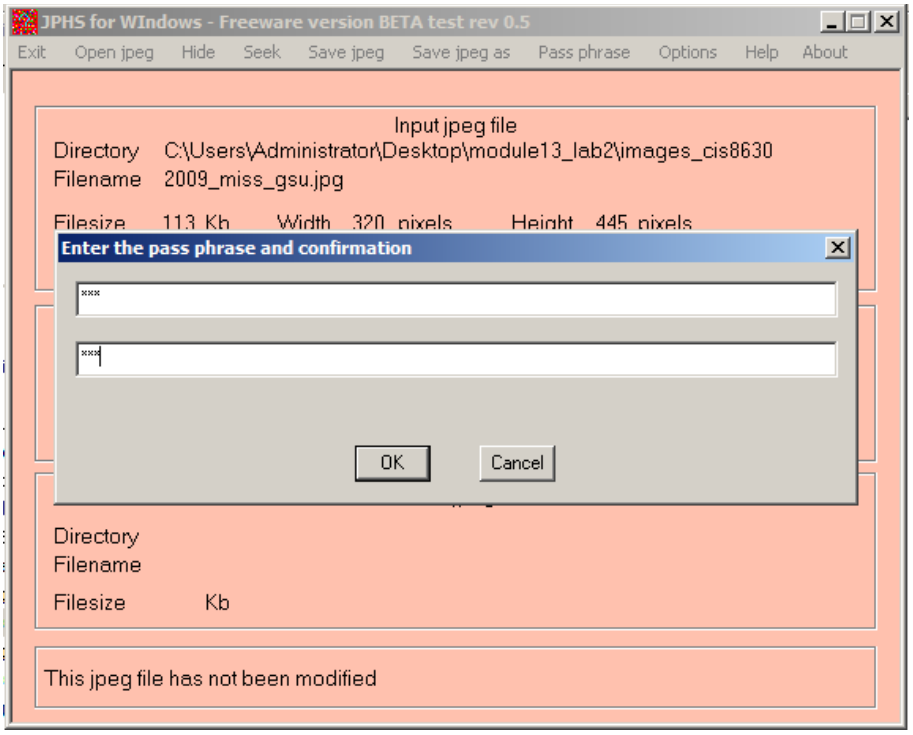
Dung lượng dữ liệu tối đa có thể thêm vào: 17Kb

Dung lượng dữ liệu đề nghị thêm vào: 11Kb

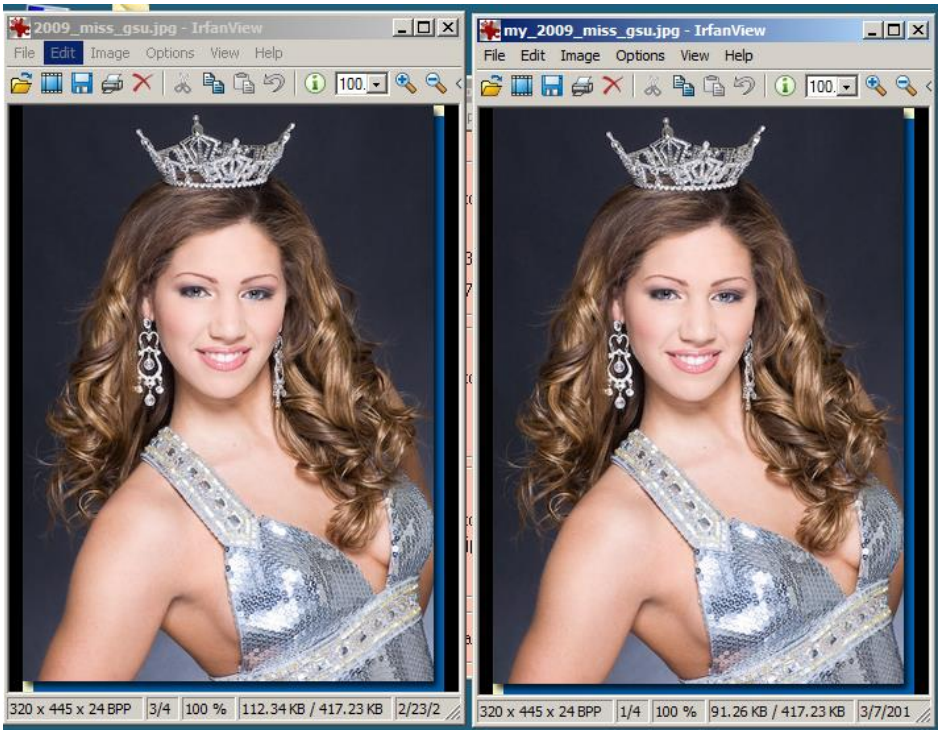
- Thông tin file PTAC.pdf muốn giấu:



- Thêm dữ liệu muốn giấu. Chọn "hide", đặt password là "UIT" và nhấn OK.



- Sau khi chọn file cần giấu và đặt mật khẩu, chọn **Save jpeg as, đặt tên là my_2009_miss_gsu.jpg** để lưu ảnh chứa thông tin. Cho biết sự khác biệt về 2 bức ảnh (dung lượng, chất lượng).



- Sau khi ẩn giấu thông tin, thực hiện khôi phục thông tin được giấu trong ảnh bằng chức năng Seek của phần mềm. Nhấn "Seek" trên menu để trích xuất thông tin trong ảnh. Nhập password "UIT" (Password đặt khi thực hiện giấu tin). Lưu dữ liệu vừa trích xuất với tên "my_PTAC.pdf". Nhận xét về thông tin của file dữ liệu vừa trích xuất với dữ liệu ban đầu.

Kịch bản 01-c. Phát hiện dữ liệu được giấu trong ảnh JPEG sử dụng StegDetect.

- Tài nguyên: image_session03.zip
- Công cụ: stegdetect04_session03.zip. Thực hiện giải nén và chạy file "xsteg.exe"
- Chọn thư mục chứa ảnh cần phân tích. Thực hiện quét và đưa ra kết quả, nhận xét.
- Thực hiện bẻ khóa mật khẩu trong quá trình giấu tin. (Chuẩn bị: my_2009_miss_gsu.jpg - ảnh đã giấu thông tin ở kịch bản 02 bên trên, Zebras2.jpg, Stegbreak.exe). Mật khẩu tìm thấy là gì? Nhận xét về khả năng tìm thấy của công cụ?
- Giải nén thông tin chứa trong file ảnh có phát hiện ẩn giấu thông tin bằng mật khẩu tìm được.

- Mở file vừa giải mã được từ Zebras2.jpeg. Xác định đây có thể là định dạng file gì? Xem nội dung của file này.

Đáp án:

Gợi ý:

Bẻ khóa mật khẩu dùng trong quá trình giấu tin:

- Đưa ảnh và stegbreak.exe cùng thư mục
- Thử mật khẩu với file ảnh để giải mã thông tin bằng cách sử dụng câu lệnh:

  stegbreak -r rules.ini -f meddict.dic my_2009_miss_gsu.jpg

  - -r rules.ini: file đính kèm theo chương trình
  - -f meddict.dic: load file từ điển vào chương trình
  - my_2009_miss_gsu.jpg: tên ảnh muốn giải mã.

Kịch bản 02. Ẩn giấu dữ liệu bằng công cụ Our Secret

- Tài nguyên: uit_dsc_0002_1.jpg, blossom.mp4
- Phần mềm Our Secret: có thể tải tại liên kết sau:

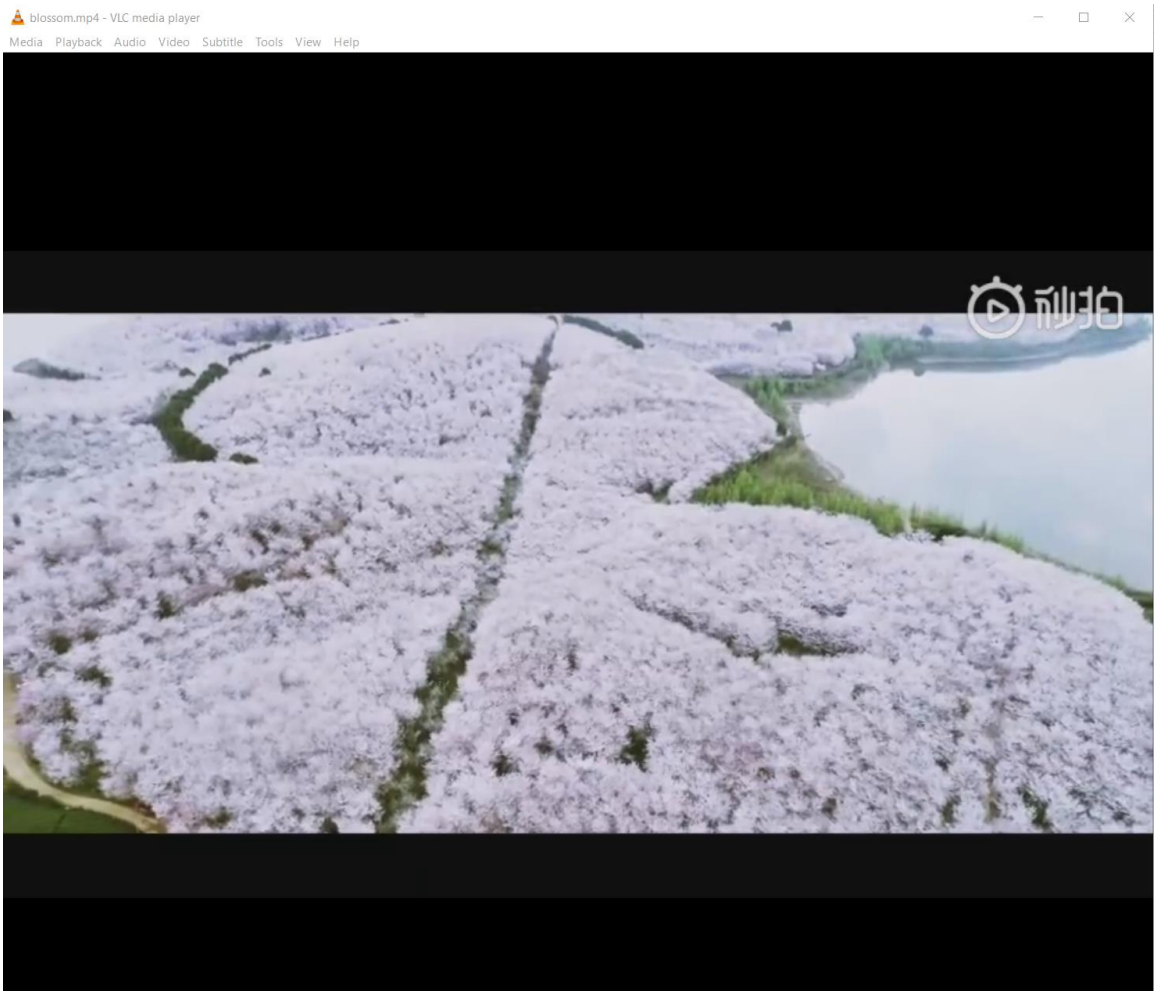
  http://steganography.findmysoft.com/
- Cài đặt phần mềm, sau đó giấu ảnh uit_dsc_0002_1.jpg vào tập tin mp4. Đặt mật khẩu trong quá trình giấu tin là "E81". Nhận xét về sự thay đổi của video (thời gian, dung lượng, chất lượng) khi thêm ảnh vào đoạn phim blossom.mp4.
- Giải mã thông tin giấu trong đoạn phim blossom.mp4. Nhận xét về nội dung file giải mã được với file ban đầu (file/thông tin được chọn để giấu).

Đáp án:

Gợi ý:

Giao diện phần mềm

Tập tin blossom.mp4:



Kịch bản 03. Điều tra thông tin được ẩn giấu

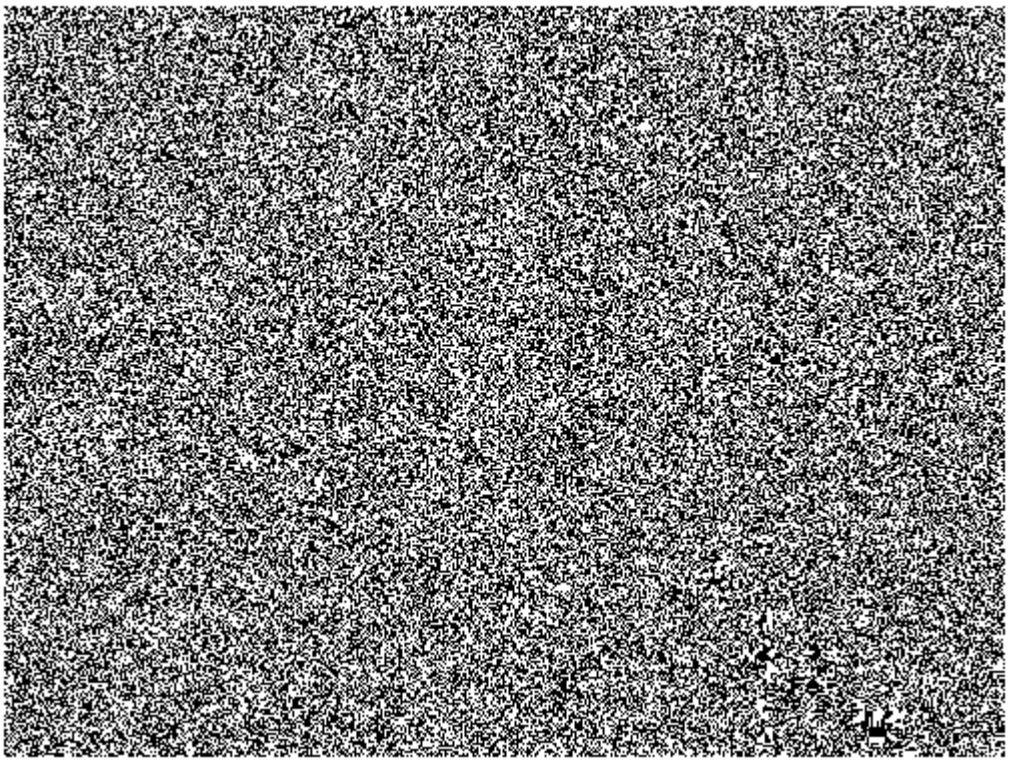
- Tài nguyên: kb03-suspicion.png
- Mô tả: Bức ảnh scan này đã được phục hồi từ các tập tin của một cựu nhân viên của Hiệp hội Ngớ ngẩn Miêu Quốc. Nhân viên điều tra cần phải tìm ra số sê-ri của máy in này để có thể xác định vị trí của thiết bị. Tìm số sê-ri của máy in.

Đáp án:

Hình ảnh đã cho được hiển thị như sau:

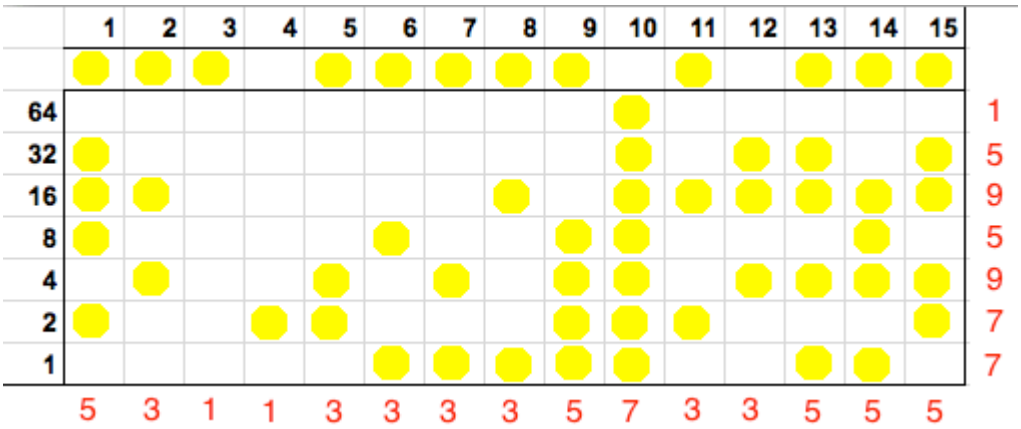
- Thử dùng lệnh strings để tìm chuỗi khả nghi. Quan sát.
- Sử dụng công cụ StegSolve để điều chỉnh bản màu của hình ảnh, với mong muốn tìm ra flag. Liên kết tải: https://github.com/eugenekolo/sec-tools/tree/master/stego/stegsolve/stegsolve
- Sau khi dùng công cụ này, ta được kết quả như sau:

FLAG{[illegible]}

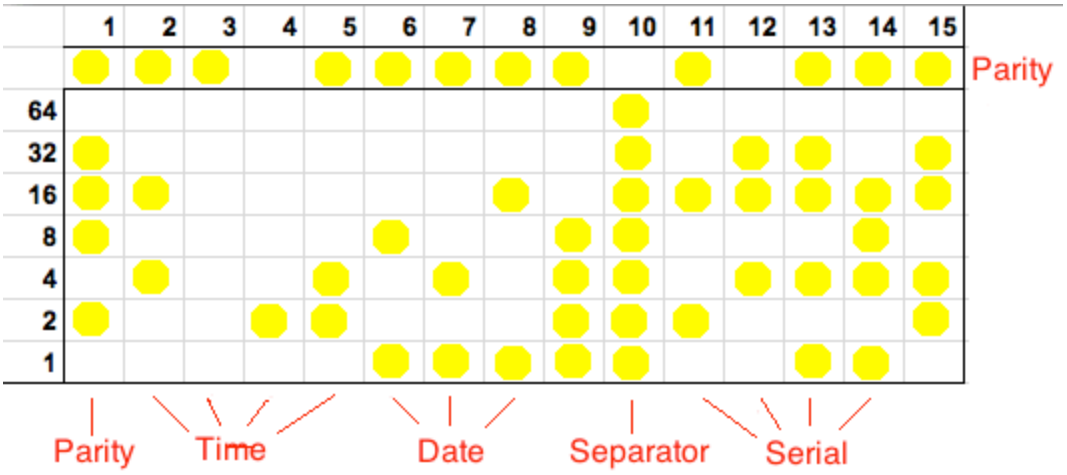


- Dựa vào thông tin được tìm thấy, ta thấy chuỗi thông tin có dạng như các ký tự trong ngôn ngữ Braille dành cho người khiếm thị. Chữ Braille được Louis Braille phát minh năm 1821. Mỗi chữ Braille được tạo thành từ 6 điểm, các điểm này được sắp xếp một trong khung hình chữ nhật gồm 2 cột và 3 dòng. Tập hợp các điểm nổi/chìm trong 6 vị trí sẽ tạo ra một bộ 64 (2^6) kiểu. Tuy nhiên, ở đây, chiều cao của cột là 8, khác với Braille. Điều khả nghi ở đây là gì?

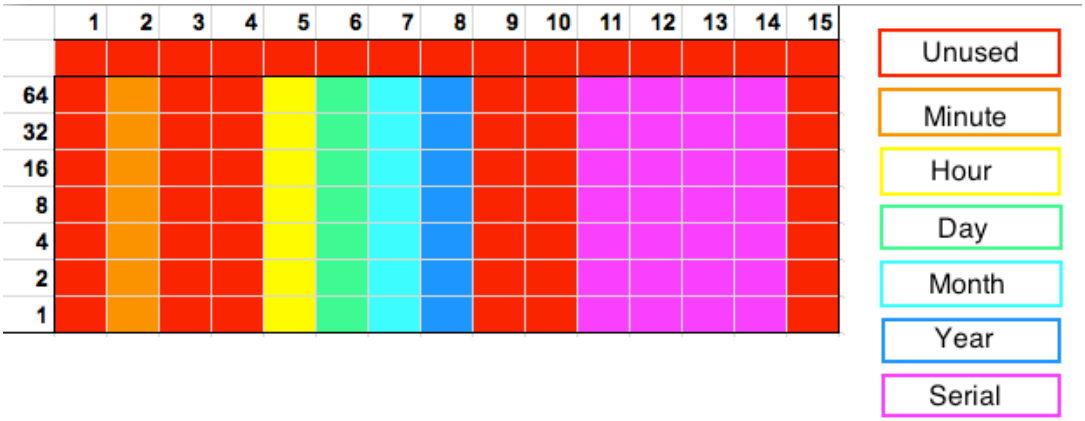
- Đọc lại thông tin mô tả của kịch bản, cần tìm thông tin sê-ri của máy in. Do đó, có thể nghĩ đến kỹ thuật Printer Steganography, tức là máy in sẽ in các chấm tròn theo cấu trúc để biểu diễn thông tin về nó. Đây chính là một kỹ thuật ẩn giấu thông tin được dùng phổ biến trong cuộc sống hiện đại ngày nay. Hầu hết tất cả các máy in Laser bao gồm của các hãng Brother, Canon, Dell, Epson, HP, IBM, Konica, Kyocera, Lanier, Lexmark, Ricoh, Toshiba, Xerox đều sử dụng kỹ thuật này.
- Hình dưới đây là một ví dụ về dữ liệu chấm tròn đại diện cho thông tin của máy in trên các tài liệu được in ra, theo kỹ thuật Printer Steganography.



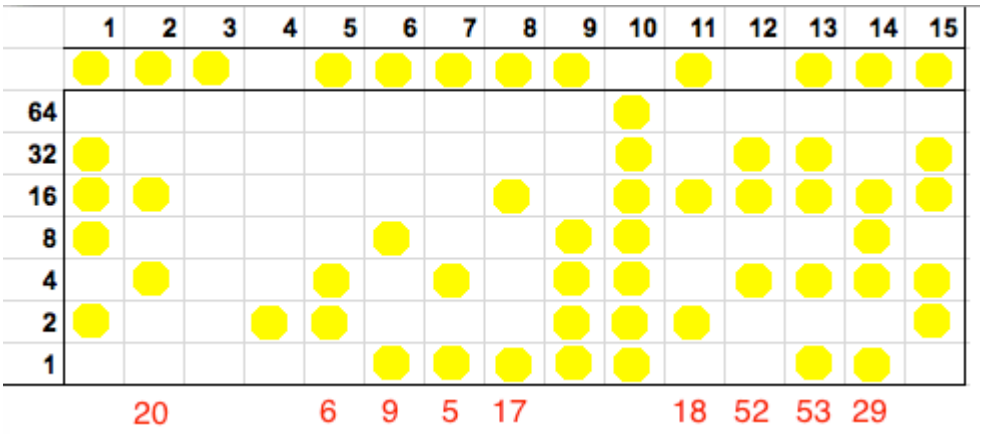
- Trong Printer Steganography, dữ liệu được biểu diễn dưới một bảng, trong đó ngoại trừ dòng trên cùng, tất cả các dòng và cột khác đều chứa số lẻ các chấm tròn (màu vàng). Các cột chỉ ra vùng lưu trữ dữ liệu và các hàng chỉ ra một số từ 0 đến 127. Cột số 9 và cột số 15 chưa/không sử dụng. Tương tự, các cột 3 và 4 cũng không được sử dụng cho việc hiển thị thời gian. Dòng 1 và cột 1 chỉ được dùng cho sự cân bằng (đảm bảo cho số chấm tròn của các hàng và cột luôn là số lẻ).



- Tóm lại các vùng dữ liệu như sau:



- Sau khi cộng các giá trị theo cột, ta có bảng sau:



- Giá trị các trường được đọc theo hướng từ phải sang trái. Ngoài ra, số sê-ri của máy in có thể bao gồm/ không bao gồm giá trị ở cột 14. Do đó, với bảng chấm tròn màu vàng theo như ví dụ này. Thông tin liên quan đến máy in khi in ra dữ liệu này như sau:
  - Printer Serial Number: 535218 (hoặc 29535218)

  - Date: 5/9/17 (tức ngày 9 tháng 5 năm 2017)
  - Time: 6:20

- Quay trở lại với kịch bản, hãy tìm số sê-ri của máy in để nhân viên điều tra xác định vị trí in tài liệu.

Đọc thêm: Lịch sử về kỹ thuật Printer Steganography

Bạn đã bao giờ in một lá thư nặc danh và gửi cho ai đó, sau đó hy vọng họ không thể theo dõi bạn từ nó hay chưa? Trong thực tế, lá thư này có thể dùng để theo dõi bạn trở lại. Điều này có nghĩa là thư nặc danh của bạn không còn vô danh nữa, do có sự tồn tại của một mã ẩn có giá trị duy nhất trên tờ giấy được in từ một máy in hiện đại ngày nay.

Vào thập niên 90 của thế kỉ trước, các phương pháp dùng cho việc nhận dạng máy in bởi các cơ quan chức năng ở Hoa Kỳ được nghiên cứu. Cụ thể, Hãng Xerox đã công bố và đưa thông tin nhận dạng vào việc in ấn trên các máy in của mình, nhằm phục vụ mục đích theo dõi những máy in này không được dùng để in các tài liệu giả mạo, có thể truy xuất nguồn gốc. Tuy nhiên, việc in các tài liệu có thể nhận dạng bằng số sê-ri ẩn cũng có thể có các lợi ích bảo mật khác và biện pháp này hiện nay đã được tất cả các nhà sản xuất máy in lớn áp dụng.

Trong khi tài liệu đang được in, các chấm nhỏ màu vàng được thêm vào mỗi trang giấy. Những chấm này hầu như không nhìn thấy được bằng mắt người. Các chấm màu vàng này chứa các dấu thời gian được mã hóa và số sê-ri liên kết các trang tài liệu giấy được in với 1 máy in cụ thể. Để xem những chấm này, cần sử dụng ánh sáng màu xanh hoặc kính lúp.

Ngược lại với những lợi ích mà kỹ thuật này mang lại, tội phạm hay các tổ chức khủng bố cực đoan có thể dùng phương pháp này để in ấn tài liệu, chứa các thông điệp bí mật cho hoạt động phạm tội của chúng. Do đó, khi bắt được tài liệu giấy ngay cả khi không chứa những gì bất thường, nó vẫn có thể ẩn giấu các thông tin quan trọng đằng sau đó.

B3. Kịch bản tổng hợp

Kịch bản 04. Điều tra thông tin được ẩn giấu

- Tài nguyên: star-wars.jpg
- Yêu cầu – Gợi ý: Bức ảnh được nhân viên điều tra tìm thấy trong một máy tính của một nghi phạm có sở thích xem ảnh của họa sĩ John Bramblitt. Tìm thông điệp được ẩn giấu, biết rằng thông điệp bắt đầu bằng "become".

Đáp án:

Kịch bản 05. Thực hiện phân tích, tìm thông tin ẩn giấu trong ảnh

- Tài nguyên thực hiện: qn001.jpg
- Yêu cầu – Gợi ý: Tìm thông điệp (flag) được ẩn giấu. Thông tin flag liên quan đến Đội tuyển bóng đá nam Việt Nam.

Gợi ý:

Kịch bản 06. Thực hiện phân tích:

- Tài nguyên: tiengiang003.jpg
- Yêu cầu – Gợi ý: Tìm thông điệp (flag) được ẩn giấu. Thuật toán dùng tìm ra flag liên quan đến việc thay thế các kí tự trong chuỗi ban đầu thành chuỗi chỉ gồm 2 kí tự a và b.

Đáp án:

Kịch bản 07. Thực hiện phân tích, tìm thông tin ẩn giấu:

- Tài nguyên: kb07-res (Tìm thông tin ẩn giấu trong Em-Gai-Mua-Huong-Tram.mp3, capture-the-flag.png)

Đáp án:

Kịch bản 08. Thực hiện phân tích, tìm thông tin ẩn giấu:

- Tài nguyên: LoveLetter.txt
- Yêu cầu – Gợi ý: Có gì đó đáng ngờ trong bức thư tình mà bạn đang đọc. Nhân viên điều tra cũng nghĩ rằng bức thư tình này chứa một thông điệp bí mật nào đó. Hãy tìm thông điệp được ẩn giấu (flag). Flag có dạng "FLAG-*"
- Link CTF: https://ringzer0ctf.com/challenges/215

Đáp án:

Kịch bản 09. Thực hiện phân tích, tìm thông tin ẩn giấu:

- Tài nguyên: transmission.png
- Yêu cầu – Gợi ý: Tìm thông điệp được ẩn giấu bằng các công cụ đã học trong buổi này.

Đáp án:

Kịch bản 10. Thực hiện phân tích, tìm thông tin ẩn giấu:

- Tài nguyên: thecatreturns.mp4
- Yêu cầu – Gợi ý: Tìm sự khác biệt giữa các khung hình (frame) trong đoạn phim đã cho. Chuyển nội dung đoạn phim thành các khung hình để phân tích. Công cụ ffmpeg, ImageJ.

Đáp án:

C. THAM KHẢO

Steganography: Forensic, Security, and Legal Issues (2008), Merrill Warkentin, Ernst Bekkering, Mark B. Schmidt, Journal of Digital Forensics, Security and Law.

https://www.garykessler.net/library/fsc_stego.html

https://people.duke.edu/~ng46/collections/steg-eurion-constellation.htm

https://www.cl.cam.ac.uk/~mgk25/eurion.pdf

https://www.cybrary.it/0p3n/printer-steganography/

D. YÊU CẦU

Bài thực hành được chia làm 2 phần riêng biệt.

- **Class Part (CP):** Sinh viên hoàn thành trên lớp (Bắt buộc).

  0% <= CP< 50%: 1đ

  50% <= CP < 90 %: 5đ

  90% <= CP <= 100%: 10đ

- **Home Part (HP):** Hoàn thành phần còn lại và làm báo cáo sau khi kết thúc buổi thực hành (nộp trên Course môn học theo deadline).
- Điểm Thực hành của mỗi Buổi (Session): **S = (CP + HP)/2**

– Sinh viên tìm hiểu và thực hành theo hướng dẫn.

– Nộp báo cáo kết quả gồm chi tiết những việc bạn đã thực hiện, quan sát thấy và kèm ảnh chụp màn hình kết quả (nếu có); giải thích cho quan sát (nếu có).

– Sinh viên báo cáo kết quả thực hiện và nộp bài.

Báo cáo:

– File .DOCX và .PDF. Tập trung vào nội dung, không mô tả lý thuyết.

– Chỉ dùng duy nhất 1 loại Font chữ (Times New Roman – cỡ chữ 12)

– Đặt tên theo định dạng: [Mã lớp]-SessionX_GroupY. (trong đó X là Thứ tự buổi Thực hành, Y là số thứ tự Nhóm Thực hành đã đăng ký với GVHD-TH).

Ví dụ: [NT101.H11.1]-Session1_Group3.

– Nếu báo cáo có nhiều file, nén tất cả file vào file .ZIP với cùng tên file báo cáo.

– Không đặt tên đúng định dạng – yêu cầu, sẽ **KHÔNG** chấm điểm bài Lab.

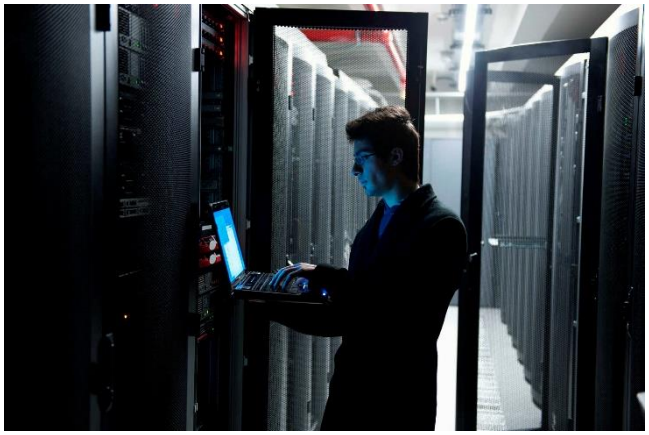
– Nộp file báo cáo trên theo thời gian đã thống nhất tại courses.uit.edu.vn.

Đánh giá: Sinh viên hiểu và tự thực hiện được bài thực hành. Khuyến khích:

– Chuẩn bị tốt và đóng góp tích cực tại lớp.

– Có nội dung mở rộng, ứng dụng trong kịch bản phức tạp hơn, có đóng góp xây dựng bài thực hành.

Bài sao chép, nộp trễ, thực hiện không nghiêm túc ... sẽ được xử lý tùy mức độ vi phạm.



HẾT