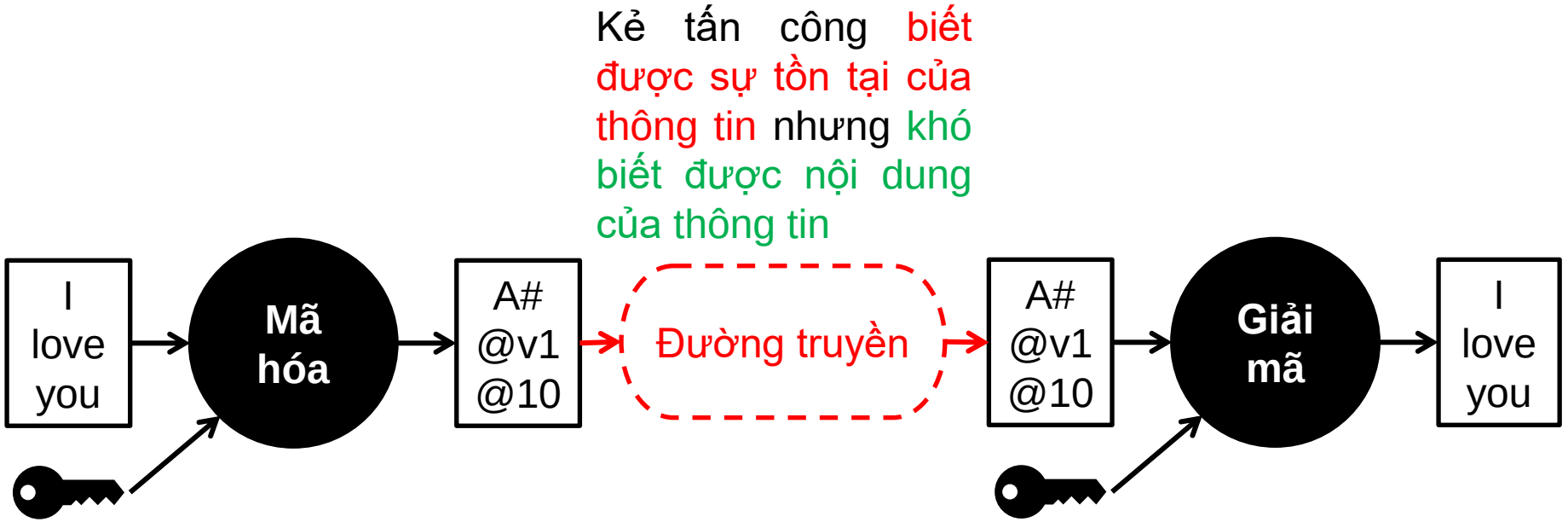
Ẩn dữ liệu & chia sẻ thông tin mật (Information hiding & secret sharing)

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Nhu cầu truyền thông an toàn

Giả sử Bob muốn gửi tin cho Alice thông qua đường truyền Internet. Làm sao để thông tin truyền đi được an toàn?

Giải pháp 1: **mã hóa** thông tin cần truyền

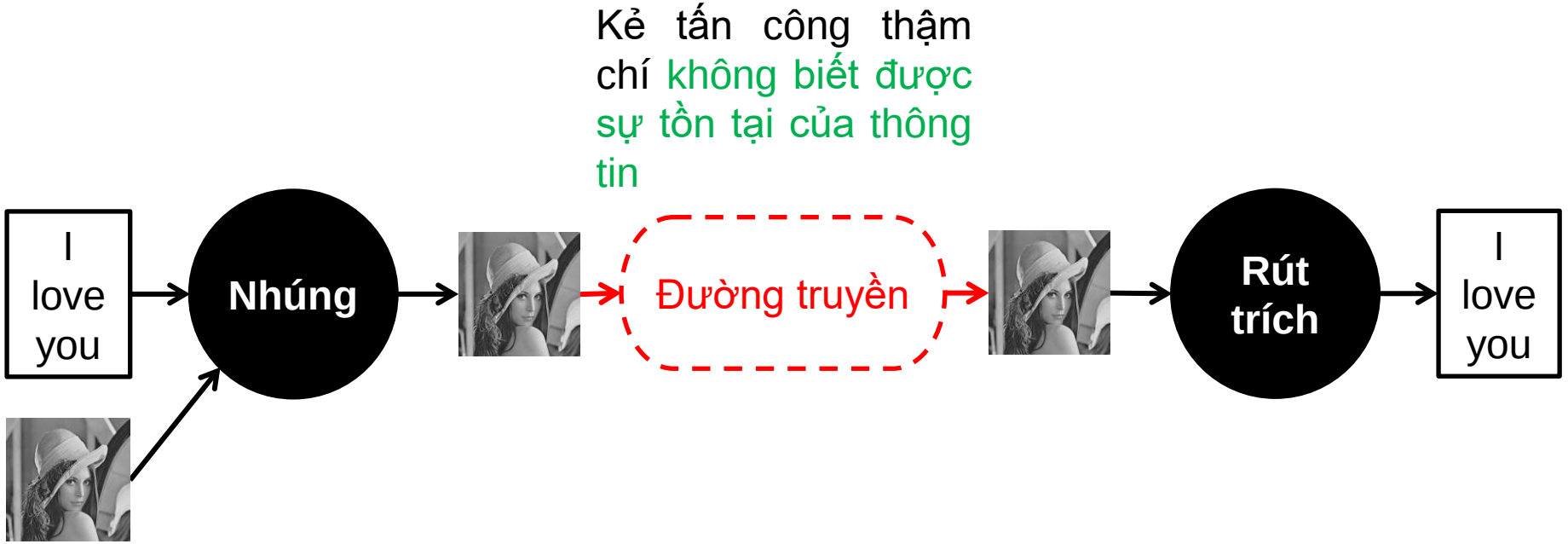


Nhu cầu truyền thông an toàn

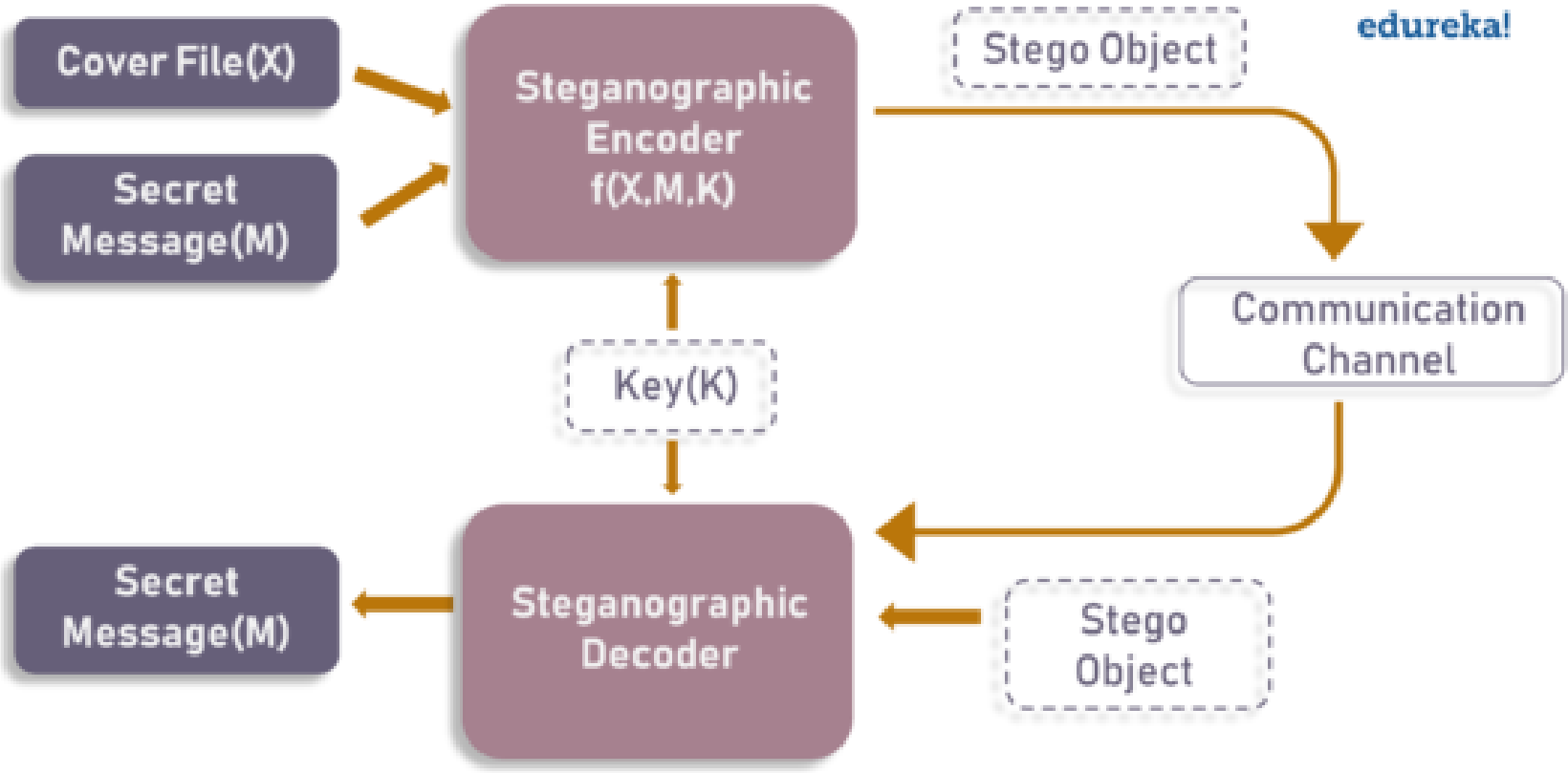
Giả sử Bob muốn gửi tin cho Alice thông qua đường truyền Internet. Làm sao để thông tin truyền đi được an toàn?

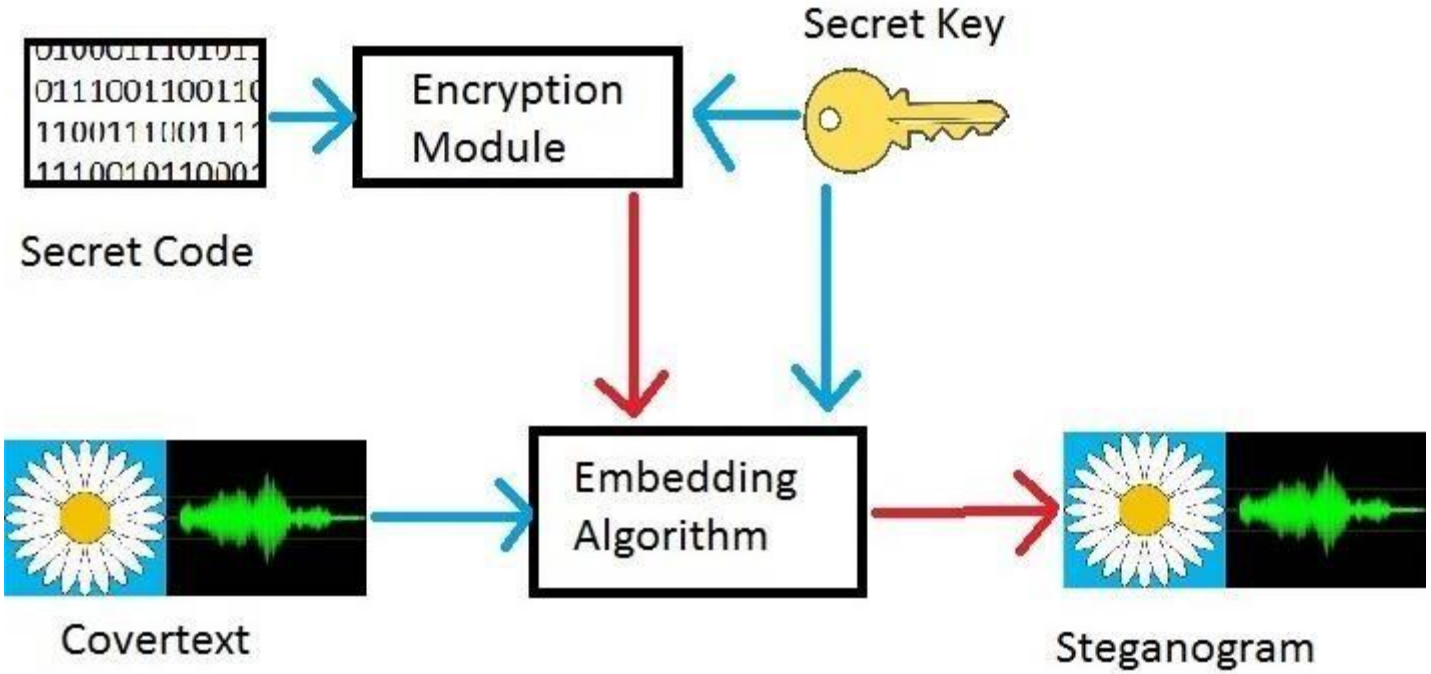
Giải pháp 2: **ẩn** thông tin cần truyền trong một thông tin khác

Hướng tiếp cận này được gọi là **steganography (covered writing)** và được xem là một nhánh của **ẩn dữ liệu - data/information hiding**



Steganography




Secret Key
Encryption Module
Secret Code
Embedding Algorithm
Covertext
Steganogram

Nói thêm về steganography

- Đã có từ thời xa xưa, làm trên các đối tượng vật lý
  - Cạo trọc đầu những người truyền tin, xăm thông tin cần truyền, rồi chờ tóc mọc lại để che hình xăm
  - Dùng mực vô hình (ví dụ, chanh) viết lên giấy
  - In siêu nhỏ thông tin cần truyền, khi nhìn bằng mắt thường chỉ là một dấu chấm trong một giấy tờ bình thường nào đó
  - ...
- Thời nay chủ yếu là ẩn dữ liệu trên các đối tượng kỹ thuật số như file văn bản, file hình ảnh, file âm thanh, file video, ...
  - Người ta đã phát hiện ra có những kẻ khủng bố trao đổi ngầm với nhau thông qua những bức ảnh trên Amazon, eBay, ...
  - Liệu rằng những ảnh hay status trên facebook có chứa thông tin ẩn?
  - Liệu rằng slide của Thầy có chứa thông tin ẩn?

	STEGANOGRAPHY	CRYPTOGRAPHY
Definition	It is a technique to hide the existence of communication	It's a technique to convert data into an incomprehensible form
Purpose	Keep communication secure	Provide data protection
Data Visibility	Never	Always
Data Structure	Doesn't alter the overall structure of data	Alters the overall structure of data
Key	Optional, but offers more security if used	Necessary requirement
Failure	Once the presence of a secret message is discovered, anyone can use the secret data	If you possess the decryption key, then you can figure out original message from the ciphertext

Nhu cầu truyền thông an toàn

Giả sử Bob muốn gửi tin cho Alice thông qua đường truyền Internet. Làm sao để thông tin truyền đi được an toàn?

Giải pháp 3: **mã hóa** thông tin cần truyền rồi **ẩn** trong một thông tin khác

Nhu cầu bảo vệ bản quyền cho các tác phẩm kỹ thuật số

Giả sử bạn chụp được một bức ảnh đẹp và đưa lên mạng. Những người khác sẽ có thể dễ dàng sao chép ảnh của bạn, thậm chí có thể nói ảnh là của họ, và sử dụng cho lợi ích của cá nhân họ. Làm sao để hạn chế vấn đề này?

Một giải pháp: **thủy vân số - digital watermarking** (đây là một nhánh khác của ẩn dữ liệu)

- Nhúng watermark là thông tin của chủ sở hữu vào tác phẩm
- Yêu cầu nhúng: **bền vững (robust)** - kẻ tấn công phải khó có thể xóa hoặc chỉnh sửa watermark

Nguồn ảnh: http://www.aoaophoto.com/watermark-images.htm

Nhu cầu bảo vệ bản quyền cho các tác phẩm kỹ thuật số

Giả sử bạn là nhà xuất bản sách điện tử, bạn bán nhiều bản sao của một file sách cho nhiều khách hàng. Một khách hàng có thể tự ý tạo ra các bản sao, gửi cho người này người kia; cuối cùng, kết quả là file sách xuất hiện trên mạng và mọi người có thể tự do down. Làm sao để hạn chế vấn đề này?

Một giải pháp: **thủy vân số - digital watermarking**

- Khi bán ra một bản sao cho một khách hàng, nhúng vào đó một mã dành riêng cho khách hàng đó. Nếu file sách bị rò rỉ trên mạng, nhà xuất bản có thể rút trích ra mã để biết được nguồn rò rỉ là từ khách hàng nào.
- Yêu cầu nhúng: **bền vững** và **vô hình**

Steganography vs watermarking

Trong 3 tiêu chí vô hình – bền vững – sức chứa

- Steganography thường đặt nặng vào tiêu chí vô hình và sức chứa
- Watermarking thường đặt nặng vào tiêu chí bền vững và vô hình

Nội dung môn học

- Ẩn dữ liệu (data/information hiding)
  - Steganography
  - Watermarking
- Chia sẻ thông tin mật (secret sharing)

- trên văn bản
- trên hình ảnh
- trên âm thanh

Nhu cầu lưu trữ thông tin mật

Giả sử bạn muốn lưu trữ một thông tin bí mật (password, mã phóng tên lửa, ...). Ngoài mong muốn thông tin mật không bị người khác biết được, bạn cũng mong muốn thông tin mật sẽ không bị mất mát

Giải pháp 1: lưu một bản duy nhất ở một nơi (ví dụ, lưu trong máy tính)

- Thông tin mật khó bị người khác biết được
- Nhưng dễ bị mất: lỡ xui nơi lưu trữ duy nhất đó "hy sinh" (ví dụ, máy tính bị hư)

Nhu cầu lưu trữ thông tin mật

Giả sử bạn muốn lưu trữ một thông tin bí mật (password, mã phóng tên lửa, ...). Ngoài mong muốn thông tin mật không bị người khác biết được, bạn cũng mong muốn thông tin mật sẽ không bị mất mát

Giải pháp 2: lưu nhiều bản sao ở nhiều nơi

- Lưu càng nhiều bản thì thông tin mật sẽ càng khó bị mất
- Nhưng lại tăng nguy cơ thông tin mật bị người khác biết được (kẻ tấn công chỉ cần lấy được thông tin mật ở một nơi lưu trữ)

Nhu cầu lưu trữ thông tin mật

Giả sử bạn muốn lưu trữ một thông tin bí mật (password, mã phóng tên lửa, ...). Ngoài mong muốn thông tin mật không bị người khác biết được, bạn cũng mong muốn thông tin mật sẽ không bị mất mát

Giải pháp 3: **chia sẻ thông tin mật (secret sharing)** - chia thông tin mật ra thành n phần và lưu ở n nơi, để khôi phục lại được thông tin ban đầu cần phải có ít nhất k phần bất kỳ ($k \leq n$ và do người dùng chỉ định, ví dụ $k = \frac{n}{2}$)

- Thông tin mật khó bị mất: chỉ khi có hơn $n - k$ nơi lưu trữ "hy sinh" thì thông tin mật mới bị mất
- Thông tin mật cũng khó bị người khác biết được: kẻ tấn công phải lấy được ít nhất là k phần từ k nơi thì mới có thể lấy được thông tin mật

Nói thêm về chia sẻ thông tin mật

Chia sẻ thông tin mật cũng có thể được dùng khi ta có nhu cầu chia sẻ một thông tin mật cho một nhóm người

- Ta mong muốn chỉ khi có đủ một số lượng người nhất định thì mới có thể khôi phục lại được thông tin mật
- Ta cũng mong muốn nếu có một số lượng người gặp sự cố thì vẫn có thể khôi phục lại được thông tin mật từ những người còn lại

Hiding a Message inside Text

- **Partially effective**

randoM capitalosis is a rarE disEase ofTen contrAcTed by careless inTernet users. tHis sad illnEss causes the aFfected peRsON To randomly capitalize letters in a bOdy oF texT. please do not confuse this disease witH a blatant attEmpt aT steganogRAPhy.

Reveals: MEET AT THE FRONT OF THE TRAP

Hiding a Message inside Images

- Phương pháp phổ biến nhất
- ***Least-significant bit (LSB) modifications***
  - *24-bit vs. 8-bit images*
  - *Tools to implement LSB: EzStego and S-Tools*
- ***Masking and Filtering***
- ***Algorithms and Transformations***

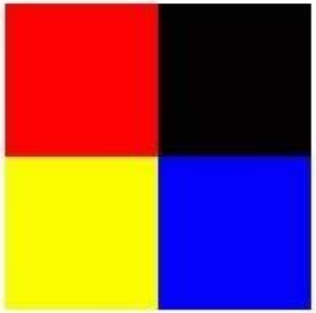
Least Significant Bit Steganography

Original Image

```
11111111  00000000
00000000  00000000
00000000  00000000

11111111  00000000
11111111  00000000
00000000  11111111
```

Stego Image



```
11111101  00000011
00000010  00000001
00000000  00000010

11111100  00000011
11111101  00000001
00000001  11111100
```

c	a	t
01 10 00 11	01 10 00 01	01 11 01 00

Steganography tools

1. Anubis
2. BMP Secrets
3. DarkCryptTC
4. OpenPuff
5. OpenStego
6. StegFS
7. StegoShare

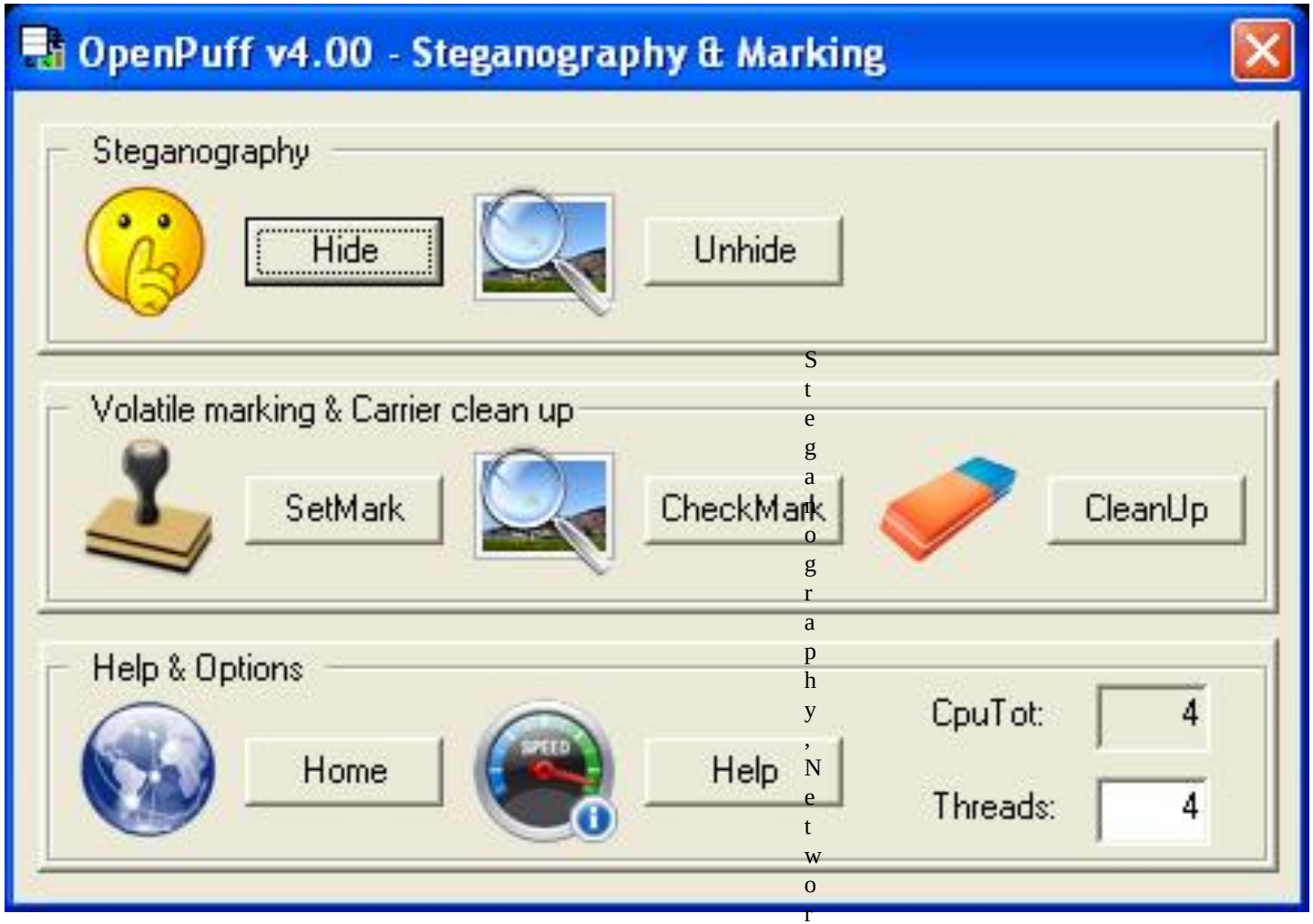
☐Other tools: https://en.wikipedia.org/wiki/Steganography_tools

☐12 best steganography tools: http://www.topbestalternatives.com/best-steganography-software/

Steganography tools (contt.)

StegoShare allows embedding of large files into multiple images

Steganography tools (contt.)



Free tool that implements three layers of hidden data

Đánh giá môn học

- **Các bài tập thực hành (30%)**
  - Ngôn ngữ lập trình: Python
  - Làm trên IPython Notebook (cho phép vừa có thể soạn thảo văn bản, vừa có thể viết và chạy Python code)
- **Seminar lý thuyết (20%)**
  - Làm nhóm, 2-3 SV / nhóm
  - Công việc: đọc hiểu tài liệu (tiếng Anh) về một chủ đề và trình bày lại trước lớp
  - Thời gian trình bày: vào các buổi học lý thuyết cuối
- **Thi lý thuyết cuối kỳ (50%)**
  - Trắc nghiệm + tự luận

Đánh giá môn học

- Nên nhớ mục tiêu chính ở đây là học, học một cách chân thật. Bạn có thể thảo luận ý tưởng với bạn khác, nhưng code và bài làm phải là của bạn, dựa trên sự hiểu của bạn.
- Nếu vi phạm thì sẽ bị 0 điểm cho toàn bộ môn học.

Lời khuyên

- Đi học đầy đủ
- Ghi chép bài
- Lấy khó làm niềm vui
- Hạn chế những việc không cần thiết, tập trung vào những việc cần thiết