

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
KHOA AN TOÀN THÔNG TIN**



**HỌC PHẦN: CÁC KỸ THUẬT GIẤU TIN
MÃ HỌC PHẦN: INT14102**

Chủ đề: Giấu tin trong âm thanh

Lab: stego_audio_code_parity-coding

Sinh viên thực hiện:	Lê Tuấn Anh
Mã sinh viên:	B21DCAT028
Nhóm:	04
Giảng viên hướng dẫn:	Đỗ Xuân Chợt

HÀ NỘI 2025

Bài lab Các kỹ thuật giấu tin: stego_code_stsm

1. Mục đích

Giúp sinh viên hiểu được thuật toán giấu tin trong âm thanh sử dụng phương pháp mã hoá chẵn lẻ (parity coding)

2. Yêu cầu đối với sinh viên

Quen thuộc với hệ điều hành Linux và có kiến thức về kỹ thuật giấu tin.

3. Nội dung lý thuyết

Parity Coding (mã hóa chẵn lẻ) là một kỹ thuật giấu tin trong âm thanh dựa trên việc điều chỉnh tính chẵn lẻ của một nhóm mẫu (sample) để mã hóa từng bit của thông điệp bí mật. Thay vì thay đổi trực tiếp giá trị của từng mẫu, kỹ thuật này nhóm các mẫu lại thành khối, tính tổng số bit 1 của nhóm và điều chỉnh bit cuối nếu cần, để đảm bảo biểu diễn chính xác thông điệp muốn giấu.

Dữ liệu âm thanh (thường là file WAV) sẽ được chia thành các nhóm mẫu. Một nhóm có thể gồm 3, 4 hoặc 5 mẫu liên tiếp, tùy theo thiết kế. Mỗi nhóm dùng để giấu 1 bit thông điệp.

Trong mỗi nhóm, ta trích bit LSB (least significant bit – bit cuối) của từng mẫu, tạo thành chuỗi bit con. Sau đó, tính tổng số bit 1 trong chuỗi đó để xác định parity:

- Nếu số bit 1 là chẵn \rightarrow parity = 0
- Nếu số bit 1 là lẻ \rightarrow parity = 1

Sau khi tính parity, ta so sánh với bit thông điệp cần giấu:

- Nếu parity đã trùng với bit thông điệp \rightarrow không cần thay đổi
- Nếu khác nhau, ta lật (flip) LSB của một mẫu bất kỳ trong nhóm để đổi từ chẵn sang lẻ (hoặc ngược lại)

Sau khi giấu hết các bit, các nhóm được ghép lại để tạo thành file âm thanh mới chứa thông điệp.

4. Nội dung thực hành

Khởi động bài lab:

Vào terminal, gõ:

labtainer stego_audio_code_parity-coding

(Chú ý: Sinh viên sử dụng mã sinh viên của mình để nhập thông tin email người thực hiện bài lab khi có yêu cầu, để sử dụng khi chấm điểm)

Nhiệm vụ 1: Giấu tin vào file âm thanh sử dụng Parity Coding

- Sau khi khởi động xong bài lab, có 2 terminal Bob và Alice
- Trong terminal Alice, đã có sẵn một file âm thanh input.wav, một file code python dùng để giấu tin vào trong file âm thanh.
- Kiểm tra file code python dùng để giấu tin, sửa tên file âm thanh đầu vào thành input.wav. Sinh viên có thể sử dụng câu lệnh:

nano hide.py

- Tạo một file message.txt chứa nội dung cần giấu
- Sau khi kiểm tra file code python, sinh viên tiến hành sử dụng file code để giấu tin trong file message.txt vào file âm thanh input.wav. Sinh viên sử dụng câu lệnh:

python3 hide.py message.txt sounds.wav secret.wav

Nhiệm vụ 2: Giải mã thông tin đã giấu trong file âm thanh

- Sau khi giấu tin thành công, chuyển file âm thanh đã được giấu tin sang cho bob bằng câu lệnh, mật khẩu là “ubuntu”:

scp secret.wav ubuntu@<ip bob>:/home/ubuntu

- Kiểm tra file code python dùng để giải mã, sửa tên file âm thanh đầu vào thành output.wav. Sinh viên có thể sử dụng câu lệnh:

nano extract.py

- Thực hiện giải mã thông điệp đã được giấu bằng câu lệnh

python3 extract.py secret.wav result.txt

Nhiệm vụ 3: Đọc file thông tin đã được giấu trong file âm thanh

- Sau khi thực hiện giải mã thông điệp đã được giấu trong file âm thanh, trích xuất thành công thông điệp ra file message.txt, sinh viên tiến hành đọc thông tin đã được giấu trong âm thanh:

cat result.txt

Nhiệm vụ 4: So sánh file âm thanh gốc với file âm thanh đã giấu tin bằng Parity Coding

- Mục đích của nhiệm vụ này là tìm hiểu sự khác biệt giữa 2 file wav trước và sau giấu tin. Tại terminal của Alice, sử dụng công cụ soxi để xem thông tin kỹ thuật của 2 file âm thanh

soxi sounds.wav

soxi secret.wav

- Kết quả cho thấy 2 file có cùng thông số kỹ thuật. Tuy nhiên, điều đó không có nghĩa là nội dung âm thanh giống nhau, do đó cần so sánh raw

data của 2 file âm thanh. Chúng ta thử xác định xem raw data của 2 file wav này khác nhau bắt đầu từ byte nào. Sử dụng lệnh

cmp -l <(sox sounds.wav -t .raw -) <(sox secret.wav -t .raw -) | head -n 20

- Kết quả cho thấy ký tự thứ 5 của dòng 1 là nơi đầu tiên xảy ra sự khác biệt. Như vậy, mặc dù 2 file có cùng định dạng và thông số kỹ thuật nhưng nội dung 2 file khác nhau hoàn toàn.

Kết thúc bài lab:

Kiểm tra checkwork:

checkwork

Trên terminal đầu tiên sử dụng câu lệnh sau để kết thúc bài lab:

stoplab