

Bài thực hành: Giấu tin trong văn bản bằng phương pháp thay đổi kích thước của các ký tự

1. Mục đích

- Giúp sinh viên hiểu về phương pháp giấu tin trong văn bản bằng phương pháp thay đổi kích thước các ký tự.

2. Yêu cầu đối với sinh viên

- Sinh viên có kiến thức xử lý chuỗi, file trong Python

3. Nội dung thực hành

- Tải bài lab: Sinh viên tải bài lab từ link github sau:

<https://github.com/Mbr-clown-lord/stego-fontsize-modification>

- Khởi động bài lab:

labtainer -r stego-fontsize-modification

(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)

Ý tưởng của phương pháp giấu tin bằng việc thay đổi kích thước của ký tự: Các văn bản định dạng .docx có thể định dạng riêng từng ký tự, bao gồm dấu câu. Khi thay đổi kích thước font của dấu câu (ví dụ 11pt thành 11.5pt), sự thay đổi nhỏ này gần như không thể nhận ra bằng mắt thường. Tuy nhiên thông tin sẽ bị mất nếu văn bản được đánh tay hoặc định dạng lại. Chính vì vậy phương pháp này không được đánh giá cao và chỉ phù hợp với văn bản in.

Thuật toán giấu tin:

Đầu vào:

- Thông điệp bí mật
- Văn bản phủ

Các bước thực hiện:

- Bước 1: Chuyển thông điệp bí mật thành dạng nhị phân

- Bước 2: Đọc dạng nhị phân của thông tin bí mật và văn bản phủ để tìm các dấu câu có trong văn bản sau đó thay đổi kích thước của dấu câu theo quy ước: bit 0(giảm đi 0.5pt), bit 1(tăng lên 0.5pt).

Đầu ra:

- Văn bản phủ có chứa thông điệp

Sau khi khởi động xong một terminal ảo có tên stego-fontsize-modification xuất hiện.

(chú ý: sau khi khởi động bài lab thành công, sinh viên cần cài đặt *python-docx* để hoàn thành cách nhiệm vụ của bài lab)

pip3 install python-docx

3.1. Nhiệm vụ 1: Chuẩn bị văn bản giấu tin

Yêu cầu:

- Thông điệp cần giấu: Số cuối trong mã sinh viên
- Sẽ có 8 bit nhị phân cần giấu vì thế văn bản docx phải có ít nhất 8 giấu câu. Chạy chương trình *check_text.py* để kiểm tra nội dung file *input.docx* có đủ để giấu 8 bit nhị phân của số cuối mã sinh viên không.

python3 check_text.py

- Dựa vào thông báo chương trình đưa ra, sửa nội dung file *input.docx* và chạy lại tới khi chương trình đưa ra thông báo đủ điều kiện để giấu tin.

Gợi ý: Sử dụng LibreOffice(vì LibreOffice không có giao diện đồ họa nên sinh viên cần sử dụng chế độ *--headless* để chuyển file sang dạng file *.txt* sửa bằng công cụ nano sau đó chuyển lại file docx để giấu tin)

3.2. Nhiệm vụ 2: Giấu thông điệp

Yêu cầu:

- Sau khi có được file *input.docx* thỏa mãn điều kiện giấu tin ở Task 1, sinh viên tiến hành giấu tin bằng cách chạy chương trình *stego.py*:

python3 stego.py

- Kết quả thu sau khi dấu tin sẽ được lưu vào file *output.docx*
- Đề xem sự khác biệt của văn bản trước và sau khi giấu tin, sinh viên cần dùng LibreOffice với giao diện đồ họa.

Gợi ý: Sử dụng ssh sao chép 2 file *input.docx* và *output.docx* sang máy host(máy đang chạy Labtainer) và mở file bằng LibreOffice của máy host)

- Để kiểm tra kết quả của quá trình giấu tin có thành công hay không, sinh viên chạy chương trình decode.py:
python3 decode.py
- Kết quả thành công nếu kết quả ra đúng số cuối trong mã sinh viên.
- Để kiểm tra việc thông tin sẽ bị mất khi văn bản bị định dạng lại, sinh viên tiến hành định dạng lại file output.docx và chạy lại chương trình decode.py.(Sử dụng LibreOffice ở chế độ --headless)
- Để tránh việc file docx sau khi giấu tin sẽ bị sửa đổi, có thể chuyển sang định dạng file pdf bằng LibreOffice:

libreoffice --headless --convert-to pdf <tên file>.docx

3.3. Nhiệm vụ 3: Tách thông điệp

- Yêu cầu: File stego.docx là file đã giấu thông điệp “Win”. Giả sử sinh viên chưa biết thông điệp được giấu trong file là gì mà chỉ biết quy ước giấu tin là quy ước đã nêu ở phần trước. Hãy sửa lại file decode.py để giải mã ra thông điệp.

Gợi ý: Đọc và tìm hiểu kỹ cách thức hoạt động của decode.py

- Kết thúc bài lab:

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

stoplab stego-fontsize-modification

- o Khi bài lab kết thúc, một tệp zip lưu kết quả được tạo và lưu vào một vị trí được hiển thị bên dưới stoplab.

- Khởi động lại bài lab:

- o Trong quá trình làm bài sinh viên cần thực hiện lại bài lab, dùng câu lệnh:

labyrinth -r stego-fontsize-modification