Labtainer : Giấu tin trong văn bản- phương pháp ngữ nghĩa-trich rút câu, thực hiện tấn công MitM

1. **Mục đích**
   - Bài thực hành này giúp sinh viên nắm bắt được phương pháp giấu tin trong văn bản bằng cách sử dụng từ đồng nghĩa
     - Sử dụng **encode.py** để giấu thông điệp vào văn bản
     - Sử dụng **decode.py** để giải mã thông điệp mà bên gửi gửi cho bên nhận
     - Sử dụng **python3.py** để phát hiện văn bản có giấu tin hay không trên máy attacker.

File **spoof_tcp.py** là một **script Python** sử dụng thư viện **Scapy** để thực hiện **tấn công mạng** dạng **man-in-the-middle (MITM)**, cụ thể là:

- **ARP Spoofing (Giả mạo ARP)**:
  - Script gửi các gói **ARP** giả mạo tới **A** (clienta, 182.20.0.10) để lừa **A** rằng **C** (attackerc, 182.20.0.30) là **B** (clientb, 182.20.0.20).
  - Kết quả: Bảng ARP của **A** ánh xạ IP của **B** (182.20.0.20) tới địa chỉ MAC của **C** (02:42:b6:14:00:1e), khiến các gói tin từ **A** gửi đến **B** thực chất được gửi tới **C**.
- **IP Spoofing (Giả mạo IP)**:
  - Khi **A** gửi gói tin TCP tới **B** (qua port 1234), **C** giả mạo IP của **B** (182.20.0.20) để trả lời, khiến **A** nghĩ rằng nó đang giao tiếp với **B**.
  - Script xử lý giao thức TCP, bao gồm **SYN**, **SYN-ACK**, **ACK**, và dữ liệu.
- **Chặn và ghi dữ liệu**:
  - Script chặn các gói tin TCP từ **A** gửi đến **B** qua port 1234 (thường là dữ liệu từ lệnh nc 182.20.0.20 1234 < text2.txt trên **A**).
  - Dữ liệu trong các gói tin (VD: nội dung text2.txt như I love coding. Python is funa...) được ghi đè vào file /home/ubuntu/text2.txt trên **C** mỗi khi nhận được gói tin chứa dữ liệu

2. **Nội dung lý thuyết**
   - Tìm hiểu về phương pháp giấu tin trong văn bản sử dụng trích rút câu
   - Tìm hiểu về cấu trúc code python.
   - Tìm hiểu về ARP Spoofing và IP spoofing
3. **Các bước thực hiện**

3.1. Khởi động bài lab

3.1.1. Các bước cần thực hiện

Thực hiện cài đặt lab tại**: https://github.com/B21DCAT076/KTGT**

Sinh viên thực hiện tải file steganography-semantic-medium.rar và giải nén ra

-Sau đó thực hiện copy file giải nén vào thư mục**: labtainer/trunk/labs/**

Tại Terminal của labtainer gõ lệnh

```
`labtainer –r stegano-setence-extraction2  `
```

(chú ý: sinh viên sử dụng <TÊN TÀI KHOẢN HỆ THỐNG> của mình để nhập thông tin email người thực hiện bài lab khi có yêu cầu, để sử dụng khi chấm điểm)

Tại **Terminal** của máy sẽ hiện ra 3 máy , 1 là Client, 2 là Server, 3 là Attacker, chúng có chung 1 mạng con, tiến hành kiểm tra địa chỉ IP của 3 máy bằng **ifconfig**

Trên máy client có: file encode.py: đây là code mã hóa, text1.txt(văn bản phủ), text2.txt (văn bản mã hóa)

Trên Server: file decode.py: code giải mã, text2.txt(nhận được từ text2.txt(client) bằng nc)

Trên Attacker: file python3.py: code phát hiện giấu tin hay không, text2.txt (chứa nội dung file text2.txt mà client gửi server), File **spoof_tcp.py** là một **script Python** sử dụng thư viện **Scapy** để thực hiện **tấn công mạng** dạng **man-in-the-middle (MITM)**

-Trên máy client, server, attacker thực hiện lệnh ping để đảm bảo chúng có kết nối mạng và nằm cùng 1 mạng con

Task 1: Thực hiện Giau Tin

-Trên Client ta quan sát bằng lệnh ls có 1 file encode.py(đây là file code mã hóa) và 2 file văn bản là text1.txt(file văn bản phủ) và text2.txt(file văn bản mã hóa sau khi chạy code python1.py)

Trên client tiến hành chạy file python1.py bằng lệnh:

python3 encode.py

Khi được hỏi nhập thông điệp cần giấu, bạn hãy nhập :**I am Ptiter**

Quan sát kết quả đầu ra với file text2.txt và so sánh text1.txt để xem sự khác biệt

Task 2 : Thực hiện truyền file sang Server:

-Trên Server ta thực hiện lệnh và giữ nguyên kết nối

nc –l –p 1234 > text2.txt

- Trên Attacker:

Ta thực hiện lệnh: Sudo nmap <ip máy server> để biết cổng nào đang mở trên Server.

Đảm bảo sudo sysctl -w net.ipv4.ip_nonlocal_bind=1

sudo sysctl -w net.ipv4.ip_forward=0

-sau đó ta thự hiện lệnh : **sudo python3 spoof_tcp.py**

<mục đích: bắt được file.txt mà Client gửi cho Server rồi gán vào text2.txt >

-Trên Client ta thực hiện lệnh:

nc <ip máy server> 1234 < text2.txt

<do attacker thực hiện tấn công MitM nên text2.txt sẽ được gửi cho Attacker>

Task3: Kiểm tra Attacker có bắt được file.txt Client gửi cho server:

Dùng lệnh: cat text2.txt

Task 4: Phát hiện giấu tin

-Trên **Attacker** tiến hành chạy file python3.py với đầu vào là file text2.txt bằng lệnh:

python3 python3.py

-PP phát hiện giấu tin trong văn bản bằng từ đồng nghĩa chủ yếu dựa trên ngữ nghĩa, cấu trúc ngữ pháp, lỗi logic câu, lỗi nghĩa diễn đạt câu.

Task5: Attacker thực hiện sửa file text2.txt(văn bản mã hóa) rồi gửi cho Server nhằm thay đổi thông điệp:

- sudo chmod 666 text2.txt
- nano text2.txt

- Sau đó ta tiến hành sửa 3 câu cuối văn bản mã hóa bằng cách thêm giới từ **a**.

-Sau đó chạy lệnh: nc 182.20.0.20 1234 < text2.txt

-Sau đó đóng nc bằng **Ctrl+C**

Trên máy server: ta tiến hành lệnh:

python3 decode.py

Quan sát thông điệp mã hóa và so sánh với thông điệp gốc

3.2 Kết thúc bài lab:

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

stoplab

- Khi bài lab kết thúc, một tệp lưu kết quả được tạo và lưu vào một vị trí được hiển thị bên dưới stoplab.
- Sinh viên cần nộp file *.lab* để chấm điểm.
- Để kiểm tra kết quả khi trong khi làm bài thực hành sử dụng lệnh: *checkwork <tên bài thực hành>*

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

labtainer -r stegano-setence-extraction2