Phân tích gói tin HTTP với Wireshark

Sniffing HTTP Traffic with Wireshark

Môn học: Nhập môn Mạng máy tính

Tái bản lần 3 - Tháng 09/2019
Lưu hành nội bộ

A. TỔNG QUAN

1. Mục tiêu

- Tìm hiểu cách tự xây dựng một website đơn giản.
- Sử dụng Wireshark để bắt gói tin HTTP để phân tích các đặc điểm của gói tin này: Thông điệp GET/response, cấu trúc của HTTP header, truy cập các file HTML dài, truy cập các file HTML có đính kèm các đối tượng, xác thực HTTP và bảo mật,…

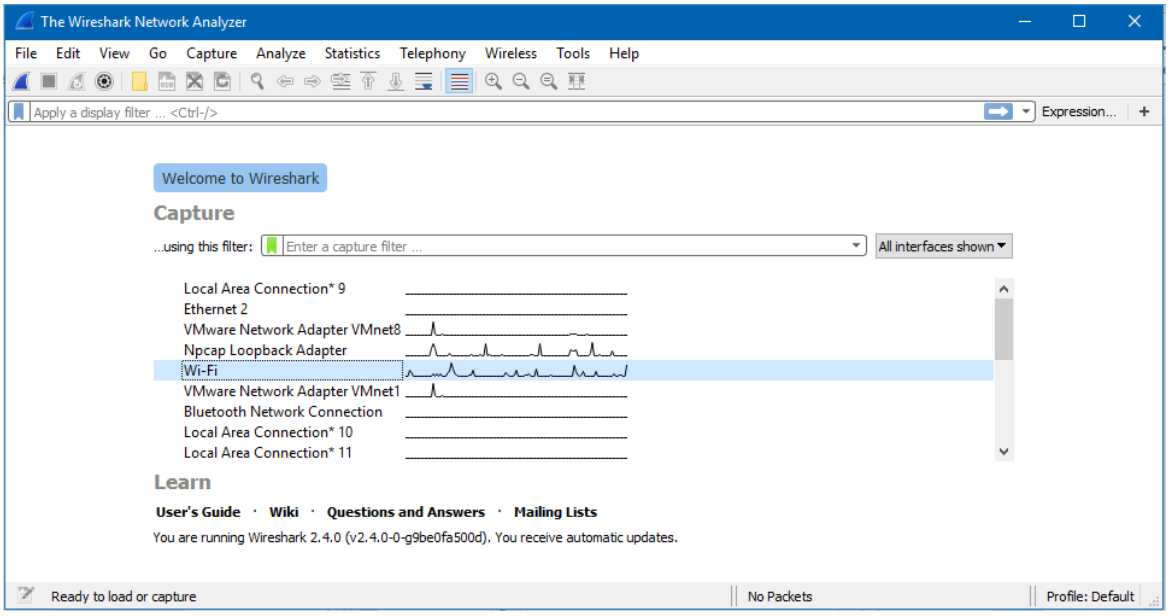
2. Kiến thức nền tảng

- Kiến thức về giao thức HTTP của chương Application

3. Môi trường & công cụ

- Một máy tính có kết nối Internet sử dụng hệ điều hành Windows/Linux.
- Phần mềm **Wireshark** được cài đặt tại máy tính trên.

  Sinh viên có thể tải về miễn phí phiên bản mới nhất theo hướng dẫn tại *https://www.wireshark.org/download.html*



Hình 1. Giao diện chính của Wireshark 2.4

B. THỰC HÀNH

1. Tạo 1 website đơn giản trên localhost

a) Tạo 1 website sử dụng HTML đơn giản

- **Bước 1:** Mở chương trình soạn thảo như: Notepad, Notepad++,…
- **Bước 2:** Tạo các dòng HTML như sau:

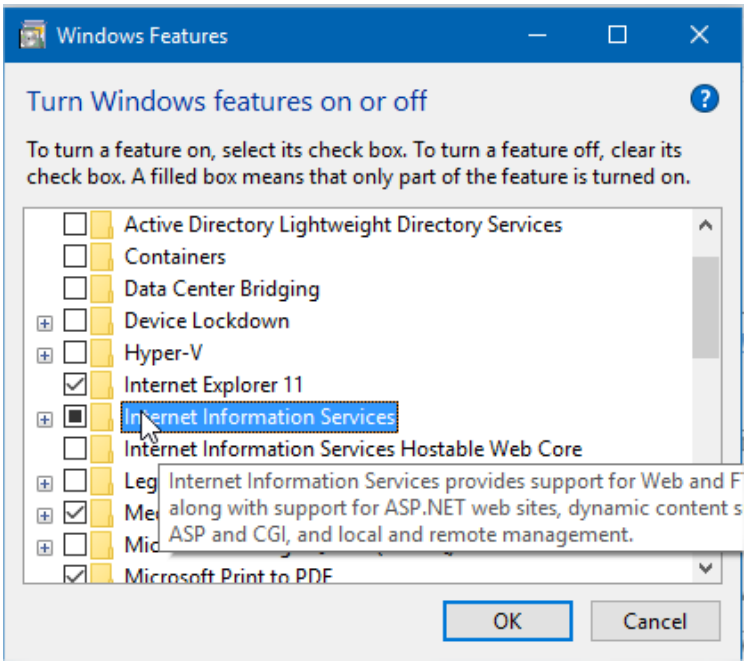
```
<!DOCTYPE html>
<html>
<head>
<title>Thực hành nhập môn mạng máy tính - 2</title>
</head>
<body>
<center><img
src="http://portal.uit.edu.vn/Styles/profi/images/logo186x150.png"/
></center>
<center><h1>MSSV: 123456</h1></center>
<center><h2> Họ và tên: Nguyễn Văn An</h2></center>
</body>
</html>
```

Sinh viên có thể tìm hiểu thêm về HTML tại: http://www.w3schools.com/ hoặc tự tạo một trang web khác tương đương bằng HTML.

- **Bước 3:** Lưu trang web với tên **MSSV.html** (MSSV là mã số sinh viên của sinh viên làm bài thực hành)

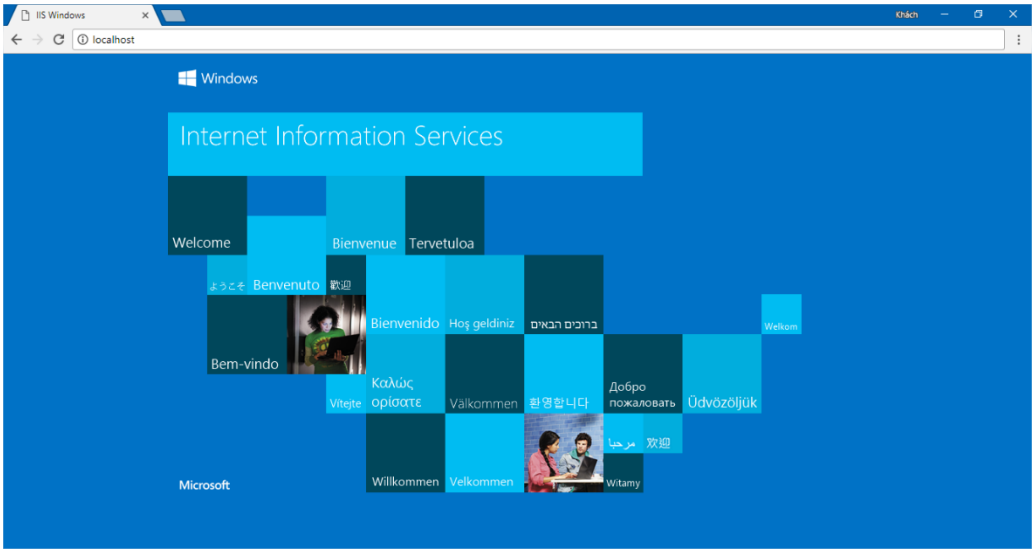
b) Cấu hình Webserver với IIS trên Windows

- **Bước 1:** Bật dịch vụ IIS (dịch vụ để tạo máy chủ Web/FTP,…) trên Windows 7 với các bước như sau: **Control Panel -> Programs and features -> Turn Windows Features on or off ->** Chọn **Internet Information Service** -> OK



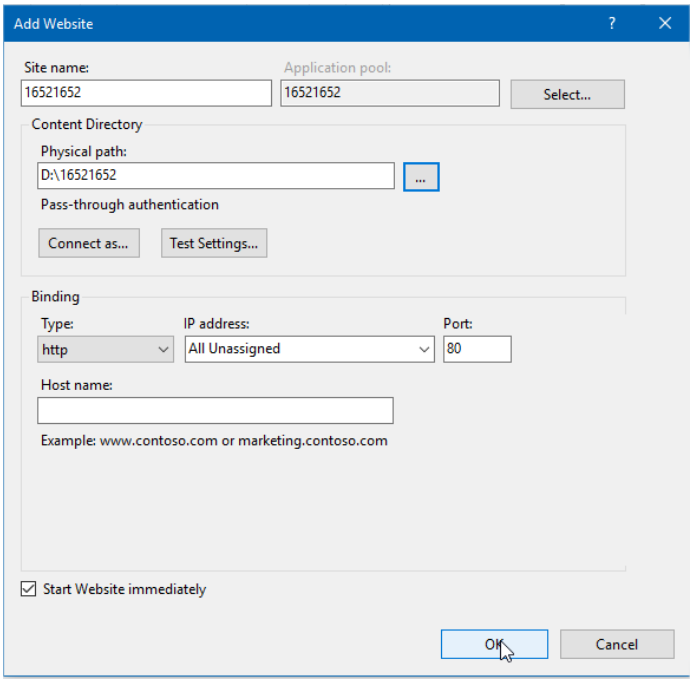
Hình 2. Chọn Internet Information Services

- **Bước 2:** Kiểm tra dịch vụ đã hoạt động hay chưa bằng cách gõ URL http://localhost vào trình duyệt bất kỳ, một trang Web mặc định của IIS hiện ra là thành công.



Hình 3. Website mặc định khi truy cập localhost của IIS

- **Bước 3:** Với các loại Windows khác nhau, vào phần search tìm phần Internet Information Services (IIS) Manager để có thể tạo trang Web trên server.
- **Bước 4:** Để tạo mới một Website, nhấp chuột phải vào **Site -> Add Website**. Đặt Sitename là MSSV (MSSV của sinh viên làm bài thực hành), Physical path là tên folder tùy ý.



Hình 4. Thêm một website mới với IIS

- **Bước 5:** Chép file MSSV.html vừa tạo ở phần 1 vào folder đã xác định trong phần Physical path.
- **Bước 6:** Truy cập vào trang web của mình với đường dẫn http://localhost/MSSV.html
- **Bước 7:** Truy cập thử trang web của sinh viên khác từ trình duyệt bằng các gõ URL như sau:

A.B.C.D/MSSV.html

với A.B.C.D là địa chỉ IP của máy tính mà bạn mình sử dụng.

MSSV.html là file html mà bạn mình tạo ra.

Gợi ý: Sinh viên có thể xem nhanh địa chỉ IP của máy mình bằng cách sau:

- Dùng tổ hợp phím Windows+R để mở cửa sổ Run
- Gõ cmd -> OK
- Gõ **ipconfig** để xem địa chỉ IP trên máy của mình
- Nếu không truy cập được, sinh viên tắt tường lửa và thử lại.

2. HTTP GET/response có điều kiện

Hầu hết các web browsers đều hỗ trợ caching và thực hiện HTTP GET có điều kiện. Trước khi thực hiện các bước sau, xóa cache của trình duyệt (đối với Firefox, chọn *Tools->Clear Recent History* và chọn *Cache box* hoặc đối với Internet Explorer thì chọn *Tools->Internet Options->Delete File*).

Thực hiện các bước sau:

- **Bước 1:** Khởi động trình duyệt và cần đảm bảo cache của trình duyệt đã được xóa.
- **Bước 2:** Khởi động Wireshark và bắt đầu bắt gói tin
- **Bước 3:** Truy cập vào trang web của sinh viên khác trong lớp.
- **Bước 4:** Nhanh chóng nhập URL đó và truy cập đến một lần nữa (hoặc chọn refresh button trên trình duyệt).

Dừng bắt gói tin và nhập "http" vào cửa sổ display-filter để hiển thị các thông điệp HTTP.

Trả lời các câu hỏi sau kèm theo hình ảnh minh chứng kết quả từ Wireshark:

1. Trình duyệt đang sử dụng phiên bản HTTP 1.0 hay 1.1? Phiên bản HTTP server đang sử dụng là bao nhiêu?
2. Địa chỉ IP của máy tính bạn là bao nhiêu? Của web server là bao nhiêu?
3. Mã trạng thái (status code) trả về từ server là gì?
4. Server đã trả về cho trình duyệt bao nhiêu bytes nội dung?
5. Xem xét nội dung của HTTP GET đầu tiên. Bạn có thấy dòng "IF-MODIFIED-SINCE" hay không?
6. Xem xét nội dung phản hồi từ server. Server có thật sự trả về nội dung của file HTML hay không? Tại sao?
7. Xem xét nội dung của HTTP GET thứ 2. Bạn có thấy dòng "IF-MODIFIED-SINCE" hay không? Nếu có, giá trị của IF-MODIFIED-SINCE là gì?
8. Mã trạng thái HTTP được trả về từ server tương ứng với HTTP GET thứ 2 là gì? Ý nghĩa nó là gì? Server có thật sự gửi về nội dung của file hay không? Giải thích.
9. Trình duyệt đã gửi bao nhiêu HTTP GET? Đến những địa chỉ IP nào?

3. Truy cập các trang HTTP dài

Trong các ví dụ của chúng ta, trang được truy cập là những files HTML ngắn và đơn giản. Chúng ta sẽ xem xét điêu gì xảy ra khi download một file HTML dài.

Thực hiện các bước sau khi có kết nối Internet:

- **Bước 1:** Khởi động web browser và đảm bảo cache được xóa.
- **Bước 2:** Khởi động Wireshark và bắt đầu bắt gói tin.
- **Bước 3:** Từ trình duyệt, truy cập đến địa chỉ sau: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html
- **Bước 4:** Dừng bắt gói tin và nhập "http" vào display-filter window để hiển thị các thông điệp HTTP.

Trong packet-listing window, bạn sẽ thấy theo sau HTTP GET là nhiều gói tin TCP phản hồi. Ở trường hợp của chúng ta, file HTML có nội dung dài, 4500 bytes là quá lớn để có thể chứa trong một gói tin TCP. Chính vì thế HTTP response được TCP tách ra thành nhiều gói nhỏ, mỗi gói chứa trong một TCP segment. Trong các phiên bản Wireshark gần đây, Wireshark xác định mỗt TCP segment là một gói tin riêng biệt và thông điệp HTTP response được phân rã ra thành nhiều gói tin TCP được xác định bởi dòng "TCP segment of reassembled PDU" trong cột Info. Các phiên bản Wireshark cũ hơn thì sử dụng "Continuation".

Trả lời các câu hỏi sau

10. Trình duyệt đã gửi bao nhiêu HTTP GET? Dòng "**THE BILL OF RIGHTS**" được chứa trong gói tin phản hồi thứ mấy?
11. Cần bao nhiêu TCP segments để chứa hết HTTP response và nội dung của The Bill of Rights?

4. Chứng thực HTTP

Truy cập vào một website được bảo vệ bởi password và quan sát chuỗi thông điệp HTTP trao đổi giữa trình duyệt và website đó. Website http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html được bảo vệ bởi password với username là "wireshark-students" (không có ngoặc kép), và password là "network" (không có ngoặc kép).

Thực hiện các bước sau khi có kết nối Internet:

- **Bước 1:** Khởi động trình duyệt và đảm bảo cache đã được xóa
- **Bước 2:** Khởi động Wireshark và bắt đầu bắt gói tin
- **Bước 3:** Từ trình duyệt, truy cập đến địa chỉ sau http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html
- **Bước 4:** Nhập username và password.

- **Bước 5:** Dừng bắt gói tin và nhập "http" vào display-filter window để hiển thị các thông điệp HTTP.
- **Bước 6:** Kiểm tra kết quả của Wireshark.

Trả lời các câu hỏi sau:

12. Mã trạng thái và ý nghĩa nó trong HTTP response tương ứng với HTTP GET đầu tiên là gì?
13. Khi trình duyệt gửi HTTP GET lần thứ 2, trường dữ liệu nào mới nào xuất hiện trong HTTP GET?

C. YÊU CẦU & ĐÁNH GIÁ

1. Yêu cầu

- Sinh viên tìm hiểu và thực hành theo hướng dẫn. Thực hiện **cá nhân.**
- Sinh viên báo cáo kết quả thực hiện và nộp bài bằng file. Trong đó:
  - Trình bày chi tiết quá trình thực hành và trả lời các câu hỏi nếu có (kèm theo các ảnh chụp màn hình tương ứng).
  - Đính kèm các file *pcapng* từ Wireshark.

Nén tất cả các file và đặt tên file theo định dạng theo mẫu: **MSSV_HoTen_BaoCaoLabX** Ví dụ: 17521007_NguyenVanA_Lab2

  - Nộp báo cáo trên theo thời gian đã thống nhất tại website môn học.

2. Đánh giá:

Sinh viên hiểu và tự thực hiện được bài thực hành, trả lời đầy đủ các yêu cầu đặt ra, khuyến khích trình bày báo cáo chi tiết, rõ ràng.

D. TÀI LIỆU THAM KHẢO

Bài thực hành được xây dựng dựa trên *Wireshark Lab: Getting Started* - Supplement to Computer Networking: A Top-Down Approach, 7th ed., J.F Kurose and K.W Ross.

HẾT

Chúc các em hoàn thành tốt