ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

NGUYỄN TRỌNG VIỆT

NGHIÊN CỨU MỘT SỐ DẠNG TẤN CÔNG WEBSITE, PHƯƠNG PHÁP VÀ CÔNG CỤ KIỂM SOÁT, PHÒNG TRÁNH TẤN CÔNG

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

Hà Nội – 2015

ĐẠI HỌC QUỐC GIA HÀ NỘI

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

NGUYỄN TRỌNG VIỆT

NGHIÊN CỨU MỘT SỐ DẠNG TẤN CÔNG WEBSITE, PHƯƠNG PHÁP VÀ CÔNG CỤ KIỂM SOÁT, PHÒNG TRÁNH TẤN CÔNG

Ngành: Công nghệ thông tin
Chuyên ngành: Hệ thống thông tin
Mã số: 60480104

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. TRẦN MINH

Hà Nội – 2015

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn này hoàn toàn do tôi thực hiện. Các trích dẫn và số liệu sử dụng trong luận văn đều được dẫn nguồn và có độ chính xác cao nhất trong phạm vi hiểu biết của tôi.

Tôi xin chịu trách nhiệm về nghiên cứu của mình.

Tác giả luận văn

Nguyễn Trọng Việt

LỜI CẢM ƠN

Sau gần 6 tháng nỗ lực thực hiện, luận văn "Nghiên cứu một số dạng Tấn công Website, phương pháp và công cụ kiểm soát, phòng tránh tấn công" đã hoàn thành. Ngoài sự cố gắng hết mình của bản thân, tôi đã nhận được sự khích lệ rất nhiều từ phía nhà trường, thầy cô, gia đình và bạn bè.

Tôi xin gửi lời cảm ơn tới các thầy cô trong khoa Công nghệ thông tin trường Đại học Công nghệ - Đại học Quốc Gia Hà Nội đã truyền đạt những kiến thức quý báu cho chúng tôi trong suốt quá trình học tập. Đặc biệt, tôi xin chân thành cảm ơn thầy Trần Minh (Viện Công nghiệp phần mềm và Nội dung số Việt Nam), người đã tận tình hướng dẫn và giúp đỡ tôi trong quá trình làm luận văn tốt nghiệp.

Xin cảm ơn tất cả bạn bè đã động viên, giúp đỡ tôi trong quá trình học tập và hoàn thành luận văn tốt nghiệp này.

LỜI MỞ ĐẦU

Cùng với sự phát triển của công nghệ thông tin, công nghệ mạng máy tính và sự phát triển của mạng Internet ngày càng đa dạng và phong phú. Các dịch vụ, ứng dụng Web đã thâm nhập vào hầu hết các lĩnh vực trong đời sống xã hội. Các thông tin trên Internet cũng đa dạng về nội dung và hình thức, trong đó có rất nhiều thông tin cần được bảo mật cao hơn bởi tính kinh tế, tính chính xác và tính tin cậy của nó.

Bên cạnh đó, các hình thức phá hoại ứng dụng Web cũng trở nên tinh vi và phức tạp hơn. Do đó đối với mỗi hệ thống, nhiệm vụ bảo mật được đặt ra cho người quản trị mạng là hết sức quan trọng và cần thiết. Xuất phát từ những thực tế đó, luận văn đi sâu tìm hiểu về các cách tấn công phổ biến nhất hiện nay và các cách phòng chống các loại tấn công này.

Chính vì vậy, thông qua việc nghiên cứu một số phương pháp tấn công và cách bảo mật các loại tấn công này, tác giả mong muốn góp một phần sức nhỏ vào việc nghiên cứu và tìm hiểu về các vấn đề an ninh mạng giúp cho việc học tập và nghiên cứu.

Lý do chọn đề tài

Trong những năm gần đây, Việt Nam ngày càng phát triển về mặt công nghệ thông tin, đặc biệt là về ứng dụng Web. Hầu hết mọi người ai cũng từng nghe và làm việc trên ứng dụng Web. Website trở nên phổ biến và trở thành một phần quan trọng của mọi người và nhất là các doanh nghiệp, công ty. Bên cạnh lý do an toàn, bảo mật cho ứng dụng Web luôn là vấn đề nan giải cho người quản trị Website. Chính vì vậy, luận văn đã đi sâu vào tìm hiểu ứng dụng Web, cách thức tấn công và bảo mật web.

Mục tiêu

Giúp cho người đọc có thể hiểu rõ hơn về các ứng dụng Web, các mối đe dọa về vấn đề an toàn thông tin khi làm việc trên ứng dụng Web hàng ngày, hiểu rõ hơn về các kỹ thuật tấn công và bảo mật web.

Phạm vi

Tìm hiểu các kỹ thuật tấn công phổ biến nhất hiện nay như Denial of Service, SQL Injection, Cross-Site Scripting,…cách bảo mật, phòng thủ các loại tấn công phổ biến trên một cách tổng quan nhất.

Bố cục của luận văn

Nội dung của luận văn bao gồm các phần sau (3 chương):

- **Chương 1: Tổng quan về an ninh mạng và ứng dụng Web -** Tìm hiểu chung về ứng dụng Web, khái niệm và hoạt động của ứng dụng Web trên Internet, đồng thời cũng đề cập khái quát về bảo mật ứng dụng Web.
- **Chương 2: Các kỹ thuật tấn công và bảo mật ứng dụng Web -** Trình bày về các kỹ thuật tấn công ứng dụng Web và cách phòng chống.
    - Chiếm hữu phiên làm việc
    - Từ chối dịch vụ
    - Chèn câu truy vấn (SQL Injection)
    - Chèn mã lệnh thực thi trên trình duyệt nạn nhân (Cross Site Scripting)
    - Các công cụ phát hiện lỗ hổng bảo mật Web
- **Chương 3: Tấn công thực nghiệm -** Một vài ví dụ tấn công ứng dụng Web bằng các kỹ thuật tấn công đã trình bày ở chương 2.
- **Kết luận:** Trình bày các kết quả đạt được của luận văn và hướng phát triển trong tương lai.

MỤC LỤC

LỜI CAM ĐOAN i
LỜI CẢM ƠN ii
LỜI MỞ ĐẦU iii
DANH MỤC CÁC TỪ VIẾT TẮT viii
DANH MỤC HÌNH VẼ viii
CHƯƠNG 1: TỔNG QUAN VỀ AN NINH MẠNG VÀ ỨNG DỤNG WEB 0
1.1 TỔNG QUAN VỀ AN NINH MẠNG 0
1.1.1 An ninh mạng là gì? 0
1.1.2 Kẻ tấn công là ai? 0
1.1.3 Lỗ hổng bảo mật? 1
1.2 TỔNG QUAN VỀ ỨNG DỤNG WEB 2
1.2.1 Giới thiệu về Website 2
1.2.2 Khái niệm về ứng dụng Web 3
1.2.3 Một số thuật ngữ trong ứng dụng Web 4
1.2.3.1 Session 4
1.2.3.2 Cookie 5
1.2.3.3 Proxy 7
1.2.4 Kiến trúc một ứng dụng Web 8
1.2.5 Nguyên lý hoạt động một ứng dụng Web 9
CHƯƠNG 2: CÁC KỸ THUẬT TẤN CÔNG VÀ BẢO MẬT ỨNG DỤNG WEB **Error! Bookma**
2.1 CHIẾM HỮU PHIÊN LÀM VIỆC ..... **Error! Bookmark not defined.**
2.1.1 Ấn định phiên làm việc (Session Fixation) ..... **Error! Bookmark not defined.**
2.1.1.1 Kỹ thuật tấn công ..... **Error! Bookmark not defined.**
2.1.1.2 Một số biện pháp bảo mật khắc phục ..... **Error! Bookmark not defined.**
2.1.2 Đánh cắp phiên làm việc (Session Hijacking) ... **Error! Bookmark not defined.**
2.1.2.1 Kỹ thuật tấn công ..... **Error! Bookmark not defined.**
2.1.2.2 Một số biện pháp bảo mật khắc phục ..... **Error! Bookmark not defined.**
2.2 TỪ CHỐI DỊCH VỤ (DOS) ..... **Error! Bookmark not defined.**
2.2.1 Những mục tiêu tấn công của DOS ..... **Error! Bookmark not defined.**
2.2.2 Kỹ thuật tấn công ..... **Error! Bookmark not defined.**
2.2.2.1 Tấn công thông qua kết nối ..... **Error! Bookmark not defined.**

2.2.2.2 Lợi dụng tài nguyên của nạn nhân để tấn công**Error! Bookmark not defined.**
2.2.2.3 Sử dụng băng thông.....**Error! Bookmark not defined.**
2.2.2.4 Sử dụng tài nguyên khác .....**Error! Bookmark not defined.**
2.2.3 Một số biện pháp bảo mật khắc phục.....**Error! Bookmark not defined.**
2.3 CHÈN CÂU TRUY VẤN SQL (SQL Injection) .....**Error! Bookmark not defined.**
2.3.1 Kỹ thuật tấn công .....**Error! Bookmark not defined.**
2.3.1.1 Tấn công SQL Injection vượt form đăng nhập đơn giản**Error! Bookmark not defi**
2.3.1.2 Tấn công dựa vào câu lệnh SELECT .....**Error! Bookmark not defined.**
2.3.1.3 Tấn công dựa vào câu lệnh INSERT.....**Error! Bookmark not defined.**
2.3.1.4 Tấn công sử dụng stored-procedures.....**Error! Bookmark not defined.**
2.3.2 Một số biện pháp bảo mật khắc phục.....**Error! Bookmark not defined.**
2.3.2.1 Kiểm soát chặt chẽ dữ liệu nhập vào.....**Error! Bookmark not defined.**
2.3.3.2 Thiết lập cấu hình an toàn cho hệ quản trị cơ sở dữ liệu**Error! Bookmark not defi**
2.4 CHÈN MÃ LỆNH THỰC THI TRÊN TRÌNH DUYỆT NẠN NHÂN (Cross Site Scripting).....**Error! Bookmark not defined.**
2.4.1 Kỹ thuật tấn công .....**Error! Bookmark not defined.**
2.4.1.1 Reflected XSS .....**Error! Bookmark not defined.**
2.4.1.2 Stored XSS .....**Error! Bookmark not defined.**
2.4.2 Một số biện pháp bảo mật khắc phục.....**Error! Bookmark not defined.**
2.4.2.1 Lọc dữ liệu.....**Error! Bookmark not defined.**
2.4.2.2 Input Encoding .....**Error! Bookmark not defined.**
2.4.2.3 Output Encoding.....**Error! Bookmark not defined.**
2.4.2.4 Web Brower’s Security .....**Error! Bookmark not defined.**
2.5 CÁC CÔNG CỤ PHÁT HIỆN LỖ HỔNG BẢO MẬT WEB**Error! Bookmark not defined**
2.5.1 Acunetix .....**Error! Bookmark not defined.**
2.5.2 Maxisploit Scanner.....**Error! Bookmark not defined.**
CHƯƠNG 3: TẤN CÔNG THỰC NGHIỆM.....**Error! Bookmark not defined.**
3.1 Công cụ cần thiết .....**Error! Bookmark not defined.**
3.1.1 XAMPP .....**Error! Bookmark not defined.**
3.1.2 DVWA .....**Error! Bookmark not defined.**
3.1.3 Firefox .....**Error! Bookmark not defined.**
3.1.4 Cài đặt và thiết lập.....**Error! Bookmark not defined.**

3.2 Thực hành tấn công.....**Error! Bookmark not defined.**
3.2.1 Tấn công SQL Injection.....**Error! Bookmark not defined.**
3.2.2 Tấn công XSS .....**Error! Bookmark not defined.**
3.2.2.1 Reflected XSS .....**Error! Bookmark not defined.**
3.2.2.2 Stored XSS .....**Error! Bookmark not defined.**
3.2.2.3 Khai thác lỗ hổng XSS và đánh cắp cookie **Error! Bookmark not defined.**
KẾT LUẬN .....**Error! Bookmark not defined.**
TÀI LIỆU THAM KHẢO.....11

DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Tiếng Anh	Tiếng Việt
ACL	Access Control List	Danh sách điều khiển truy cập
CGI	Common Gateway Interface	Bộ thông dịch Script
DDOS	Distributed Denial Of Services	Từ chối dịch vụ từ nhiều nguồn
DNS	Domain Name System	Hệ thống tên miền
DOS	Denial Of Services	Từ chối dịch vụ
FTP	File Tranfer Protocol	Giao thức truyền file đơn giản
IIS	Internet Information Services	Dịch vụ cung cấp thông tin Internet
SSL	Secure Socket Layer	Giao thức mã hóa SSL
XSS	Cross Site Scripting	Tấn công XSS ứng dụng Web

DANH MỤC HÌNH VẼ

Hình 1.1: Kiến trúc của một ứng dụng Web 8
Hình 1.2: Nguyên lý hoạt động của một ứng dụng Web.....9
Hình 2.1: Nguyên lý tấn công ấn định phiên làm việc**Error! Bookmark not defined.**
Hình 2.2: Cơ chế thiết lập kết nối trước khi truyền số liệu**Error! Bookmark not defined.**
Hình 2.3: Tấn công DoS truyền thống .....**Error! Bookmark not defined.**
Hình 2.4: Tấn công SYN flood .....**Error! Bookmark not defined.**
Hình 2.5: Tấn công DDOS .....**Error! Bookmark not defined.**
Hình 2.6: Tấn công Smurf Attack .....**Error! Bookmark not defined.**
Hình 2.7: Một site bị lỗi SQL Injection.....**Error! Bookmark not defined.**
Hình 2.8: Tấn công SQL Injection .....**Error! Bookmark not defined.**
Hình 2.9: Tấn công dạng Reflected.....**Error! Bookmark not defined.**
Hình 2.10: Tấn công XSS thông qua email.....**Error! Bookmark not defined.**
Hình 2.11: Các bước thực hiện XSS đánh cắp Cookie người dùng**Error! Bookmark not de**

Hình 2.12: Chèn câu lệnh Javascript để khai thác lỗ hổng XSS**Error! Bookmark not defin**
Hình 2.13: Popup hiển thị chứng tỏ web bị lỗi XSS**Error! Bookmark not defined.**
Hình 2.14: Kiểu tấn công Stored XSS.....**Error! Bookmark not defined.**
Hình 2.15: Giao diện của Acunetix Web Vulnerability Scanner**Error! Bookmark not defi**
Hình 2.16: Kết quả Scan của Acunetix Web Vulnerability Scanner**Error! Bookmark not d**
Hình 2.17: Giao diện của Maxisploit Scanner .....**Error! Bookmark not defined.**
Hình 2.18: Kết quả Scan của Maxisploit Scanner...**Error! Bookmark not defined.**
Hình 3.1: Hiển thị từng bản ghi trong CSDL .....**Error! Bookmark not defined.**
Hình 3.2: Hiển thị tất cả các bản ghi trong CSDL ..**Error! Bookmark not defined.**
Hình 3.4: Hiển thị phiên bản CSDL .....**Error! Bookmark not defined.**
Hình 3.5: Hiển thị tên CSDL.....**Error! Bookmark not defined.**
Hình 3.6: Hiển thị nội dung các cột của bảng User trong CSDL**Error! Bookmark not defi**
Hình 3.7: Giải mã MD5 để lấy password.....**Error! Bookmark not defined.**
Hình 3.8: Lấy cookie của trang web.....**Error! Bookmark not defined.**
Hình 3.9: Website bị lỗi Stored XSS.....**Error! Bookmark not defined.**
Hình 3.10: Cookie được gửi về Hacker.....**Error! Bookmark not defined.**
Hình 3.11: Đăng nhập bằng cookie đã ăn cắp được**Error! Bookmark not defined.**

CHƯƠNG 1: TỔNG QUAN VỀ AN NINH MẠNG VÀ ỨNG DỤNG WEB

1.1 TỔNG QUAN VỀ AN NINH MẠNG

1.1.1 An ninh mạng là gì?

An ninh mạng là một trong những lĩnh vực mà hiện nay giới công nghệ thông tin khá quan tâm. Một khi Internet ra đời và phát triển, nhu cầu trao đổi thông tin trở nên cần thiết. Mục đích của việc kết nối mạng là làm cho mọi người có thể sử dụng chung tài nguyên mạng từ những vị trí địa lý khác nhau. Chính vì vậy mà các tài nguyên dễ dàng bị phân tán, hiển nhiên một điều là chúng ta dễ bị xâm phạm, gây mất mát dữ liệu cũng như các thông tin có giá trị. Kết nối càng rộng thì càng dễ bị tấn công, đó là một quy luật tất yếu. Từ đó, vấn đề bảo vệ thông tin cũng đồng thời xuất hiện và như thế an ninh mạng ra đời.

Ví dụ: User A gửi một tập tin cho User B trong phạm vi là nước Việt Nam thì nó khác xa so với việc User A gửi tập tin cho User C ở Mỹ. Ở trường hợp đầu thì dữ liệu có thể mất mát với phạm vi nhỏ là trong nước nhưng trường hợp sau thì việc mất mát dữ liệu với phạm vi rất rộng là cả thế giới.

Mỗi một lỗ hổng trên mạng đều là mối nguy hiểm tiềm tàng. Từ một lổ hổng bảo mật nhỏ của hệ thống, nhưng nếu biết khai thác và lợi dụng kỹ thuật hack điêu luyện thì cũng có thể trở thành mối tai họa.

1.1.2 Kẻ tấn công là ai?

Kẻ tấn công người ta thường gọi là Hacker, là những kẻ tấn công vào hệ thống mạng với nhiều mục đích khác nhau. Trước đây Hacker được chia làm 2 loại nhưng hiện nay thì được chia thành 3 loại:

Hacker mũ đen

Đây là tên trộm chính hiệu, với những Hacker có kinh nghiệm thì đặc biệt nguy hiểm đối với hệ thống mạng. Mục tiêu của chúng là đột nhập vào hệ thống mạng của đối tượng để lấy cấp thông tin, nhằm mục đích bất chính. Hacker mũ đen là những tội phạm thật sự cần sự trừng trị của pháp luật.

Hacker mũ trắng

Họ là những nhà bảo mật và bảo vệ hệ thống. Họ cũng xâm nhập vào hệ thống, mục đích là tìm ra những kẽ hở, những lổ hổng chết người và sau đó tìm cách vá lại

chúng. Tất nhiên, Hacker mũ trắng cũng có khả năng xâm nhập và cũng có thể trở thành Hacker mũ đen.

Hacker mũ xám

Loại này được sự kết hợp giữa hai loại trên. Thông thường họ là những người còn trẻ, muốn thể hiện mình. Trong một thời điểm, họ đột nhập vào hệ thống để phá phách. Nhưng trong thời điểm khác họ có thể gửi đến nhà quản trị những thông tin về lổ hổng bảo mật và đề xuất cách vá lỗi.

Ranh giới phân biệt các Hacker rất mong manh. Một kẻ tấn công là Hacker mũ trắng trong thời điểm này nhưng ở thời điểm khác họ lại là một tên trộm chuyên nghiệp.

1.1.3 Lỗ hổng bảo mật?

Các lỗ hổng bảo mật trên một hệ thống là các điểm yếu có thể tạo ra sự ngưng trệ của dịch vụ, thêm quyền đối với người sử dụng hoặc cho phép các truy nhập không hợp pháp vào hệ thống. Các lỗ hổng cũng có thể xuất hiện ngay trong hạ tầng mạng hoặc nằm ngay trên các dịch vụ cung cập như Sendmail, Web, Ftp,... Ngoài ra các lỗ hổng còn tồn tại ngay chính các hệ điều hành như: Windows XP, 7, Linux,... hoặc trong các ứng dụng mà người sử dụng thường xuyên sử dụng như: Office, trình duyệt,...

Theo bộ quốc phòng Mỹ, các lỗ hổng bảo mật một hệ thống được chia như sau:

Lỗ hổng loại A

Các lỗ hổng này cho phép người sử dụng ở ngoài có thể truy nhập vào hệ thống bất hợp pháp. Lỗ hổng này rất nguy hiểm, có thể phá hủy toàn bộ hệ thống.

Lỗ hổng loại B

Các lỗ hổng này cho phép người sử dụng thêm các quyền trên hệ thống mà không cần thực hiện kiểm tra tính hợp lệ. Mức độ nguy hiểm trung bình. Những lỗ hổng này thường có trong các ứng dụng trên hệ thống, có thể dẫn đến mất hoặc lộ thông tin dữ liệu.

Lỗ hổng loại C

Các lỗ hổng loại này cho phép thực hiện các phương thức tấn công theo DoS. Mức độ nguy hiểm thấp, chỉ ảnh hưởng tới chất lượng dịch vụ, có thể làm ngưng trệ, gián đoạn hệ thống, không làm phá hỏng dữ liệu hoặc được quyền truy nhập bất hợp pháp.

1.2 TỔNG QUAN VỀ ỨNG DỤNG WEB

1.2.1 Giới thiệu về Website

Website là một "trang web" được lưu trữ tại các máy chủ hay các hosting hoạt động trên Internet. Đây là nơi giới thiệu những thông tin, hình ảnh về doanh nghiệp, sản phẩn và dịch vụ của doanh nghiệp hay giới thiệu bất kỳ thông tin gì để khách hàng có thể truy cập bất kỳ ở đâu, bất cứ lúc nào.

Website là tập hợp của nhiều web page. Khi doanh nghiệp, công ty xây dựng website nghĩa là đang xây dựng nhiều trang thông tin về sản phẩm, dịch vụ hay giới thiệu,... Để tạo nên một website cần có 3 yếu tố sau:

Tên miền (domain)

Thực chất một website không cần đến tên miền nó vẫn có thể hoạt động bình thường vì nó còn có địa chỉ IP của trang web đấy, chúng ta chỉ cần gõ vào trình duyệt IP của trang web thì ngay lập tức trình duyệt sẽ load trang web đấy về trình duyệt của bạn. Sỡ dĩ chúng ta cần phải có tên miền thay cho IP là vì IP là mỗi chuỗi số thập phân, có những địa chỉ IP thì rất là dễ nhớ nhưng đa số địa chỉ IP thì rất là khó nhớ. Với cái tên nó rất gần gũi với ngôn ngữ tự nhiên của con người nên rất là dễ nhớ cũng chính vì vậy mà người ta đã thay tên miền cho IP và từ đó công nghệ DNS ra đời.

Ví dụ đơn giản để hiểu thêm tính năng của tên miền: Trong danh bạ điện thoại của chúng ta nếu chúng ta lưu số điện thoại mà không gán với một tên thì chắc chắn một điều là chúng ta không thể nhớ hết được số điện thoại của từng người và cũng không thể nào biết được số điện thoại này là của ai nhưng nếu chúng ta lưu số một ai đó với một cái tên thì sau này khi cần gọi cho người đó sẽ tìm trong danh bạ dễ dàng hơn.

Nơi lưu trữ website (hosting)

Nơi lưu trữ website thì bắt buộc chúng ta phải có, nó có thể là một máy chủ để lưu trữ hay một hosting chúng ta thuê từ nhà cung cấp dịch vụ.

Nội dung các trang thông tin (web page)

Nội dung trang thông tin này thì phải có rồi vì mục đích của chúng ta lập nên website nhằm đăng thông tin của chúng ta lên website hay giới thiệu các thông tin của công ty.

Nói đến một website người ta thường nói website đấy là web động hay tĩnh, đa số các website bây giờ đến là website động.

Website tĩnh có thể hiểu như thế này: người dùng gửi yêu cầu một tài nguyên nào đó và máy chủ sẽ trả về tài nguyên đó. Các trang Web không khác gì là một văn bản được định dạng và phân tán. Lúc mới đầu phát triển website thì web tĩnh được sử dụng rất nhiều vì lúc đấy nhu cầu của việc đăng tải trên website là chưa cao như đăng thông tin về các sự kiện, địa chỉ hay lịch làm việc qua Internet mà thôi, chưa có sự tương tác qua các trang Web.

Website động là thuật ngữ được dùng để chỉ những website được hỗ trợ bởi một phần mềm cơ sở web, nói cho dễ hiểu thì web động là web có cơ sở dữ liệu. Ngày nay, đa số các trang web đều có cơ sở dữ liệu vì mục đích, nhu cầu của con người càng ngày gia tăng. Thực chất, website động có nghĩa là một website tĩnh được "ghép" với một phần mềm web (các modules ứng dụng cho Web). Với chương trình phần mềm này, người chủ website thực sự có quyền điều hành nó, chỉnh sửa và cập nhật thông tin trên website của mình mà không cần phải nhờ đến những người chuyên nghiệp.

Trước đây, năm 1995 đến 2004 thì sử dụng công nghệ web 1.0 với công nghệ này thì chỉ được đọc nội dung trang web mà người dùng không thể chỉnh sửa, bình luận hay nói cách khác website lúc bất giờ chỉ hoạt động một chiều mà thôi.

Hiện nay, đã phát triển công nghệ web 2.0 hoạt động hai chiều có nghĩa là người dùng cũng có thể chỉnh sửa, bình luận hay xóa nội dung trang web. Trên đà phát triển đó người ta tiếp tục nghiên cứu và phát triển web 3.0 hướng hẹn rất nhiều điều thú vị còn ở phía trước.

1.2.2 Khái niệm về ứng dụng Web

Ứng dụng Web là một ứng dụng máy chủ/máy khách sử dụng giao thức HTTP để tương tác với người dùng hay hệ thống khác. Trình duyệt Web dành cho người dùng như Internet Explore hoặc Firefox hay Chrome,... người dùng gửi và nhận các thông tin từ máy chủ Web thông qua việc tác động vào các trang Web. Các ứng dụng Web có thể là trang trao đổi mua bán, các diễn đàn, gửi và nhận email, games online,...

Với công nghệ hiện nay, website không chỉ đơn giản là một trang tin cung cấp các bài tin đơn giản. Những ứng dụng web viết trên nền web không chỉ được gọi là một phần

của website nữa, giờ đây chúng được gọi là phần mềm viết trên nên web. Có rất nhiều phần mềm chạy trên nền web như Google Word (xử lý các file văn bản), Google spreadsheets (xử lý tính bảng tính), Google Translate (từ điển, dịch văn bản),...

Ngày nay, ứng dụng web phát triển rất cao, gần như bây giờ người ta đều sử dụng ứng dụng web như xem phim online, nghe nhạc online, chia sẻ mạng xã hội (facebook, zing), chơi games online, ngân hàng trực tuyến,... và bắt đầu xuất hiện những Hacker muốn thu lợi ích về phần mình từ các ứng dụng web.

1.2.3 Một số thuật ngữ trong ứng dụng Web

1.2.3.1 Session

Session là khoảng thời gian người sử dụng giao tiếp với một ứng dụng. Session bắt đầu khi người sử dụng truy cập vào ứng dụng lần đầu tiên, và kết thúc khi người sử dụng thoát khỏi ứng dụng. Mỗi session sẽ có một định danh (ID), mỗi session khác nhau sẽ có ID khác nhau. Trong ngữ cảnh ứng dụng web, website sẽ quyết định khi nào session bắt đầu và kết thúc. Trong một session, website có thể lưu trữ một số thông tin như đánh dấu bạn đã login hay chưa, những bài viết nào bạn đã đọc qua…

HTTP là giao thức hướng đối tượng tổng quát, phi trạng thái, nghĩa là HTTP không lưu trữ trạng thái làm việc giữa trình duyệt với trình chủ. Sự thiếu sót này gây khó khăn cho một số ứng dụng web, bởi vì trình chủ không biết trước đó trình duyệt đã có những trạng thái nào. Vì thế, để giải quyết vấn đề này, ứng dụng web đưa ra một khái niệm phiên làm việc (Session). Còn Session ID là một chuỗi để chứng thực phiên làm việc. Một số trình chủ sẽ cung cấp một Session ID cho người dùng khi họ xem trang web trên trình chủ.

Để duy trì phiên làm việc thì Session ID thường được lưu vào:

- Biến trên URL
- Biến ẩn form
- Cookie

Phiên làm việc chỉ tồn tại trong một thời gian cho phép, thời gian này được cấu hình quy định tại trình chủ hoặc bởi ứng dụng thực thi. Trình chủ sẽ tự động giải phóng phiên làm việc để khôi phục lại tài nguyên của hệ thống.

1.2.3.2 Cookie

Cookie là những phần dữ liệu nhỏ có cấu trúc được chia sẻ giữa trình chủ và trình duyệt của người dùng.

Các cookie được lưu trữ dưới những file dữ liệu nhỏ dạng text, được ứng dụng tạo ra để lưu trữ/truy tìm/nhận biết các thông tin về người dùng ghé thăm trang web và những vùng mà họ đi qua trong trang. Những thông tin này có thể được bao gồm tên/định danh người dùng, mật khẩu, sở thích, thói quen… cookie được trình duyệt của người dùng chấp nhận lưu trên đĩa cứng của máy tính, tuy nhiên không phải lúc nào trình duyệt cũng hỗ trợ cookie, mà còn tùy thuộc vào người dùng có chấp nhận chuyện lưu trữ đó hay không.

Ở những lần truy cập sau đến trang web đó, ứng dụng có thể dùng lại những thông tin trong cookie (như thông tin liên quan đến việc đăng nhập vào Facebook…) mà người dùng không phải làm lại thao tác đăng nhập hay cung cấp các thông tin khác.

Cookie được phân làm 2 loại secure/non-secure và persistent/non-persistent do đó ta sẽ có 4 kiểu cookie là:

- Persistent và Secure
- Persistent và Non-Secure
- Non-Persistent và Secure
- Non-Persistent và Non-Secure

Persistent cookie được lưu trữ dưới dạng tập tin .txt trên máy khách trong một khoảng thời gian xác định.

Non-Persistent cookie thì được lưu trữ trên bộ nhớ RAM của máy khách và sẽ bị hủy khi đóng trang web hay nhận được lệnh hủy từ trang web.

Secure cookie chỉ có thể được gửi thông qua HTTPS (SSL).

Non-Secure cookie có thể được gửi bằng cả hai giao thức HTTPS hay HTTP. Thực chất là đối với secure cookie thì trình chủ sẽ cung cấp chế độ truyền bảo mật.

Các thành phần của một cookie bao gồm:

Domain	Flag	Path	Secure	Expiration	Name	Value
www.abc.com	FALSE	/	FALSE	1154029490	Apache	64.3.40.151.1601 8996349247480

Domain: tên miền của trang web đã tạo cookie (trong ví dụ trên là www.abc.com)

Flag: mang giá trị TRUE/FALSE – Xác định các máy khác với cùng tên miền có được truy xuất đến cookie hay không.

Path: phạm vi các địa chỉ có thể truy xuất cookie. Ví dụ: Nếu path là "/tracuu" thì các địa chỉ trong thư mục /tracuu cũng như tất cả các thư mục con của nó như /tracuu/baomat có thể truy xuất đến cookie này. Còn nếu giá trị là "/" thì cookie sẽ được truy xuất bởi tất cả địa chỉ thuộc miền trang web tạo cookie.

Secure: mang giá trị TRUE/FALSE – Xác định đây là một secure cookie hay không, nghĩa là kết nối có sử dụng SSL hay không.

Expiration: thời gian hết hạn của cookie,, được tính bằng giây kể từ 00:00:00 giờ GMT ngày 01/01/1970. Nếu giá trị này không được thiết lập thì trình duyệt sẽ hiểu đây là non-persistent cookie và chỉ lưu trong bộ nhớ RAM và sẽ xóa nó khi trình duyệt bị đóng.

Name: tên biến (trường hợp này là Apache)

Value: với cookie được tạo ở trên thì giá trị của Apache là 64.3.40.151.16018996349247480, của tên miền http://www.abc.com

Kích thước tối đa của cookie là 4kb. Số cookie tối đa cho một tên miền là 20 cookie. Cookie bị hủy ngay khi đóng trình duyệt gọi là "session cookie".

Một ví dụ về cookie: Giả sử lần đầu tiên bạn vào trang facebook.com thì máy tính của bạn sẽ tải trang này rất lâu vì nó phải tải nội dung trang web về máy của bạn. Sau khi đăng nhập vào hệ thống và sử dụng như bình thường. Sang ngày hôm sau, vào lại trang facebook.com thì vào rất nhanh và nhiều khi cũng không cần phải đăng nhập tài khoản nữa nguyên nhân chính là do trình duyệt đã lưu cookie các thông tin hôm qua bạn đã vào. Cookie là một cao dao hai lưỡi, lợi ích của nó thì bạn có thể thấy được sự tiện lợi là đỡ tốn thời gian tải lại trang web nhưng ngược lại nhược điểm của nó là các Hacker có thể dựa vào các file cookie để lấy các thông tin tài khoản. Rất là nguy hiểm nên tốt nhất

không để trình duyệt lưu cookie nhưng đa số người dùng hiện nay đều để chế độ lưu cookie vì người dùng không biết đến sự nguy hiểm của nó hoặc là thấy nó tiện cho công việc của mình.

1.2.3.3 Proxy

Hiện nay, người dùng sử dụng Internet đa số là đi Internet trực tiếp nghĩa là người dùng tự mình đi đến máy chủ hỏi xin các yêu cầu. Đi trực tiếp như thế này thì có cái khuyết điểm là băng thông sẽ tốt rất nhiều cũng chính vấn đề về băng thông nên mới ra đời khái niệm "proxy".

Proxy là một Internet server làm nhiệm vụ chuyển tiếp thông tin và kiểm soát tạo sự an toàn cho việc truy cập Internet của các máy khách, còn gọi là khách hàng sử dụng dịch vụ Internet. Trạm cài đặt proxy gọi là proxy server. Proxy hay trạm cài đặt proxy có địa chỉ IP và một cổng truy cập cố định.

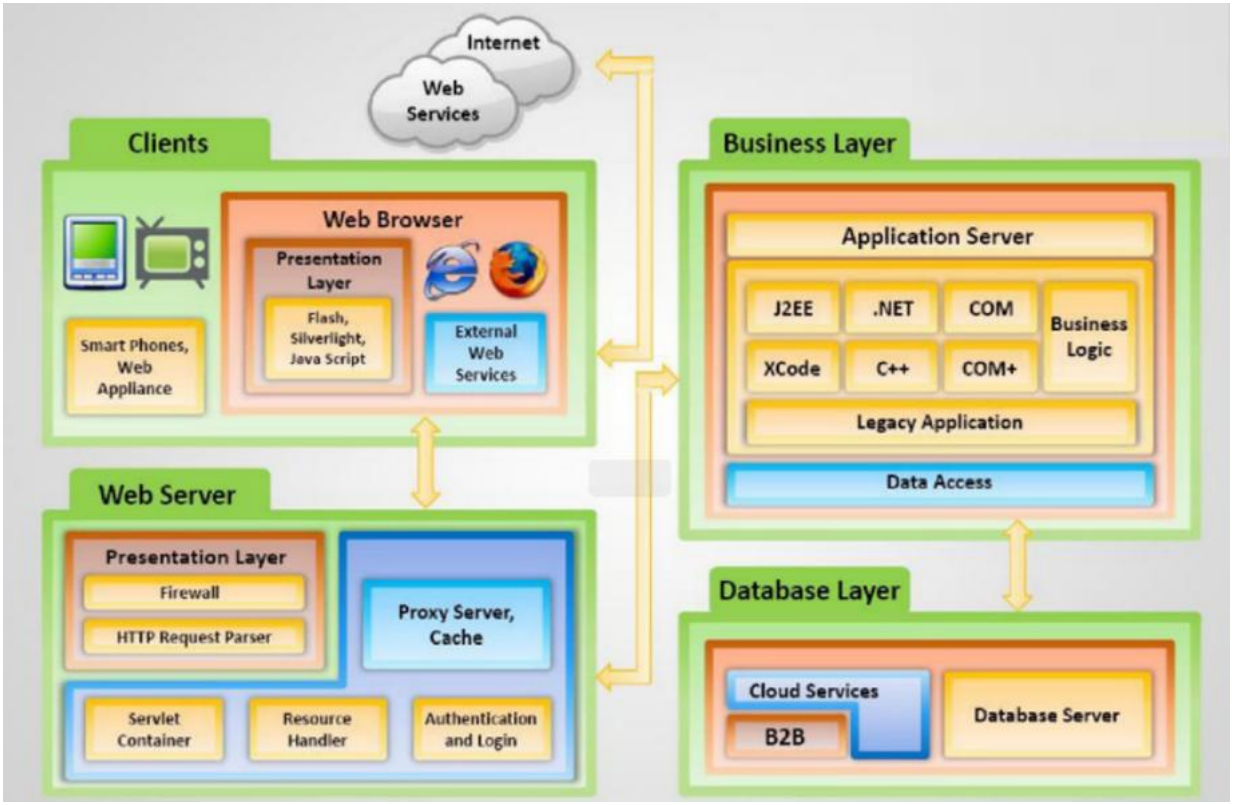
Proxy cung cấp cho người sử dụng truy xuất Internet những nghi thức đặt biệt. Những chương trình máy khách của người sử dụng sẽ qua trung gian máy chủ proxy thay thế cho máy chủ thật sự mà người sử dụng cần giao tiếp.

Máy chủ proxy xác định những yêu cầu từ client và quyết định đáp ứng hay không đáp ứng, nếu yêu cầu được đáp ứng máy chủ proxy sẽ kết nối với máy chủ thật thay cho máy khách và tiếp tục chuyển tiếp những yêu cầu từ máy khách đến máy chủ, cũng như trả lời của máy chủ đến máy khách. Vì vậy máy chủ proxy giống cầu nối trung gian giữa máy chủ và máy khách.

Thường thì máy chủ proxy được xây dựng chủ yếu là trong công ty hay các nhà cung cấp dịch vụ để phục vụ cho nhân viên hay là khách hàng của nhà cung cấp. Ví dụ: người dùng ở Việt Nam thích vào trang facebook.com nhưng hiện nay thì các nhà mạng lại chặn trang facebook. Sở dĩ nhà mạng có thể chặn được người dùng là vì nó dựa trên gói tin chạy qua Router với địa chỉ đích của facebook là bị chặn. Vậy thì người dùng không thể đi đến trang facebook đó theo phương thức truyền thống là trực tiếp nữa rồi nên người dùng mới đi theo gián tiếp là trỏ trang facebook đến một máy chủ proxy để nhờ máy chủ proxy đẩy đi đến trang facebook giúp. Như vậy thì người dùng có thể truy cập facebook mặc dù bị các nhà mạng chặn. Nói như vậy không có nghĩa là đi theo kiểu proxy

là lợi hoàn toàn, khuyết điểm lớn nhất mà proxy mắc phải là bảo mật vì nó là thằng trung gian nên nó có thể biết hết mọi thứ mà người dùng khai báo với máy chủ facebook.

1.2.4 Kiến trúc một ứng dụng Web



Hình 1.1: Kiến trúc của một ứng dụng Web

Một ứng dụng Web có đầy đủ các thành phần như sau:

Máy khách

Tại máy khách muốn truy cập vào được các ứng dụng web thì phải có trình duyệt web. Có thể dùng trình duyệt Web mặc định của các hệ điều hành như windows là Internet Explorer, Linux thường là Firefox,... còn không thì có thể cài thêm các chương trình duyệt web như Google Chrome, Opera,...

Máy chủ web

Là nơi lưu trữ nội dung trang web, tiếp nhận các yêu cầu kết nối từ máy khách, máy chủ web sử dụng phần mềm để chạy dịch vụ web phục vụ cho các máy khách như trên Windows có IIS, Linux thì có Apache, Tom cat,...

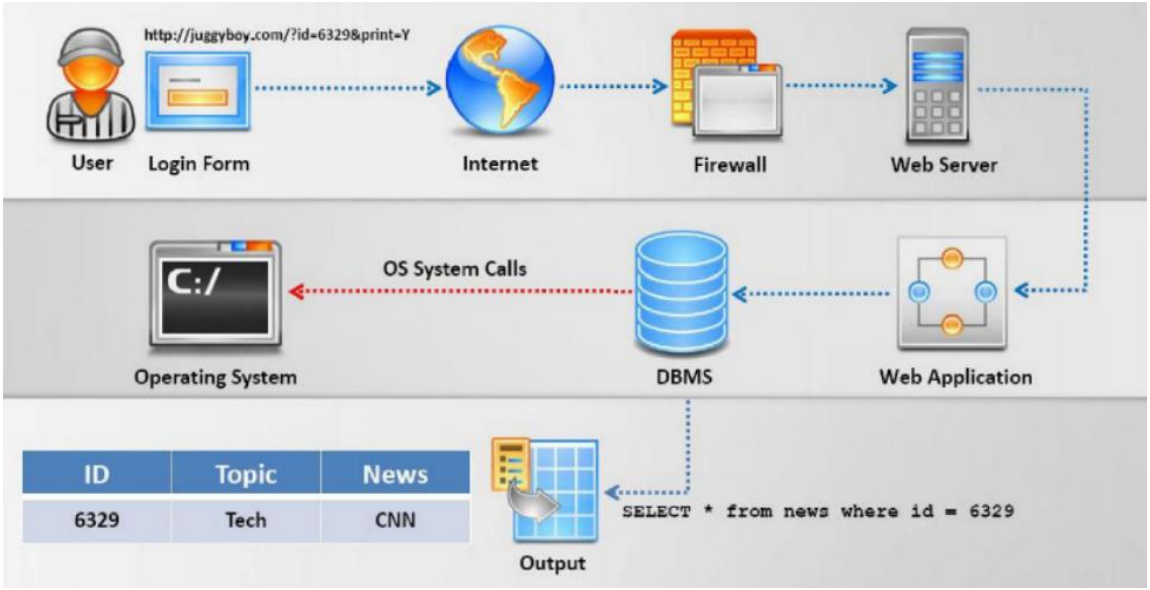
Ứng dụng web

Ứng dụng web được viết bằng các ngôn ngữ khác nhau như java, php,.. hay có thể là một đoạn flash đơn giản để nhúng các ứng dụng vào trang web. Ví dụ như games online trên facebook hay zing.

Cơ sở dữ liệu

Là một máy chủ đảm nhiệm việc lưu trữ thông tin của các ứng dụng web có thể là lưu trữ ngay trên máy chủ web hoặc là một máy chủ khác nhưng thường để bảo mật thì người ta lưu trên một máy chủ khác và sử dụng hệ quản trị cơ sở dữ liệu như SQL Server hay Oracle,... Ví dụ: như chơi games online trên web của facebook hay zing thì người chơi games xong thường lưu các giá trị của người chơi vào một cơ sở dữ liệu nào đấy và khi nào người chơi muốn tiếp tục chơi thì truy vấn lấy cơ sở dữ liệu đấy ra.

1.2.5 Nguyên lý hoạt động một ứng dụng Web



Hình 1.2: Nguyên lý hoạt động của một ứng dụng Web

Trình khách (hay còn gọi là trình duyệt): Internet Explorer, Firefox, Chrome...

Trình chủ: Apache, IIS,...

Hệ quản trị cơ sở dữ liệu: SQL Server, MySQL, DB2, Acess...

Hoạt động của một ứng dụng Web:

Trình duyệt:

- Gửi một yêu cầu (request) đến trình chủ thông qua các lệnh cơ bản GET, POST của giao thức HTTP.

Trình chủ:

- Thực thi một chương trình được xây dựng từ nhiều ngôn ngữ như Perl, C/C++...
- Yêu cầu bộ diễn dịch thực thi các trang ASP, JSP, PHP...
- Trình chủ trả về cho trình khách một luồng dữ liệu có định dạng theo giao thức HTTP gồm 2 phần: Header – mô tả các thông tin về gói dữ liệu và các thuộc tính, trạng thái trao đổi giữa trình duyệt và WebServer. Body – phần nội dung dữ liệu mà Server gửi về Client, nó có thể là một file HTML, một hình ảnh, một đoạn phim hay một văn bản bất kỳ.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Lê Đình Duy (2004), *Tấn công kiểu SQL Injection – Tác hại và phòng tránh*, Kỷ yếu hội thảo Công nghệ thông tin 2004, Trường Đại học Khoa học Tự nhiên Thành phố Hồ Chí Minh.
2. Bùi Duy Hùng (2009), *Phát hiện lỗ hổng an ninh trên các ứng dụng Web*, Đồ án tốt nghiệp, Trường Đại học Bách Khoa Hà Nội.
3. Đặng Thị Thu Hương (2009), *Một số vấn đề bảo mật Web*, Luận văn tốt nghiệp, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.
4. Võ Đỗ Thắng (2014), *Web Application & Defense*, Slide bài giảng, Trung tâm Đào tạo Quản trị mạng và an ninh mạng Athena.
5. Võ Đỗ Thắng (2014), *Các công cụ xác định lỗ hổng Website*, Slide bài giảng, Trung tâm Đào tạo Quản trị mạng và an ninh mạng Athena.
6. Nguyễn Duy Thắng, Nguyễn Thu Minh, *Nghiên cứu một số vấn đề về bảo mật ứng dụng Web trên Internet*, Luận văn tốt nghiệp, Trường Đại học Khoa học Tự nhiên Thành phố Hồ Chí Minh.

Tiếng Anh

7. Stuart Mcclure (2012), *Hacking Exposed 7: Network Security Secrets & Solutions*, VP of Operations & Strategy for the Risk & Compliance Business Unit at McAfee.
8. Kevin Spett (2002), *SQL Injection - Are your web Application vulnerable?*, SPI Dynamics, Inc.