HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

PHẠM THỊ TRANG

NGHIÊN CỨU ĐẢM BẢO AN TOÀN THÔNG TIN TRONG MÔI TRƯỜNG WEB SỬ DỤNG KỸ THUẬT MẬT MÃ

Chuyên ngành: Khoa học máy tính

Mã số: 60.48.01

Người hướng dẫn khoa học: GS. TS NGUYỄN BÌNH

TÓM TẮT LUẬN VĂN THẠC SĨ

HÀ NỘI – 2012

MỞ ĐẦU

Trong thời đại công nghệ thông tin như hiện nay, khi mà internet trở nên thân quen và dần trở thành một công cụ không thể thiếu trong cuộc sống thì lợi ích của website đối với các cơ quan nhà nước nói chung, người dân và các doanh nghiệp là vô cùng lớn. Tuỳ từng lĩnh vực, loại hình, đối tượng mà lợi ích của website mang lại khác nhau.

Các phương thức tấn công qua mạng ngày càng tinh vi, phức tạp có thể dẫn đến mất mát thông tin, thậm chí có thể làm sụp đổ hoàn toàn hệ thống thông tin của đơn vị. Vì vậy nhu cầu an toàn và bảo mật web được xem là một trọng tâm trong lĩnh vực an toàn thông tin của nhiều nước và nhiều hãng an toàn nổi tiếng trên thế giới.

Tại Việt Nam, việc nghiên cứu và ứng dụng các chuẩn an toàn web, các sản phẩm an toàn và bảo mật web còn hạn chế vì một số lý do như các tổ chức, cá nhân chưa đánh giá hết mức độ quan trọng của dữ liệu được trao đổi thông tin qua web, ý thức bảo mật thông tin của các tổ chức, cá nhân còn thiếu, đầu tư cho an toàn và bảo mật chưa đồng bộ, giá thành của các sản phẩm an toàn và bảo mật web còn khá cao. Vì vậy việc nghiên cứu về vấn đề đảm bảo an toàn thông tin trong môi trường web có ý nghĩa hết sức quan trọng.

Xuất phát từ những thực tế trên tôi đã chọn đề tài "***Nghiên cứu đảm bảo thông tin trong môi trường web sử dụng kỹ thuật mật mã***" nhằm nghiên cứu một số kỹ thuật mật mã ứng dụng vào trong quá trình đảm bảo an toàn thông tin trước sự tấn công khai thác thông tin trái phép trong môi trường web.

CHƯƠNG 1 - TỔNG QUAN VỀ AN TOÀN VÀ BẢO MẬT WEB

1.1. Quá trình phát triển của web

Ý tưởng về web do Tim Berners Lee, ông đã phát minh ra một giao thức cơ bản cho việc chuyển tải các tài liệu lên mạng là HTTP, ngôn ngữ đánh dấu siêu văn bản HTML để mô tả cấu trúc của một tài liệu. Kể từ khi ra đời web đã phát triển không ngừng và đã trải qua các thế hệ web 1.0, web 2.0, web 3.0.

1.2. Mô hình hoạt động của web

Từ một trình duyệt, người dùng gõ vào địa chỉ của một trang web, trình duyệt sẽ thực hiện một kết nối tới máy chủ tên miền để chuyển đổi tên miền ra địa chỉ IP tương ứng. Sau đó, trình duyệt sẽ gửi tiếp một kết nối tới máy chủ của website có địa chỉ IP này qua cổng 80. Dựa trên giao thức HTTP, trình duyệt sẽ gửi yêu cầu GET đến máy chủ. Khi đó máy chủ sẽ xử lý các yêu cầu của người dùng, rồi gửi trả kết quả về cho phía client.

1.3. Các hiểm hoạ đối với an toàn web

1.3.1. Tấn công vào vùng ẩn

Dựa vào chức năng "View Source" của trình duyệt mà có thể biết được thông tin về các phiên làm việc của client do đó có thể tìm ra sơ hở của trang web mà ta muốn tấn công và từ đó có thể truy cập vào hệ thống của máy chủ...

1.3.2. Can thiệp vào tham số trên URL

Dùng các câu lệnh SQL để khai thác CSDL trên server bị lỗi, điển hình nhất là tấn công bằng "SQL INJECT". Khi hacker gửi những dữ liệu (thông qua các form), ứng dụng web sẽ thực hiện và trả về trình duyệt kết quả câu truy vấn hay những thông báo lỗi có liên quan đến CSDL và từ đó có thể điều khiển toàn bộ hệ thống ứng dụng.

1.3.3. Tấn công dùng cookie

Cookie là những phần dữ liệu nhỏ có cấu trúc được chia sẻ giữa trình chủ và trình duyệt của người dùng chứa thông tin về người dùng đã ghé thăm trang web và những vùng mà họ đã đi qua trong trang web và lưu trên đĩa cứng của máy tính. Những người biết tận dụng những thông tin này có thể gây nên những hậu quả nghiêm trọng.

1.3.4. Các lỗ hổng bảo mật

Hiện nay các lỗ hổng được phát hiện ra ngày càng nhiều về hệ điều hành, máy chủ web, và các phần mềm của các hãng khác... Tuy khi phát hiện ra được và có bổ sung các bảng path nhưng các khách hàng không cập nhật thường xuyên nên là cơ hội cho các hacker tấn công.

1.3.5. Cấu hình không an toàn

Các phần mềm và hệ điều hành trên máy chủ không được cập nhật với bản vá lỗi bảo mật mới nhất, không phân đúng quyền cho các thư mục và tập tin trong trang web, những chức năng quản lý và debug được triển khai không cần thiết, phần mềm web server đăng quá nhiều thông tin trong trang báo lỗi, cấu hình SSL và các hàm mã hóa không đúng.

1.3.6. Tràn bộ đệm

Tin tặc gửi một đoạn mã được thiết kế đặc biệt đến ứng dụng, tin tặc có thể làm cho ứng dụng web thi hành bất kỳ đoạn mã nào, điều này tương đương với việc chiếm quyền làm chủ máy server.

1.3.7. Tấn công từ chối dịch vụ DoS (Denial of Service)

Là các cuộc tấn công trên hệ thống mạng nhằm ngăn cản những truy xuất tới các dịch vụ, làm cho các dịch vụ mạng bị tê liệt, không còn khả năng đáp ứng được yêu cầu bằng cách làm tràn ngập số lượng kết nối, quá tải server hoặc ứng dụng chạy trên server.

1.4. Các vấn đề bảo mật ứng dụng web

1.4.1. Giao thức IPSec

IPSec là một tiêu chuẩn nhằm bổ sung an toàn cho Internet,

được xác định trong RFC 1825, so với giao thức IP, IPSec thêm vào hai trường mào đầu IP để cung cấp tính năng xác thực và bảo mật tại lớp IP.

IPSec có hai cơ chế mã hoá là Tunnel Mode và Transport Mode, sử dụng thuật toán mã hoá đối xứng để mã hoá truyền thông. Các thực thể tham gia truyền thông sử dụng các khoá chia sẻ được tạo ra bằng thuật toán Diffie-Hellman kèm với việc xác thực để đảm bảo khoá đối xứng được thiết lập giữa những bên truyền thông một cách chính xác.

Trước khi IPSec trao đổi dữ liệu đã xác thực hoặc mã hoá, cả bên gửi và bên nhận phải thống nhất với nhau về thuật toán mã hoá và khoá (hoặc các khoá) sử dụng.

1.4.2. Giao thức SSL và TLS

1.4.2.1. Lịch sử SSL, TLS

Giao thức SSL đầu tiên do Netscape phát triển để bảo mật dữ liệu gửi và nhận trên Internet của các giao thức thuộc lớp ứng dụng như HTTP, LDAP hay POP3. Các phiên bản gồm: SSL 1.0, SSL 2.0 - 1994, SSL 3.0 - 1996. SSL nhanh chóng hoàn thiện qua các phiên bản sau đó nó trở thành giao thức phổ biến nhất cho an toàn truyền thông trên WWW. TLS v1.0 (được biết đến như là SSL v3.1)-1999 tuy nhiên các giao thức hoạt động không được đầy đủ. TLS còn được nâng cấp hơn nữa các chức năng qua các phiên bản TLS v1.1 vào năm 2006 và TLS v1.2 vào năm 2008.

1.4.2.2. Nhiệm vụ và kiến trúc SSL

a. Nhiệm vụ: Xác thực server, xác thực client, mã hoá kết nối.

b. Kiến trúc SSL: gồm 4 giao thức con sau: SSL Handshake, SSL Change Cipher Spec, SSL Alert, SSL Record Layer.

SSL là một lớp (bảo mật) trung gian giữa lớp vận chuyển và lớp ứng dụng. SSL được xếp lớp lên trên một dịch vụ vận chuyển định hướng nối kết và đáng tin cậy, SSL nằm trong tầng ứng dụng của giao thức TCP/IP.

- SSL Record Protocol: Sử dụng để trao đổi tất cả các kiểu dữ liệu trong một phiên – bao gồm các thông điệp, dữ liệu của các giao thức SSL khác và dữ liệu của ứng dụng. SSL Record Protocol liên quan đến việc bảo mật và đảm bảo toàn vẹn dữ liệu, mục đích là thu nhận những thông điệp mà ứng dụng chuẩn bị gửi, phân mảnh dữ liệu cần truyền, đóng gói, bổ xung header tạo thành một đối tượng bản ghi được mã hoá và có thể truyền bằng giao thức TCP.

- Handshake Protocol: Giao thức này được sử dụng để khởi tạo phiên SSL giữa client và server, nhờ giao thức này các bên sẽ xác thực lẫn nhau và thoả thuận các tham số cho phiên làm việc sẽ được thiết lập.

- Alert Protocol: Sử dụng để mang các thông điệp của phiên liên quan tới việc trao đổi dữ liệu và hoạt động của các giao thức.

- Change Cipher Spec Protocol: Chứa một thông điệp mang giá trị 1 làm chuyển trạng thái của một phiên từ "đang chờ" sang "bền vững".

1.4.2.3. Hoạt động của SSL

Khi trình duyệt của một máy khách đến một Website bí mật của một máy chủ, máy chủ gửi một lời chào tới trình duyệt. Trình duyệt đáp lại bằng một lời chào. Việc tiến hành trao đổi lời chào, hoặc bắt tay cho phép 2 máy tính quyết định các chuẩn mã hoá và nén (mà chúng cùng hỗ trợ).

Trình duyệt máy khách yêu cầu máy chủ đưa ra một chứng chỉ số. Máy chủ gửi cho trình duyệt một chứng chỉ đã được công nhận bởi CA. Trình duyệt kiểm tra chữ ký số có trên chứng chỉ của máy chủ, dựa vào khoá công khai của CA, khoá này được lưu giữ trong trình duyệt. Hoạt động này xác thực máy chủ thương mại.

Máy khách và máy chủ thoả thuận rằng mọi trao đổi phải được giữ bí mật, bởi vì những thông tin này là quan trọng. Để thực hiện bí mật, SSL sử dụng mã hoá khoá công khai (không đối xứng) và mã hoá khoá riêng (đối xứng). Thoạt đầu, trình duyệt sinh ra một khoá

riêng dùng chung cho cả hai. Sau đó, trình duyệt mã hoá khoá riêng bằng khoá công khai của máy chủ. Khoá công khai của máy chủ được lưu giữ trong chứng chỉ số, máy chủ gửi chứng chỉ này cho trình duyệt trong quá trình xác thực. Một khi khoá được mã hoá, trình duyệt gửi nó cho máy chủ. Ngược lại, máy chủ giải mã thông báo bằng khoá riêng của nó và tìm ra khoá riêng dùng chung. Tất cả các thông báo giữa máy khách và máy chủ được mã hoá bằng khoá riêng dùng chung (cũng được biết đến như là một khoá phiên).

Sau khi kết thúc phiên giao dịch, khoá phiên bị huỷ bỏ. Một kết nối mới lại bắt đầu tương tự.

1.4.2. Giao thức SET

1.4.2.1. Giới thiệu tổng quan về SET

SET là một giao thức chuẩn để đảm bảo an toàn cho các giao dịch thẻ tín dụng trong các mạng không an toàn và Internet. SET không phải là một hệ thống thanh toán mà là một bộ các giao thức và khuôn dạng an toàn cho phép người sử dụng triển khai cơ sở hạ tầng thanh toán bằng thẻ tín dụng trên một mạng một cách an toàn.

1.4.2.2. Nguyên tắc thanh toán bằng thẻ với giao thức SET

Trong giao thức SET, có 5 thực thể gồm: Chủ thẻ, Thương nhân, Ngân hàng thương nhân, Tổ chức cấp thẻ, Cơ quan chứng thực.

1.4.2.3. Mô tả một quá trình giao dịch

Bước 1: Chủ thẻ và thương nhân đăng ký với một CA để nhận được chứng chỉ số.

Bước 2: Khách hàng duyệt website và đặt mua hàng với hình thức thanh toán là SET.

- Thương nhân gửi một bản sao chứng chỉ của mình để khách hàng xác minh rằng đây là một kho hàng hợp lệ. Thương nhân cũng gửi chứng chỉ số của ngân hàng thanh toán.

- Khách hàng nhận và xác minh chứng chỉ của Thương nhân để khẳng định thương nhân đó có hợp lệ hay không.

- Khách hàng gửi thông điệp đặt hàng cho thương nhân bao gồm thông tin đặt hàng, thông tin thanh toán và thông tin để đảm bảo thanh toán chỉ có thể được thực hiện với lệnh cụ thể này. Thông tin đặt hàng được mã hoá bằng khoá công khai của Thương nhân, còn thông tin thanh toán được mã hoá bằng khoá công khai của ngân hàng.

- Thương nhân xác minh khách hàng và yêu cầu Ngân hàng thương nhân uỷ quyền thanh toán bằng cách gửi lệnh đặt hàng bao hàm khoá công khai của ngân hàng, thông tin thanh toán của khách hàng và chứng chỉ của Thương nhân.

- Ngân hàng thương nhân xác minh và uỷ quyền thanh toán.

- Thương nhân xác nhận lệnh và giao hàng và điền uỷ quyền này vào lệnh rồi gửi xác nhận lệnh cho khách hàng, sau đó giao hàng cho khách hàng. Thông qua ngân hàng thanh toán yêu cầu ngân hàng của chủ thẻ thanh toán. Để khởi động thanh toán, thương nhân tạo và ký một yêu cầu cầm giữ và gửi cho cổng nối thanh toán. Do đã có uỷ quyền, cổng nối thanh toán chuyển yêu cầu cầm giữ thành nguồn tiền chuyển vào tài khoản của Thương nhân.

- Tổ chức cấp thẻ in hoá đơn thẻ tín dụng cho khách hàng.

1.4.2.4. Mã hóa SET

a. Sử dụng khóa đối xứng: Gói dữ liệu được mã hóa bằng cách dùng một khóa đối xứng ngẫu nhiên (DES 56 bit). Khóa này được mã hóa với khóa công khai (RSA) trong thông báo của người nhận. Kết quả thu được gọi là "Phong bì số" của thông báo.

b. Sử dụng khóa bất đối xứng, chữ ký số:

+ Mật mã khóa phi đối xứng: Mật mã khóa công khai được dùng để mã hóa các khóa DES và dùng để xác thực, mỗi lần SET thực hiện xử lý dùng hai cặp khóa bất đối xứng: một cặp khóa trao đổi để mã hóa và giải mã khóa phiên, và một cặp "signature" để tạo và xác minh các chữ ký số (160 bit).

+ Chữ ký số: Nhằm bảo đảm tính xác thực và toàn vẹn của

thông báo, người nhận ký số có thể chắc chắn rằng thông báo thật sự đến từ người gửi.

+ Chứng chỉ số: Dùng để xác nhận bên tham gia, CA sẽ tạo ra một thông báo chứa tên của người tham gia và khóa công khai của nó.

+ Chữ ký kép: Chữ ký kép liên kết 2 thông điệp dành cho hai đối tượng nhận khác nhau gồm thông tin đặt hàng OI cho thương nhân và thông tin thanh toán PI cho ngân hàng.

1.4.2.5. Ưu điểm của SET

Đảm bảo tính chính xác của thông tin cho bên gửi và bên nhận, sự toàn vẹn của thông tin trong quá trình truyền dữ liệu thông qua việc sử dụng chữ ký số, khó bẻ khoá, bảo vệ tất cả những người tham gia hợp pháp trong giao dịch và sử dụng một cách an toàn nhất, hạn chế tình trạng từ chối dịch vụ và lừa đảo qua mạng do có cơ chế xác thực cả hai phía.

1.4.2.6. Hạn chế của SET

Yêu cầu phần mềm, phần cứng chuyên dụng với chi phí cao, độ trễ khi giao dịch do tính phức tạp của các thuật toán mã hóa công khai và thường xuyên tiến hành giao dịch với các ngân hàng trung gian, hệ thống cồng kềnh và quá trình giao dịch chậm, các tổ chức tài chính phải trả thêm phí cài đặt và duy trì PKI cho CA, các giao dịch dựa trên tài khoản như: séc điện tử không hỗ trợ trong SET.

1.4.3. So sánh giữa SET và SSL

SSL: Không sử dụng cổng nối thanh toán và Thương nhân nhận được cả thông tin về việc đặt hàng lẫn thông tin thẻ tín dụng, thực hiện xác thực tại thời điểm khởi đầu của mỗi phiên, không yêu cầu cơ quan chứng thực gốc.

SET: Giấu thông tin về thẻ tín dụng của khách hàng đối với Thương nhân và cùng giấu thông tin về đơn hàng đối với các ngân hàng để bảo vệ việc riêng tư, xác thực tại mỗi lần yêu cầu/đáp ứng, Yêu cầu cơ quan chứng thực gốc và kiến trúc phân cấp.

CHƯƠNG 2 - HỆ MẬT MÃ, MÃ KHOÁ ĐỐI XỨNG, MÃ KHOÁ CÔNG KHAI, CHỮ KÝ SỐ

2.1. Tổng quan về mật mã học

2.1.1. Giới thiệu về mật mã học

Mật mã học là ngành khoa học ứng dụng toán học vào việc biến đổi thông tin thành một dạng khác với mục đích che dấu nội dung, ý nghĩa thông tin cần mã hóa. Đây là một ngành quan trọng và có nhiều ứng dụng trong đời sống xã hội.

Có 4 loại hệ mật mã sau: Hệ mật mã dòng, hệ mật mã khối đối xứng, hệ mật mã có hồi tiếp mật mã, hệ mật mã khoá công khai.

2.1.2. Sơ lược về lịch sử của mật mã học

Năm 1949 khi Claude Shannon đưa ra lý thuyết thông tin từ đó một loạt các nghiên cứu quan trọng của ngành mật mã học đã được thực hiện chẳng hạn như các nghiên cứu về mã khối, sự ra đời của các hệ mã mật khoá công khai và chữ ký điện tử.

Đầu những năm 1970 là sự phát triển của các thuật toán mã hoá khối đầu tiên: Lucipher và DES. DES sau đó đã có một sự phát triển ứng dụng rực rỡ cho tới đầu những năm 90. Cuối những năm 1970 thuật toán mã hoá khoá công khai của Whitfield Diffie và Martin Hellman làm nền tảng cho sự ra đời của các hệ mã hoá công khai và các hệ chữ điện tử.

2.1.3. Định nghĩa Hệ mật mã

Một hệ mật là bộ 5 (P, C, K, E, D) thoả mãn các điều kiện sau:

1) P là tập hữu hạn các bản rõ có thể

2) C là tập hữu hạn các bản mã có thể

3) K là tập hữu hạn các khoá có thể

4) Đối với mỗi $k \in K$ có một quy tắc mã hoá $e_k \in E, e_k : P \to C$

và một quy tắc giải mã tương ứng: $d_k \in D, d_k : C \to P$

sao cho: $d_k(e_k(x)) = x$ với $\forall x \in P$.

Tính chất 4 là tính chất quan trọng nhất của mã hoá, nếu mã hoá bằng e_k và bản mã nhận được sau đó được giải mã bằng hàm d_k thì kết quả nhận được phải là bản rõ ban đầu x. Hàm $e_k(x)$ phải là một đơn ánh vì nếu không thì sẽ không giải mã được. Vì nếu tồn tại x_1 và x_2 sao cho $y=e_k(x_1)=e_k(x_2)$ thì khi nhận được bản mã y sẽ không biết nó được mã từ x_1 hay x_2.

Trong một hệ mật bất kỳ ta luôn có $|C| \geq |P|$ vì mỗi quy tắc mã hoá là một đơn ánh. Khi $|C| = |P|$ thì mỗi hàm mã hoá là một hoán vị.

2.1.4. Mô hình truyền tin cơ bản của mật mã học và luật Kirchoff

Người gửi S muốn gửi một thông điệp X tới người nhận R, S mã hoá X tạo ra một đoạn văn bản được mã hoá Y không thể đọc được sử dụng khoá K_1. Giải mã là quá trình ngược lại cho phép người nhận thu được thông tin X ban đầu từ đoạn mã hoá Y sử dụng khoá giải mã K_2.

2.1.5. Một số ứng dụng của mật mã học

Ứng dụng của mật mã học gồm: Bảo mật, xác thực, toàn vẹn, dịch vụ không thể chối từ.

2.2. Các hệ mật mã khoá đối xứng

2.2.1. Hệ mật mã cổ điển

Các hệ mã cổ điển gồm: Mã dịch chuyển, mã thay thế, mã Apphin, mã Vigenère, mã hill, mã hoán vị, mã dòng...

Thuật toán đối xứng hay còn gọi thuật toán mã hoá cổ điển là thuật toán mà khoá mã hoá có thể tính toán ra được từ khoá giải mã.

$E_K(P)=C$

$D_K(C)=P$

K1 có thể trùng K2 hoặc K1 có thể tính toán từ K2, hoặc K2 có thể tính toán từ K1.

2.2.2. Hệ mật mã chuẩn DES

2.2.2.1. Tổng quan

DES do IBM phát triển và được công bố vào năm 1975.

DES là thuật toán mã hoá khối, hay chính là mã hoá một khối dữ liệu 64 bit bằng một khoá 56 bit. Một khối bản rõ 64 bit đưa vào thực hiện, sau khi mã hoá dữ liệu ra là một khối bản mã 64 bit. Cả mã hoá và giải mã đều sử dụng cùng một thuật toán và khoá.

2.2.2.2. Mô tả DES

- Bước 1: Với bản rõ cho trước x, một xâu bit x_0 sẽ được xây dựng bằng cách hoán vị các bit của x theo phép hoán vị cố định ban đầu IP. Ta viết: $x_0 = IP(x) = L_0R_0$, trong đó L_0 gồm 32 bit đầu và R_0 là 32 bit cuối.

- Bước 2: Tính toán 16 lần lặp theo một hàm xác định. Ta sẽ tính L_iR_i, với $1 \leq i \leq 16$ theo quy tắc sau:

$L_i = R_{i-1}$; $R_i = L_{i-1} \oplus f(R_{i-1}, k_i)$.

- Bước 3: Áp dụng phép hoán vị ngược IP^{-1} cho xâu bit $R_{16}L_{16}$, ta thu được bản mã y. Tức là $y = IP^{-1}(R_{16}L_{16})$. Hãy chú ý thứ tự đã đảo của L_{16} và R_{16}

2.2.2.3. Giải mã DES

DES sử dụng cùng chức năng để giải mã hoặc mã hoá một khối. Chỉ có sự khác nhau đó là các khoá phải được sử dụng theo thứ tự ngược lại. Nghĩa là, nếu các khoá mã hoá cho mỗi vòng k_1, $k_2 \ldots k_{16}$ thì các khoá giải mã là k_{16}, $k_{15} \ldots k_1$. Thuật toán dùng để sinh khoá được sử dụng cho mỗi vòng theo kiểu vòng quanh. Khoá được dịch phải, và số những vị trí được tính từ cuối bảng lên thay vì từ trên xuống.

2.2.2.4. Ứng dụng DES

Ứng dụng cho các văn bản trong giao dịch ngân hàng sử dụng các tiêu chuẩn được hiệp hội các ngân hàng Mỹ phát triển. DES được sử dụng để mã hoá các số định danh cá nhân (Pins) và việc chuyển khoản bằng máy thủ quỹ tự động (ATM). DES cũng được dùng để xác thực các giao dịch trong hệ thống chi trả giữa các nhà băng của ngân hàng hối đoái (CHIPS). DES còn được sử dụng rộng rãi trong cá tổ chức chính phủ như: Bộ năng lượng, Bộ tư pháp và Hệ thống

lưu trữ liên bang.

2.3. Hệ mật mã khoá công khai

2.3.1. Giới thiệu về mật mã khóa công khai

Năm 1976 Diffie và Hellman đã đưa ra hệ mã hoá công khai hay hệ mã hoá phi đối xứng, khoá sử dụng vào việc mã hoá là khác so với khoá giải mã và khoá giải mã không thể tính toán được từ khoá mã hoá. Người gửi A có được khoá công khai của người nhận B và có bản tin P cần gửi đi thì có thể dễ dàng tạo ra được bản mã C.

$C = E_{KB}(P) = E_B(P)$

Người nhận B khi nhận được bản tin mã hóa C với khoá bí mật k_B thì có thể giải mã bản tin trong thời gian đa thức.

$P = D_{kB}(C) = D_B[E_B(M)]$

Một số hệ mật khoá công khai quan trọng gồm: Hệ mật RSA, Hệ mật xếp ba lô Merkle - Hellman, Hệ mật McEliece, Hệ mật ElGamal, Hệ mật Chor-Rivest, Hệ mật trên các đường cong Elliptic.

2.3.2. Hệ mật RSA

2.3.2.1. Mở đầu

Hệ mật RSA được mô tả như sau: Ta có sơ đồ chung của hệ mật mã khoá công khai được cho bởi:

$S=(P, C, K, E, D)$ (1)

Trong đó P là tập ký tự bản rõ, C là tập ký tự bản mã, K là tập các khoá k, mỗi khoá k gồm có hai phần k=(k', k"), k' là khoá công khai dành cho việc lập mật mã, còn k" là khoá bí mật dành chi việc giải mã. Với mỗi ký tự bản rõ $x \in P$, thuật toán lập mã E cho ta ký tự mã tương ứng $y=E(k', x) \in C$, và với ký tự mã y thuật toán giải mã D sẽ cho ta lại ký tự bản rõ x: D(k", y)=D(k", E(k', x))=x.

Để xây dựng một hệ mật mã khoá công khai RSA, ta chọn trước một số nguyên n=p.q là tích của hai số nguyên tố lớn, chọn một số e sao cho $\gcd(e,\phi(n))=1$, và tính số d sao cho: $e.d \equiv 1 (\mathrm{mod}(\phi(n))$

Mỗi cặp k=(k',k"), với k'=(n,e) và k"=d sẽ là một cặp khoá của một hệ mật mã RSA cụ thể cho một người tham gia.

Như vậy, sơ đồ chung của hệ mật mã RSA được định nghĩa bởi danh sách (1), trong đó:

P=C=Z_n, trong đó n là một số nguyên Blum, tức là tích của hai số nguyên tố;

K={k=(k', k"): k'=(n, e) và k"=d, gcd(e, $\phi(n)$)=1, e.d $\equiv$ 1(mod$\phi(n)$)};

E và D được xác định bởi:

E(k', x) = x^e mod n, với mọi x $\in$ P

D(k", y) = y^d mod n, với mọi y $\in$ C

2.3.2.2. Thực hiện hệ mật mã RSA

Để thực hiện hệ mật mã RSA cho một mạng truyền tin bảo mật, ngoài việc xây dựng các chương trình tính toán hàm E (với tham biến đầu vào là n, e và x) và hàm D (với tham biến đầu vào là n, d và y), ta còn phải chọn cho mỗi người tham gia một bộ (n,e,d) để tạo các khoá công khai k' và khoá bí mật k". Hệ mã của mỗi người tham gia chỉ có khả năng bảo mật khi n=p.q là số nguyên rất lớn (và do đó p, q cũng phải là những số nguyên tố rất lớn); rất lớn có nghĩa là p, q phải có biểu diễn thập phân cỡ hơn 100 chữ số, do đó n có cỡ hơn 200 chữ số thập phân, hay n$\geq 10^{200}$.

2.3.2.3. Tính bảo mật của mật mã RSA

Bài toán thám mã (khi chỉ biết bản mã) đối với mật mã RSA là: biết khoá công khai k'=(n,e), biết bản mã y=x^e mod n, tìm x. Với bài toán này có độ khó tương đương với bài toán phân tích số nguyên (Blum) thành thừa số nguyên tố. Do đó, giữ tuyệt mật d, hay giữ tuyệt mật các thừa số p, q là có ý nghĩa rất quyết định đến việc bảo vệ tính an toàn của hệ mật mã RSA.

Bên cạnh đó có một số sơ hở mà người thám mã có thể lợi dụng để tấn công như: dùng môđun n chung, dùng số mũ lập mã e nhỏ, lợi dụng tính nhân của hàm lập mã, tấn công bằng cách lặp phép mã.

2.3.2.4. Ứng dụng của RSA

Hệ mã hóa RSA được ứng dụng rộng rãi chủ yếu cho web và

các chương trình email, các công nghệ bảo mật sử dụng cho thương mại điện tử.

2.4. Chữ ký số

2.4.1. Khái niệm về chữ ký số

Một sơ đồ chữ ký số là bộ 5 (P, A, **K**, S,V) thoả mãn các điều kiện dưới đây:

1) P là tập hữu hạn các bức điện (thông điệp, bản rõ) có thể.

2) A là tập hữu hạn các chữ ký có thể.

3) K là tập không gian khoá (tập hữu hạn các khoá có thể)

4) Với mỗi khoá K∈**K** tồn tại một thuật toán ký $sig_K \in S$ và một thuật toán xác minh $ver_K \in V$. Mỗi sig_K: P→A và ver_K: PxA → {TRUE, FALSE} là những hàm sao cho mỗi bức điện x∈P và mỗi chữ ký y ∈ A thoả mãn phương trình dưới đây nếu:

$$ver(x,y) = \begin{cases} TRUE \text{ nõu } y = sig(x) \\ FALSE \text{ nõu } y \neq sig(x) \end{cases}$$

Với mỗi K ∈ **K**, hàm sig_K và ver_K là các hàm đa thức thời gian. Hàm ver_K sẽ là hàm công khai còn hàm sig_K là hàm bí mật. Không thể dễ dàng tính toán để giả mạo chữ ký của B trên bức điện x, nghĩa là với x cho trước chỉ có B mới có thể tính được y để ver(x,y)=TRUE.

2.4.2. Hệ chữ ký RSA

Cho n= p*q, trong đó p,q là các số nguyên tố. Đặt P=A=Z_n và định nghĩa:

K={(n,p,q,a,b): n=p*q, p và q là các số nguyên tố, $ab \equiv 1 (\text{mod } \phi(n))$}.

Các giá trị n và b là công khai; còn q, p, a là bí mật.

Với K=(n, p, q, a, b), ta xác định:

Hàm ký: $sig_K(x)=x^a \bmod n$

và kiểm tra chữ ký: **$ver_K(x,y)$ = TRUE $\Leftrightarrow x \equiv y^b \pmod n$** với x,y ∈ Z_n.

CHƯƠNG 3 - MỘT SỐ CHUẨN AN TOÀN WEB

3.1. Chuẩn mã hoá XML (XMLEnc)

3.1.1. Giới thiệu về XML

XML ra đời vào tháng 2 năm 1998 cho phép người dùng có thể tự định nghĩa các thẻ.

Các thành phần của XML gồm: Khai báo, chú thích, phần tử (Elements), phần tử gốc (Root), thuộc tính (Attributes).

3.1.2. Mã hoá XML

XML cung cấp cơ chế mã hoá: một phần dữ liệu sẽ được trao đổi, các phiên giao dịch an toàn giữa nhiều hơn hai bên. Mỗi bên có thể duy trì trạng thái bảo mật hoặc không bảo mật với bất cứ nhóm giao tiếp nào. Cả dữ liệu bảo mật và không bảo mật đều có thể được trao đổi trong cùng văn bản.

3.1.3. Các cách mã hoá XML

3.1.3.1. Mã hóa các tài liệu trọn vẹn với XML Encryption

Cấu trúc biểu diễn dữ liệu mã hoá toàn bộ tệp có các phần tử chính như sau:

```
<?xml version='1.0' ?>
<EncryptedData xmlns='http://www.w3.org/2001/04/xmlenc#'
   Type='http://www.isi.edu/in-notes/iana/assignments/media-types/text/xml'>
        <CipherData><CipherValue>A23B45C56</CipherValue>
        </CipherData>
</EncryptedData>
```

+ Phần tử <*EncryptedData*>: Là phần tử lớn nhất.

+ Thuộc tính *Type* chỉ ra kiểu dữ liệu được mã hoá.

+ Thuộc tính *xmlns*, chỉ ra không gian tên sử dụng để mã hóa

+ Phần tử <*CipherData*> chỉ ra dữ liệu được mã hoá, chứa giá trị dữ liệu trong phần tử con <*CipherValue*>.

3.1.3.2. Mã hóa một phần tử đơn với XML Encryption

```
<EncryptedData Type='http://www.w3.org/2001/04/xmlenc#Element'
                    xmlns='http://www.w3.org/2001/04/xmlenc#'>
                       <CipherData>
```

```
                    <CipherValue>A23B45C564587</CipherValue>
                </CipherData>
</EncryptedData>
```

Thuộc tính *Type*:

http://www.w3.org/2001/04/xmlenc#Element, không còn sử dụng kiểu IANA nữa mà thay vào đó Sử dụng kiểu mà XML Encryption đã chỉ ra. *#Element* có nghĩa là *EncryptedData* - nó thay thế một phần tử.

3.1.3.3. Mã hóa nội dung của một phần tử

```
<EncryptedData Type='http://www.w3.org/2001/04/xmlenc#Content'
                xmlns='http://www.w3.org/2001/04/xmlenc#'>
    <CipherData> <CipherValue>A23B45C564587</CipherValue>
    </CipherData>
</EncryptedData>
```

Sử dụng http://www.w3.org/2001/04/xmlenc#Content làm giá trị thuộc tính Type.

3.1.3.4. Mã hóa dữ liệu không phải XML

```
<?xml version='1.0' ?>
<EncryptedData xmlns='http://www.w3.org/2001/04/xmlenc#'
    Type='http://www.isi.edu/in-notes/iana/assignments/media-types/jpeg' >
    <CipherData> <CipherValue>A23B45C56</CipherValue> </CipherData>
</EncryptedData>
```

Tệp JPEG hoàn chỉnh là một chuỗi đã được mã hóa các byte và sẽ xuất hiện như là nội dung của phần tử CipherValue.

Thuộc tính *Type* của phần tử *EncryptedData* bao gồm kiểu IANA cho định dạng JPEG.

3.2. Chuẩn quản lý khoá XML (XMKS)

3.2.1. Quá trình trao đổi khoá

Bên A gửi khoá công khai của nó cho bên B để trao đổi khoá:

```
<?xml version='1.0' ?>
<SecureCommunicationDemonstration>
<EncryptedKey CarriedKeyName="Muhammad Imran"
        xmlns='http://www.w3.org/2001/04/xmlenc#'>
 <ds:KeyInfo xmlns:ds='http://www.w3.org/2000/09/xmldsig#'>
```

```
        <ds:KeyValue>1asd25fsdf2dfdsfsdfds2f1sd23</ds:KeyValue>
 </ds:KeyInfo>
</EncryptedKey>
</SecureCommunicationDemonstration>
```

- Phần tử *EncryptedKey:* Là phần tử gốc chứa các phần tử *ds:KeyInfo* và *ds:KeyValue*. Các phần tử *ds: KeyInfo* và *ds:KeyValue* thuộc vào không gian tên chữ kí số.

- Phần tử có tên là *KeyValue*: Chứa khoá công khai của A.

- Thuộc tính *CarriedKeyName:* Tên của khóa đang được vận chuyển.

Bên B gửi lại mã khóa bí mật được tạo ngẫu nhiên với khóa công khai của bên A:

```
<?xml version='1.0' ?>
<SecureCommunicationDemonstration>
 <EncryptedKey CarriedKeyName="Imran Ali"
xmlns='http://www.w3.org/2001/04/xmlenc#'>
  <EncryptionMethod Algorithm="http://www.w3.org/2001/04/xmlenc#rsa-1_5"/>
        <CipherData> CipherValue>xyza21212sdfdsfs7989fsdbc</CipherValue>
        </CipherData>
 </EncryptedKey>
</SecureCommunicationDemonstration>
```

- Phần tử *CipherData* và *CipherValue* trong phần tử *EncryptedKey* sẽ vận chuyển các khóa bí mật (đã được mã hóa).

3.2.2. Sử dụng các khóa đã trao đổi

Sau khi đã trao đổi một khóa bí mật thì cần sử dụng khóa đó để mã hóa dữ liệu. Giả sử A gửi đi một đoạn tin XML, dữ liệu được mã bằng khóa bí mật và đặt trong phần tử <*CipherValue*>. A sẽ giải mã khóa bí mật này với khóa riêng của A, A có thể sử dụng khóa bí mật này mã hóa dữ liệu mà A muốn gửi cho B và đặt đoạn mã vào trong phần tử *CipherValue*. Phần tử *ds:KeyInfo* chứa một phần tử *KeyName*. Việc kết hợp này chỉ tới tên của khóa mà A sử dụng cho việc mã hóa dữ liệu.

3.3. Chuẩn chữ ký XML (XMLSig)

- Chuẩn cung cấp tính toàn vẹn dữ liệu, xác thực nguồn gốc và chống chối bỏ, cơ chế biểu diễn dữ liệu được ký số và chữ ký số theo cấu trúc dựa trên cú pháp XML với một số phần tử chính như sau:

```
<Signature>
  <SignedInfo>
    <CanonicalizationMethod />
    <SignatureMethod />
    <Reference>
      <Transforms>
      <DigestMethod>
      <DigestValue>
    </Reference>
    <Reference /> etc.
  </SignedInfo>
  <SignatureValue />
  <KeyInfo />
  <Object />
</Signature>
```

+ Phần tử ***<SignedInfo>***: Chứa thông tin thực sự được ký số gồm: *CanonicalzationMethod*: Thuật toán được sử dụng để hợp quy phần tử *<SignedInfo>* trước khi tóm lược, *SignatureMethod:* Là thuật toán sử dụng để chuyển *SignedInfo* đã được hợp quy thành *SignatureValue*, *Reference*: Chứa một hoặc nhiều tham chiếu tới đối tượng dữ liệu được ký số. Mỗi phần tử *Reference* bao gồm: **Transforms** là một tuỳ chọn, liệt kê các bước thực hiện trên một tài nguyên trước khi được tóm lược, **DigestMethod** là thuật toán được sử dụng để tạo ra *DigestValue*, **Thuộc tính URI** trong *Reference* chỉ ra đối tượng dữ liệu được ký số, **Thuộc tính Type** cũng cung cấp thông tin về tài nguyên mà URI trỏ tới.

+ Phần tử ***<KeyInfo>*** cung cấp thông tin về khoá được sử dụng để kiểm tra chữ ký số, thông qua chứng chỉ, tên khoá hoặc thông tin thoả thuận khoá.

3.4. Một số chuẩn của OASIS

3.4.1. Chuẩn xác thực (SAML)

Chuẩn được đưa ra bởi tổ chức OASIS định nghĩa một nền tảng

cho việc trao đổi các thông tin bảo mật dưới dạng XML. Những thông tin bảo mật này có thể là: các thông tin về chứng thực, các quyết định về phân quyền, hay có thể là những thuộc tính của các đối tượng được biểu diễn dưới dạng XML và được cấp phát bởi các nơi cung cấp chứng thực SAML.

3.4.2. Chuẩn trao quyền và kiểm soát truy nhập (XACML)

XACML là một chuẩn dùng để xác định chính sách XML cho việc kiểm soát quyền truy cập thông tin qua mạng, đại diện cho một tiêu chuẩn công nghiệp thương mại điện tử có thể cung cấp khả năng tương tác đa dạng giữa các hệ thống bằng cách sử dụng các cơ chế kiểm soát truy cập độc quyền khác nhau. XACML cũng là một chuẩn mở rộng có thể phát triển để hỗ trợ trao quyền và cơ chế kiểm soát truy cập. Phiên bản hiện tại của chuẩn là XACML1.

3.4.3. Chuẩn khai thác dịch vụ (SPML)

SPML cho phép biểu diễn và trao đổi thông tin người dùng, thông tin về tài nguyên và yêu cầu cung cấp dịch vụ theo cú pháp XML. Chuẩn đưa ra một số khái niệm liên quan đến quá trình quản lý các thuộc tính, các tài khoản và quyền khai thác dịch vụ.

3.4.4. Chuẩn quyền số (XrML)

XrML cung cấp một phương thức tổng quát nhằm xác định và quản lý các bản quyền và quy định được tích hợp giữa nội dung số và các dịch vụ. XrML hiện nay là ngôn ngữ bản quyền được sử dụng trong nhiều giải pháp quản lý bản quyền số, bao gồm giải pháp DRM của Microsoft và Content Guard của hãng Content Guard.

3.4.5. Chuẩn an toàn dịch vụ web (WS-Security)

WS-Security là nền tảng để giải quyết các vấn đề bảo mật cho các thông điệp SOAP, với ba mục tiêu chính: Sử dụng các security token trong phần đầu của các thông điệp SOAP để hỗ trợ cho việc định danh và chứng thực, sử dụng chuẩn XML-Signature đảm bảo tính toàn vẹn và xác thực của dữ liệu, sử dụng chuẩn XML-Encryption đảm bảo độ tin cậy cho dữ liệu.

CHƯƠNG 4 - AN TOÀN THÔNG TIN TRONG MÔI TRƯỜNG WEB

4.1. Vấn đề an toàn thông tin

An toàn thông tin là vấn đề đặc biệt quan trọng cần phải được đảm bảo an toàn trước việc khai thác thông tin trái phép và cần tập trung vào việc bảo vệ các tài sản khi chúng được chuyển tiếp giữa client và server phải đảm bảo tính toàn vẹn, an toàn và bao gồm cả tính xác thực. Các kỹ thuật đảm bảo cho an toàn giao dịch điện tử chính là sử dụng các hệ mật mã, các chứng chỉ số và sử dụng chữ ký số trong quá trình thực hiện các giao dịch.

4.2. Chứng chỉ số và cơ chế xác thực

4.2.1. Chứng chỉ số

Chứng chỉ số là một tệp tin điện tử được sử dụng để nhận diện một cá nhân, một máy chủ, một công ty, hoặc một vài đối tượng khác và gắn chỉ danh của đối tượng đó với một khoá công khai, để có được chứng chỉ số cần đăng ký những thông tin với nhà cấp chứng chỉ số (CA), một tổ chức có thẩm quyền xác nhận chỉ danh và cấp các chứng chỉ số. Trong chứng chỉ số chứa một khoá công khai được gắn với một tên duy nhất của một đối tượng giúp ngăn chặn việc sử dụng khoá công khai cho việc giả mạo. Ngoài ra chứng chỉ số còn chứa thêm tên của đối tượng mà nó nhận diện, hạn dùng, tên của CA cấp chứng chỉ số đó, mã số thứ tự, và những thông tin khác.

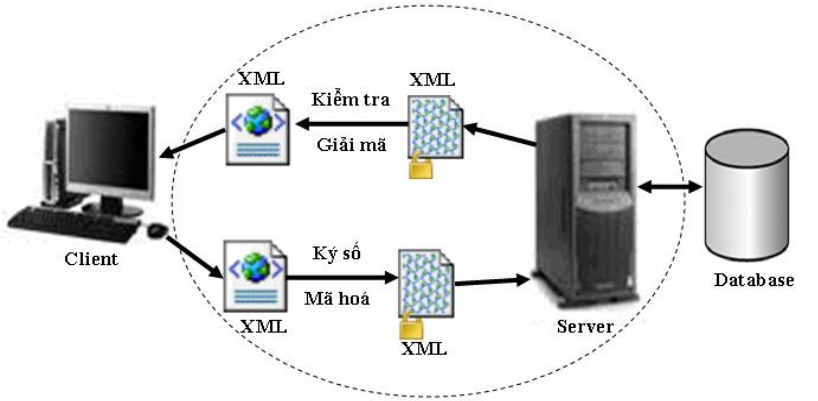
4.2.2. Cơ chế xác thực

Có hai hình thức xác thực máy khách: Xác thực dựa trên tên truy nhập và mật khẩu, xác thực dựa trên chứng chỉ số.

4.3. Mô hình an toàn và bảo mật web

Đảm bảo an toàn thông tin trao đổi trong môi trường web giữa client và máy chủ ứng dụng luận văn áp dụng các chuẩn mã hoá XML, chuẩn quản lý khoá XML và chuẩn chữ ký XML nhằm cung cấp tính xác thực, toàn vẹn, bí mật và chống chối bỏ cho các giao

dịch web.



- Định dạng dữ liệu trao đổi theo khuôn dạng XML
- Áp dụng chuẩn mã hoá XML và chuẩn chữ ký XML
- Áp dụng chuẩn quản lý khoá XML
- Phân phối khoá sử dụng kỹ thuật chứng chỉ khoá công khai
- Kiểm soát truy nhập
- Đảm bảo an toàn, tin cậy và hiệu quả

Hình 4.1 - Mô hình trao đổi dữ liệu an toàn và bảo mật web cho dòng thông tin từ client tới máy chủ ứng dụng

4.4. Cài đặt chức năng an toàn thông tin trên website

Quá trình trao đổi dữ liệu khi upload từ client lên server là các file XML, tại máy client chọn file XML bất kỳ sau đó tiến hành quá trình mã hoá 1 phần dữ liệu hoặc mã hoá toàn bộ file đó theo chuẩn mã hoá XML sử dụng thuật toán mã hoá DES, dữ liệu sau khi mã hoá được biểu diễn theo chuẩn mã hoá XML.

Giao diện thực hiện quá trình trên như sau:

Hình 4.4 - Giao diện thực hiện chức năng Upload file XML từ client lên server

Sau khi thực hiện xong quá trình mã hóa, phía client tiến hành thực hiện ký điện tử lên file, sử dụng phương thức ký RSA-SHA1. Khi thực hiện ký điện tử lên file cần phải có chứng thư số được chứng thực bởi các nhà cung cấp chữ ký điện tử như: viettel, vnpt, bkav, nacecom… File XML sau khi ký sẽ gồm phần dữ liệu mã hoá phía trên và phần thông tin ký điện tử phía dưới được biểu diễn theo chuẩn chữ ký XML.

Hệ thống chỉ chấp nhận các file đã mã hoá và ký số thì mới được upload lên server. Sau khi ký xong phía client có thể xem lại toàn bộ danh sách các file XML, client có thể load file mã hoá hoặc có thể load file giải mã. Hệ thống có giao diện như sau:

Hình 4.5 - Giao diện hiển thị danh sách file sau khi upload

4.5. Một số hàm chính trong chương trình

* File UploadFileSigner.java:

- Hàm **init**(): Khởi tạo chức năng kí điện tử: nhập số pin, kiểm tra số pin của chứng thư số, lấy thông tin certchain trong chứng thư số.

- Hàm **upload**(): Upload file lên server.

- Hàm **encrypt**(String fileName): Mã hóa file có tên là fileName.

- Hàm **encryptChildNodes**(Node nNode, Document document): Mã hóa các node con của file xml.

- Hàm **sign**(int loaiHoSo): Thực hiện kí điện tử vào file.

KẾT LUẬN

KẾT LUẬN:

Luận văn nghiên cứu các hiểm hoạ thường gặp của web, tìm hiểu mô hình và xu thế phát triển của web, một số chuẩn an toàn web, kỹ thuật và công nghệ để giải quyết vấn đề an toàn và bảo mật web hiện nay từ đó phân tích, tổng hợp một số cơ sở mật mã cần thiết để áp dụng các hệ mật một cách tin cậy trong an toàn và bảo mật web, tiến hành xây dựng mô hình an toàn và bảo mật web.

Trong luận văn này, tác giả đã đề cập đến hai kỹ thuật chính trong an toàn thông tin đó là mã hoá và ký số cùng với những vấn đề liên quan đến bảo mật ứng dụng web. Hai kỹ thuật này cũng được ứng dụng trên website góp phần vào việc đảm bảo an toàn thông tin trong quá trình trao đổi dữ liệu.

Về kỹ thuật mã hoá, mã hoá file XML cần trao đổi trong môi trường web theo chuẩn mã hoá XML nhằm đảm bảo an toàn về thông tin giao tiếp nhưng không đảm bảo liệu thông tin có bị giả mạo hoặc có bị mạo danh hay không, do đó luận văn đã nghiên cứu và ứng dụng chữ ký số vào tệp XML sau khi mã hoá. Tác giả cũng đã tìm hiểu phương thức ký RSA-SHA1 sử dụng chứng thư số cho khoá công khai nhằm đảm xác thực tính đúng đắn của đối tác trong quá trình trao đổi. Dữ liệu sau khi ký số được biểu diễn theo chuẩn chữ ký XML.

KIẾN NGHỊ VÀ HƯỚNG PHÁT TRIỂN:

Với bước đầu nghiên cứu cài đặt thử nghiệm chương trình đã tạo tiền đề ứng dụng an toàn dữ liệu trao đổi trong môi trường web và từ đó đưa chương trình vào ứng dụng thực tế. Trong thời gian tới, tôi sẽ tiếp tục phát triển đề tài với phương hướng cụ thể như sau:

Nghiên cứu một số cơ sở mật mã cần thiết để áp dụng các hệ mật một cách tin cậy trong an toàn và bảo mật web vào website thực tế, nghiên cứu và thực nghiệm với một số chuẩn an toàn web của OASIS nhằm cung cấp thêm tính năng bảo mật, cải tiến và nâng cao hiệu quả của các module đã cài đặt trên website cũng như các kỹ thuật cài đặt khác.