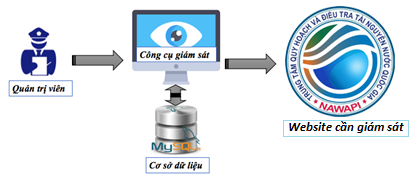
**Tài liệu phân tích thiết kế hệ thống**

1. ***Kiến trúc***



*Mô hình giải pháp đề xuất*

1. ***Chức năng của hệ thống***

Bảng thể hiện tổng quát các chức năng của hệ thống

|  |  |  |
| --- | --- | --- |
| **STT** | **Chức năng** | **Mô tả** |
| 1 | Giám sát web | chụp ảnh trang web; lấy và so sánh nội dung, cấu trúc, giao diện, trích xuất và đưa ra hình ảnh sự thay đổi trong bố cục; sử dụng công nghệ "server send event" để thông báo thời gian thực |
| 2 | xác định và cảnh báo tấn công thay đổi giao diện | * *Xác định sự thay đổi giao diện đó có phải là tấn công hay không?* * *Phân tích tập tin nhật ký tại máy chủ web* |

1. **Chi tiết chức năng của hệ thống**

**3.1. Giám sát web:**

*Đầu vào :* Địa chỉ URL của website.

Quy trình thực hiện

*Bước 1 : Giám sát website theo địa chỉ URL*

Thực hiện: Chụp ảnh trang web, lưu ảnh chụp vào ổ đĩa, lấy cấu trúc HTML (DOM) từ phantomjs, lặp qua cấu trúc DOM và trích xuất các thuộc tính style và nội dung của thẻ. Tính checksum MD5 cho thuộc tính style và nội dung của thẻ trong HTML để so sánh sự khác biệt trong lần sau. Sau đó, lưu các thông tin này dưới dạng file .json .

*Bước 2: Phát hiện thay đổi ở trên giao diện*

Thực hiện : So sánh sự khác nhau (checksum MD5) của style và nội dung so với lần trước của website.

Sau đó trả kết quả trong hai trường hợp sau :

Trường hợp 1: Nội dung hiển thị không thay đổi : lặp lại bước 1 sau khoảng thời gian được cài đặt.

Trường hợp 2: Từ so sánh sự khác nhau (checksum MD5) của style và nội dung sẽ phát hiện 4 trường hợp thay đổi giao diện có thể xảy ra :

1. Thay đổi về văn bản (text)
2. Thay đổi thêm thành phần
3. Thay đổi xóa thành phần
4. Thay đổi cấu trúc giao diện

*Đầu ra :* Có hay không sự thay đổi giao diện bao gồm các thông tin : thời gian phát hiện website có sự thay đổi này, hình ảnh website trước và sau khi bị thay đổi.

## 3.2. xác định và cảnh báo tấn công thay đổi giao diện.

*Đầu vào* : Các thông tin về sự thay đổi giao diện của website

Quy trình thực hiện :

*Bước 1: Xác định sự thay đổi giao diện đó có phải là tấn công hay không?*

Thực hiện :

Từ kết quả phát hiện 4 trường hợp thay đổi giao diện có thể xảy ra xác định tấn công thay đổi giao diện như sau:

* khi có thay đổi (1) (2) (3) (4) đồng thời cùng xuất hiện, như vậy cả về nội dung thông điện và cả cấu trúc của website đã bị thay đổi hoàn toàn. Ta có thể cảnh báo là tấn công thay đổi giao diện vì nội dung hiện lên sẽ là các thông điệp mà tin tặc muốn truyền tải.
* Khi có sự thay đổi (2) và (4) xuất hiện đồng thời, trường hợp này khả năng tin tặc chèn và thực thi mã độc là rất lớn. Ta có thể cảnh báo là website bị tấn công thay đổi giao diện bằng XSS. Tuy nhiên đây là một trong các trường hợp có thể xảy ra.
* Các trường hợp riêng rẽ còn lại không đủ dữ kiện để xác định rằng đó là tấn công thay đổi giao diện hay là những thay đổi hợp lệ từ người quản trị.

Vì vậy, để tăng độ chính xác và tin cậy. ta cần thực hiện phân tích tập tin nhật ký nhằm phát hiện ra các dấu hiệu khác nhằm xác định rằng đây có phải tấn công thay đổi giao diện hay không.

*Bước 2 : Phân tích tập tin nhật ký tại máy chủ web*

Đầu vào :

* Gửi lời gọi đến chức năng phân tích tập tin nhật ký trên máy chủ web
* Thời gian phát hiện thay đổi
* Tập tin nhật ký của máy chủ web apache với cấu trúc Host Ident Authuser Date Request Status Bytes đã giới thiệu ở phần 1.3.1

Thực hiện :

* Từ website thu thập thông tin gửi lời gọi đến chức năng phân tích tập tin nhật ký trên máy chủ web, kèm theo thông tin thời gian phát hiện sự thay đổi.
* Tại máy chủ web : đọc từng dòng của tập tin nhật ký trong khoảng một ngày trước tính từ thời gian phát hiện sự thay đổi và tiến hành phân tích :
* So sánh Host, phát hiện IP bất thường ( IP nước ngoài, IP không nằm trong danh sách hợp lệ đã thực hiện đăng nhập và thao tác)
* Đối với dòng truy cập : Phân tích Request – là dòng yêu cầu từ client - bằng biểu thức chính quy nhằm phát hiện các dấu hiệu tấn công XSS, SQL Injection,…
* **Cơ chế phát hiện tấn công thay đổi giao diện là loại tấn công Cross Site Scripting (XSS)**

Lỗi XSS xảy ra khi ứng dụng web nhận các dữ liệu độc hại và chuyển nó đến trình duyệt cho người dùng mà không xác nhận lại dữ liệu đó có hợp lệ hay không. Kiểu tấn công này cho phép kẻ tấn công thực thi các đoạn mã độc trong trình duyệt của nạn nhân và có thể cướp phiên người dùng hoặc chuyển hướng người dùng đến các trang độc hại khác.Để phát hiện các tấn công XSS, có thể sử dụng các biểu thức chính quy:[3]

(?:,\s\*(?:alert|showmodaldialog|eval)\s\*,)|(?::\s\*eval\s\*[^\s])|([^:\s\w,.\/?+-]\s\*)?(?<![az\/\_@])(\s\*return\s\*)?(?:(?:document\s\*\.)?(?:.+\/)?(?:alert|eval|msgbox|showmod(?:al|eless)dialog|showhelp|prompt|confirm|dialog|open))\s\*(?:[^.az\s\-]|(?:\s\*[^\s\w,.@\/+-]))|(?:java[\s\/]\*\.[\s\/]\*lang)|(?:\w\s\*=\s\*new\s+\w+)|(?:&\s\*\w+\s\*\)[^,])|(?:\+[\W\d]\*new\s+\w+[\W\d]\*\+)|(?:document\.\w)]

Hoặc

(?:=\s\*(?:top|this|window|content|self|frames|\_content))|(?:\/\s\*[gimx]\*\s\*[)}])|(?:[^\s]\s\*=\s\*script)|(?:\.\s\*constructor)|(?:default\s+xml\s+namespace\s\*=)|(?:\/\s\*\+[^+]+\s\*\+\s\*\/)]

* **Cơ chế phát hiện tấn công thay đổi giao diện là loại tấn công SQL Injection**

SQL Injection là một kỹ thuật cho phép những kẻ tấn công lợi dụng lỗ hổng của việc kiểm tra dữ liệu đầu vào trong các ứng dụng web và các thông báo lỗi của hệ quản trị cơ sở dữ liệu trả về để inject (chèn vào) và thi hành các câu lệnh SQL bất hợp pháp. SQL Injection có thể cho phép những kẻ tấn công thực hiện các thao tác, delete, insert, update,… trên cơ sở dữ liệu của ứng dụng, thậm chí là server mà ứng dụng đó đang chạy, lỗi này thường xảy ra trên các ứng dụng web có dữ liệu được quản lý bằng các hệ quản trị cơ sở dữ liệu như SQL Server, MySQL, Oracle, DB2, Sysbase..

Để phát hiện các tấn công SQL Injection, có thể sử dụng các biểu thức chính quy:[3]

(?:"\s\*or\s\*"?\d)|(?:\\x(?:23|27|3d))|(?:^.?"$)|(?:(?:^["\\]\*(?:[\d"]+|[^"]+"))+\s\*(?:n?and|x?or|not|\|\||\&\&)\s\*[\w"[+&!@(),.-])|(?:[^\w\s]\w+\s\*[|-]\s\*"\s\*\w)|(?:@\w+\s+(and|or)\s\*["\d]+)|(?:@[\w-]+\s(and|or)\s\*[^\w\s])|(?:[^\w\s:]\s\*\d\W+[^\w\s]\s\*".)|(?:\Winformation\_schema|table\_name\W)]

Hoặc:

(?:"\s\*\\*.+(?:or|id)\W\*"\d)|(?:\^")|(?:^[\w\s"-]+(?<=and\s)(?<=or\s)(?<=xor\s)(?<=nand\s)(?<=not\s)(?<=\|\|)(?<=\&\&)\w+\()|(?:"[\s\d]\*[^\w\s]+\W\*\d\W\*.\*["\d])|(?:"\s\*[^\w\s?]+\s\*[^\w\s]+\s\*")|(?:"\s\*[^\w\s]+\s\*[\W\d].\*(?:#|--))|(?:".\*\\*\s\*\d)|(?:"\s\*or\s[^\d]+[\w-]+.\*\d)|(?:[()\*<>%+-][\w-]+[^\w\s]+"[^,])]

* **Cơ chế phát hiện tấn công thay đổi giao diện là loại tấn công directory traversal**

Directory Traversal là một dạng tấn công cho phép kẻ tấn công truy cập vào thư mục và file cấm trên web server. Nếu như truy cập này thành công thì kẻ tấn công có thể xem được các file, thư mục cấm và thực thi các câu lên trên web server. Đa số các web server mắc lỗi này đều không kiểm soat đầu vào dữ liệu được gửi từ client. Dạng tấn công này còn có tên gọi khác như Dot-DotSlash, Directory Climbing, Path Traversal và Backtracking.

Để phát hiện các tấn công Directory Traversal, có thể sử dụng các biểu thức chính quy:[3]

(?:(?:\/|\\)?\.+(\/|\\)(?:\.+)?)|(?:\w\.exe\??\s)|(?:;\s\*\w+\s\*\/[\w\*-]+\/)|(?:\d\.\dx\|)|(?:%(?:c0\.|af\.|5c\.))|(?:\/(?:%2e){2})]

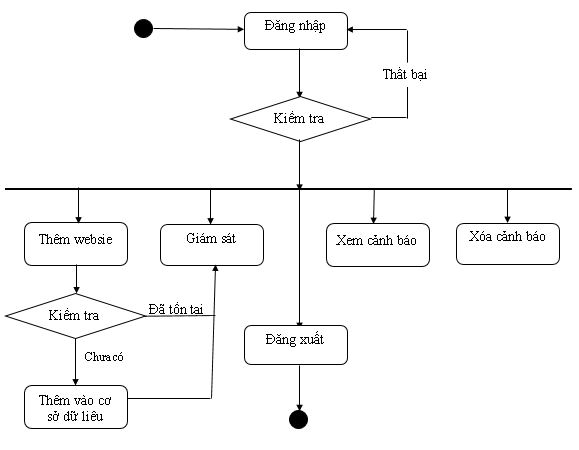
Hoặc:

(?:%c0%ae\/)|(?:(?:\/|\\)(home|conf|usr|etc|proc|opt|s?bin|local|dev|tmp|kern|[br]oot|sys|system|windows|winnt|program|%[a-z\_-]{3,}%)(?:\/|\\))|(?:(?:\/|\\)inetpub|localstart\.asp|boot\.ini)]

*Đầu ra* : Kết luận website có phải bị tấn công thay đổi giao diện hay không

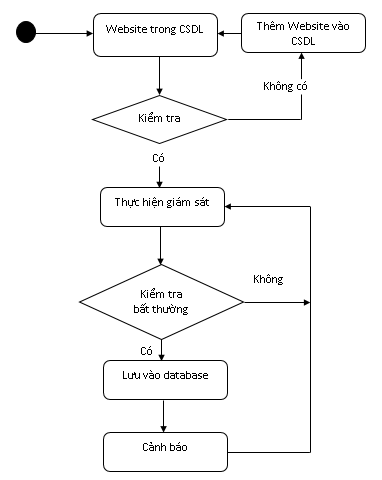
1. ***Sơ đồ chức năng***

* 1. *Chức năng quản trị viên*



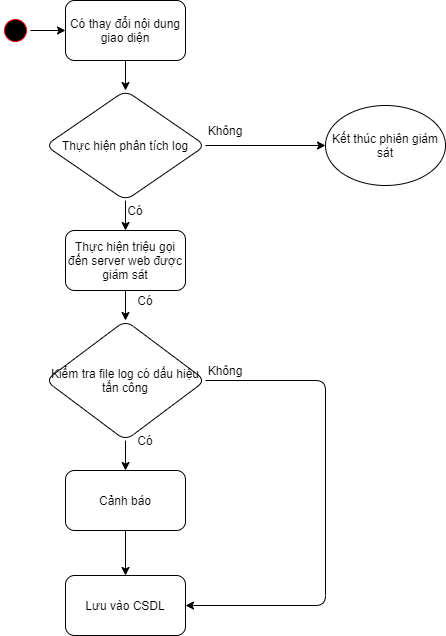
*Thiết kế chức năng quản trị viên*

* 1. *Chức năng giám sát web*

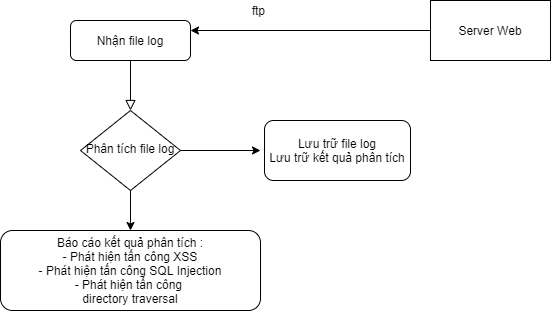


*Thiết kế chức năng giám sát*

### Thiết kế chức năng phân tích tập tin nhật ký

**

*Thiết kế chức năng phân tích tập tin nhật ký*

**

*Hoạt động của công cụ phân tích tập tin nhật ký*