HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Tô Thanh Tú

GIẢI PHÁP CẢNH BÁO KIỂU TẤN CÔNG AN NINH MẠNG DEFACE VÀ HIỆN THỰC

CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN

MÃ SỐ: 8.48.01.04

TÓM TẮT LUẬN VĂN THẠC SĨ

TP.HỒ CHÍ MINH - NĂM 2023

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: **TS. Nguyễn Đức Thái**

(Ghi rõ học hàm, học vị)

Phản biện 1:.....

Phản biện 2:.....

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ..... giờ ..... ngày ..... tháng ..... năm .....

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

Từ khi Internet lần đầu xuất hiện tại Việt Nam, đến nay Internet là đã phát triển rộng khắp đến mọi nơi và trở nên phổ biến. Cùng với đó, lĩnh vực công nghệ thông tin cũng đã phát triển mạnh mẽ và trở thành ngành khoa học đóng vai trò thúc đẩy kinh tế xã hội phát triển. Nhiều cơ hội được tạo ra, nhiều tri thức của nhân loại được tiếp cận do lực lượng lao động từng bước được nâng cao trình độ lao động, hiệu suất làm việc.

Bên cạnh những thuận lợi, các lợi ích mang lại, công nghệ thông tin cũng tạo ra nhiều thách thức. Trong các thách thức về an toàn và an ninh thông tin thì an ninh mạng có vai trò hết sức quan trọng. Các hình thức, tần suất và mức độ tấn công, phá hoại trên mạng Internet đang ngày càng trở lên tinh vi, phức tạp và gây ra nhiều hậu quả nghiêm trọng. Do vậy, đòi hỏi vai trò của người quản trị mạng phải đảm bảo an toàn hệ thống trở nên hết sức cần thiết.

Thông qua việc tìm hiểu về một số kiểu tấn công phổ biến hiện nay, đề tài sẽ tập trung tìm hiểu vào hình thức tấn công an ninh mạng deface (còn được gọi là tấn công làm thay đổi nội dung website). Một loại tấn công phổ biến nhưng có ảnh hưởng nghiêm trọng đối với nhà quản trị website và các tổ chức, doanh nghiệp. Sau đó đề xuất các biện pháp và xây dựng một hệ thống có khả năng đưa ra cảnh báo khi website bị tấn công an ninh mạng.

CHƯƠNG 1: GIỚI THIỆU ĐỀ TÀI

1.1 Tính cấp thiết của đề tài

Hiện nay, tôi đang công tác tại Sở Thông tin và Truyền thông tỉnh Tây Ninh, đây là một cơ quan phụ trách về hoạt động ứng dụng Công nghệ thông tin trên địa bàn tỉnh. Việc phát hiện các cuộc tấn công an ninh mạng, trong đó tấn công an ninh mạng deface là một trong những mối quan tâm nhằm kịp thời phòng, chống lại việc bị up các thông tin sai lệch, các hình ảnh phản động, bôi nhọ lãnh đạo của Đảng, Nhà nước trên website của Sở, ban, ngành của tỉnh. Đến nay, chưa có công cụ nhằm phát hiện cuộc tấn công nêu trên. Và đó là lý do tôi chọn đề tài này.

Từ các kiến thức tìm hiểu được, tôi mong muốn góp một phần nhỏ vào việc nghiên cứu, xác định các dấu hiệu phát hiện cuộc tấn công an ninh mạng deface. Từ đó, xây dựng hệ thống có khả năng kịp thời cảnh báo đến người quản trị hệ thống, giúp cho các website của tỉnh được vận hành ổn định.

1.2 Mục tiêu của đề tài

Tìm hiểu cách thức hoạt động, trình bày của một website, các kỹ thuật tấn công và bảo mật website.

Xác định các dấu hiệu nhằm phát hiện một cuộc tấn công an ninh mạng deface:

- Dấu hiệu phát hiện sự thay đổi tính toàn vẹn: Ứng dụng hàm băm để kiểm tra sự thay đổi của mã nguồn của website trên máy chủ web;

- Dấu hiệu phát hiện sự thay đổi dữ liệu;

- Dấu phát hiện tấn công làm tê liệt website;

Đề xuất xây dựng một hệ thống giám sát website có 02 tính năng như sau:

- Đặt tại máy chủ website nhằm phát hiện kịp thời cuộc tấn công an ninh mạng deface. Dự kiến áp dụng đối với các website của các Sở, ban, ngành trên bàn tỉnh Tây Ninh. Các thông tin cảnh báo sẽ được xử lý để quyết định có gửi thông báo đến người quản trị hay không, tần suất cảnh báo phải phù hợp. Hình thức cảnh báo: Sử dụng email và tin nhắn SMS để thông báo.

- Đặt bên ngoài máy chủ của website để tiếp tục thực hiện chức năng cảnh báo ngay cả khi website bị kẻ gian chiếm quyền điều khiển hoặc phá hủy.

1.3 Phương pháp thực hiện đề tài

- Phương pháp luận: Dựa trên bài báo, đề xuất nhằm phát hiện, chống tấn công làm thay đổi nội dung đã được trình bày, công bố trước đó.

- Phương pháp thống kê: Phương pháp này được áp dụng để tổng hợp, chọn lọc những thông tin, dữ liệu theo đúng yêu cầu đặt ra.

- Phương pháp thực nghiệm: Xây dựng hệ thống và thực nghiệm thuật toán đã đề xuất.

1.4 Cấu trúc luận văn

Luận văn được trình bày 6 Chương, cụ thể như sau:

Chương 1. Giới thiệu đề tài

Chương 2. Những công trình liên quan

Chương 3. Nền tảng lý thuyết liên quan đến tấn công an ninh mạng deface

Chương 4. Đề xuất biện pháp nhằm phát hiện cuộc tấn công an ninh mạng deface

Chương 5. Xây dựng hệ thống giám sát và cảnh báo cuộc tấn công an ninh mạng deface

Chương 6: Kết luận

CHƯƠNG 2: NHỮNG CÔNG TRÌNH LIÊN QUAN

Chương này sẽ trình bày một số công trình, các biện pháp của các nhà khoa học đã thực hiện hoặc đề xuất nhằm phát hiện một trang website đã bị tấn công deface.

Hai tác giả Andrew Cooks và Martin S Olivier (2004) đề xuất giải pháp hạn chế bị tấn công an ninh mạng deface bằng chiến thuật chỉ đọc tại bài báo [1]“curtailing web defacement using a read-only strategy”.

Tại bài báo [2]“Anti Web Site Defacement System (AWDS)” đăng bởi tác giả Mazin S.Al-Hakeem (2010) đã đề xuất một hệ thống có khả năng phát hiện và khôi phục website bằng cách áp dụng thuật toán Rabin’s Fingerprinting cải tiến.

Năm 2010, ba nhà nghiên cứu Bartoli, A.; Davanzo, G. và Medvet, E. tại bài báo [3]“A Framework for Large-Scale Detection of Web Site Defacements” đã đề xuất một biện pháp nhằm phát hiện cuộc tấn công deface ở quy mô lớn dựa vào kỹ thuật phát hiện sự bất thường (anamoly detection technique).

Năm 2011, ba tác giả G. Davanzo, E. Medvet và A. Bartoli đã đề xuất sử dụng kỹ thuật học máy để phát hiện cuộc tấn công an ninh mạng deface [4] “Anomaly detection techniques for a web defacement monitoring service”.

Trong bài báo [5] “An approach to Reveal Website Defacement” của ba tác giả Rajiv Kumar Gurjwa, Divya Rishi Sahu, Deepak Singh Tomar đã trình bày biện pháp nhằm phát hiện tấn công an ninh mạng deface dựa vào việc sử dụng mã CRC 32, kỹ thuật hàm băm (hashing), tỷ suất nhiễu so với tín hiệu (Peak Signal to Noise Ratio - PSNR), cấu trúc tương đồng (Structural Similarity - SSIM) vào năm 2013.

Ramniwas, Nikhil, và Deepak (2014) đề xuất phân tích tập tin ghi nhận sự thay đổi của website [6]: Các thông tin thay đổi sẽ được ghi nhận, xử lý và phân tích nhằm xác định website bị tấn công.

Ebot Enaw và Djoursoubo Pagou Prosper (2014) đề xuất sử dụng trí tuệ nhân tạo để phát hiện cuộc tấn công an ninh mạng tại bài báo [7].

Một phương pháp nhằm phát hiện tấn công an ninh mạng deface thông qua ứng dụng hàm băm vào năm 2015 do Rashmi và Shahzia đề xuất. Xây dựng một môđun tích hợp vào máy chủ web để phát hiện cuộc tấn công trong các website. hệ thống này còn có khả năng tự thay đổi cấu hình để trang web bị tấn công không hiển thị. Đây là một phương pháp mới so

với các phương pháp đã biết trước đây. Tuy nhiên nó cũng chỉ phù hợp cho các trang tĩnh [8].

Các tác giả Francesco Bergadano, Fabio Carretto, Fabio Cogno và Dario Ragno (2019) đề xuất phát hiện cuộc tấn công an ninh mạng deface thông qua biện pháp máy học đối thủ thụ động (Passive Adversaries)[9].

Trên Tạp chí khoa học Đại học Đà Lạt (Tập 8, Số 2, năm 2018), các tác giả Trần Đắc Tốt, Đặng Lê Nam, Phạm Nguyễn Huy Phương đề xuất xây dựng [10]"hệ thống cảnh báo tấn công thay đổi giao diện website" bằng việc kết hợp giám sát trong mạng nội bộ (Local Area Network - LAN) và giám sát từ xa. Ở từng thời điểm xác định, hệ thống sẽ tiến hành giám sát máy chủ, cơ sở dữ liệu, mã nguồn của website để phát hiện sự thay đổi bất hợp pháp.

CHƯƠNG 3: NỀN TẢNG LÝ THUYẾT LIÊN QUAN ĐẾN TẤN CÔNG AN NINH MẠNG DEFACE

3.1 Tổng quan về an ninh mạng

3.1.1. Tìm hiểu về an toàn thông tin và an ninh mạng

An ninh mạng là sự bảo đảm hoạt động trên không gian mạng không gây phương hại đến an ninh quốc gia, trật tự, an toàn xã hội, quyền và lợi ích hợp pháp của cơ quan, tổ chức, cá nhân.

3.1.2. Sự cần thiết phải bảo vệ an toàn thông tin

Cá cá nhân, tổ chức, nhất là doanh nghiệp sẽ hứng chịu nhiều thiệt hại khi gặp các sự cố về an toàn, an ninh mạng như là: Tổn thất về chi phí; Tổn thất về thời gian; Hệ thống hoạt động trì trệ, kém hiệu quả; Tổn thất đến danh dự, uy tín của doanh nghiệp; Mất cơ hội kinh doanh.

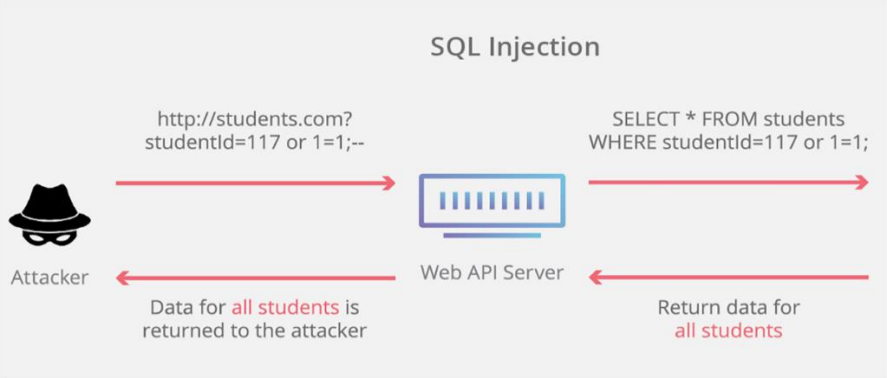
3.2. Một số lỗ hổng an ninh trên môi trường mạng

Lỗ hổng bảo mật có thể được hiểu là các điểm yếu của hệ thống mà kẻ gian có thể khai thác, tấn công nhằm xâm nhập, điều khiển trái phép hệ thống.

3.3 Một số kỹ thuật tấn công và bảo mật website

3.3.1 Kỹ thuật tấn công SQL Injection

Tấn công SQL Injection là kiểu tấn công hệ thống bằng cách đưa vào các câu vấn tin SQL và thực thi bất hợp pháp.



Hình 3.1 Mô hình tấn công SQL Injection

3.3.2 Tấn công XSS (Cross Site Scripting)

XSS là viết tắt của Cross-Site Scripting, đây là một kĩ thuật tấn công mà kẻ tấn công sẽ chèn vào các website động những đoạn mã có thể thực thi một câu lệnh bất hợp pháp.

3.4 Về tấn công an ninh mạng deface

Tấn công Deface (Website Defacement) là hình thức tấn công làm thay đổi giao diện trực quan của một trang web.

Các hậu quả của tấn công an ninh mạng deface gồm [9]:

- Thay đổi một phần hoặc toàn bộ nội dung của trang web.

- Thay đổi mã nguồn của trang web.

- Chuyển hướng của trang web.

- Hủy hoặc xóa toàn bộ trang web.

3.4.1 Nguyên nhân website bị tấn công an ninh mạng deface

- Mật khẩu của tài khoản quản trị yếu, dễ đoán.

- Cài đặt các module, plugin, extension,... trong các mã nguồn mở hiện nay (thường là các website joomla, wordpress,...).

- Để lộ mật khẩu quản trị....

3.4.2 Dấu hiệu nhận biết website bị tấn công an ninh mạng deface

Thông thường, khi các trang mặc định như: home.html, trangchu.html, default.html… bị thay đổi nghĩa là website đã bị tấn công Deface.

3.4.3 Tình hình về tấn công an ninh mạng deface

Theo Bộ thông tin và Truyền thông thống kê cho thấy trong tháng 9/2021, tại Việt Nam đã xuất hiện 1.074 cuộc tấn công an ninh mạng, giảm 6,45% so với tháng trước đó. Qua phân loại, có 192 cuộc Phishing, 743 cuộc Malware và 139 cuộc tấn công an ninh mạng Deface.

3.5 Hàm băm

3.5.1 Khái niệm hàm băm

Hàm băm là hàm chuyển đổi một thông điệp có độ dài bất kỳ thành một dãy ký tự (hoặc dãy bit) nhưng có độ dài cố định.

3.5.2 Tính chất và yêu cầu của hàm băm

Một số tính chất cơ bản của hàm băm là: Tính tất định; Tính một chiều. Tính hiệu quả.

Yêu cầu đối với một hàm băm tốt, an toàn: Tính toán nhanh; Tính kháng va chạm.

3.5.3 Một số hàm băm phổ biến

- Hàm băm MD5.
- Hàm băm SHA-1.
- Hàm băm RIPEMD-160.
- Hàm băm SHA-2.
- Hàm băm SHA-3.
- Hàm băm BLAKE2.

3.5.4. Thuật toán Rabin-Karp

Thuật toán Rabin-Karp là một thuật toán được sử dụng để tìm kiếm hoặc đối sánh các mẫu trong đoạn văn bản.

3.5.5 Thuật toán Rabin-Karp cải tiến

Đầu vào: Tài liệu (trang web công khai)

Đầu ra: Thực hiện băm để lấy dấu vân tay tài liệu.

3.6 Thuật toán đối sánh chuỗi

Đối sánh chuỗi là việc so sánh một hoặc vài chuỗi (thường được gọi là mẫu hoặc pattern) với toàn bộ văn bản để tìm ra nơi và số lần xuất hiện của chuỗi đó trong văn bản.

3.6.1 Phân loại thuật toán đối sánh chuỗi

- Theo thứ tự đối sánh:
- Theo số lượng pattern:
- Theo độ sai khác đối sánh:

- Theo sự thay đổi của pattern và văn bản

3.6.2 Dấu vân tay tài liệu (Document Fingerprint)

Dấu vân của tài liệu là tập hợp các mã được sinh ra từ các khóa nội dung của tài liệu đó. Mỗi mã đó được gọi là một giá trị băm.

3.7 Ứng dụng thuật toán Rabin Karp để so sánh tìm độ tương đồng của 02 tài liệu

3.7.1 Các bước tiền xử lý trước khi thực hiện băm tài liệu

Có 04 bước như sau [11]:

Bước 1-Mã hóa (Tokenizing): Chuyển tài liệu cần băm thành các nhóm ký tự và từ viết hoa sang viết thường.

Bước 2-Loại bỏ các từ thường xuất hiện (Stopword Removal): Bởi vì, với, và, hoặc,...

Bước 3-Đưa các từ có tiền tố (prefix) hoặc hậu tố (suffix) trở lại các từ gốc.

Bước 4-Thực hiện băm tài liệu.

3.7.2 Ví dụ tính mã băm của chuỗi "MEDAN"

Chọn hệ số K-Gram=5; hệ số mũ B (basis)=7; Chuỗi A = MEDAN

A (1) = 77; A(2) = 69; A(3) = 68; A(4) = 65; A(5) = 78

Giá trị băm = $(77 \times 7^5) + (69 \times 7^4) + (68 \times 72^3) + (65 \times 7^2) + (78 \times 7^1)$

= 1.486.863

3.7.3 Ví dụ tính mã băm của một tài liệu

- Giả sử chọn hệ số K-Gram=5, tính giá trị băm của 5 ký tự đầu tiên. Sau đó tính giá trị băm của 5 ký tự kế tiếp.

- Giả sử ta có 02 bảng giá trị băm như hình dưới đây của 02 tài liệu [11]:

Bảng 3.1 Bảng giá trị băm của 02 tài liệu

19875	16830	23124	17433	20546
21489	26753	13498	23846	16528
21848	28447	29994	10301	13009
18832	27217	23157	25854	22492
14952	14337	29348	19978	28809
13485	14188	13131	21215	12053
25669	13809	26508	19455	25356
29964	17723	2663	17445	11803
19477	27142	24814	15155	26266
28432	19007	21896	16625	20681

28432	26406	28424	13930	19187
18049	10867	18516	26753	19975
10152	13053	24120	21896	18351
12605	25101	21215	20750	15513
22949	26006	25045	25932	10695
13254	21504	20286	22492	10615
25565	29941	17403	23018	22666
19744	19769	19877	29535	13139
25669	16830	14297	20916	24640
16960	20681	13131	13009	18947

Mỗi bảng này có 50 giá trị băm, trong đó có 10 giá trị băm là giống nhau. Khi này mức độ tương đồng của 02 tài liệu được tính bằng công thức sau:

$$P = \frac{2\ x\ 10}{50 + 50}\ x\ 100$$

$$= \frac{20}{100}\ x\ 100$$

$$= 20\%$$

3.7.4 Ví dụ về tính mã băm của 02 chuỗi với hệ số K-Gram=5, hệ số B=7

Chuỗi 1= “MEDAN**CDE**MABCC”

Chuỗi 2= “MEDAN**ZXY**MABCC”

Ta có 02 bảng băm như sau:

Bảng 3.2 Bảng giá trị băm của 02 chuỗi với hệ số K-Gram=5, hệ số B=7

1	2	3	4	5	6	7	8	9
1486863	1349537	1329454	1306529	1499057	1317232	1338603	1370558	**1476594**

1	2	3	4	5	6	7	8	9
1486863	1349698	1330721	1315538	1562120	1758673	1722763	1706698	**1476594**

Mỗi bảng băm có 9 phần tử, trong đó 2 phần tử có giá trị giống nhau. Vậy, mức độ tương đồng là:

P = (2 x 2) / (9+9) x 100% = 22,2%

3.7.5 Ví dụ về tính mã băm của 02 chuỗi với hệ số K-Gram=3, hệ số B=7

Chuỗi 1= “MEDAN**CDE**MABCC”

Chuỗi 2= “MEDAN**ZXY**MABCC”

Ta có 02 bảng băm như sau

Bảng 3.3 Bảng giá trị băm của 02 chuỗi với hệ số K-Gram=3, hệ số B=7

1	2	3	4	5	6	7	8	9	10	11	12
4256	**3857**	**3787**	**3731**	4291	3759	3815	3920	**4228**	**3647**	**3703**	**3752**

1	2	3	4	5	6	7	8	9	10	11	12
4256	**3857**	**3787**	**3731**	4452	5026	4935	4900	**4228**	**3647**	**3703**	**3752**

Mỗi bảng băm có 12 phần tử, trong đó 8 phần tử có giá trị giống nhau. Vậy, mức độ tương đồng là:

P = (2 x 8) / (12+12) x 100% = 66,6%

CHƯƠNG 4: ĐỀ XUẤT BIỆN PHÁP NHẰM PHÁT HIỆN CUỘC TẤN CÔNG AN NINH MẠNG DEFACE

Một website bao gồm các thành phần chính như là Domain (tên miền), Host, Source (mã nguồn) và Database (cơ sở dữ liệu). Tương ứng với mỗi thành phần là biện pháp tấn công khác nhau.

4.1 Biện pháp giám sát việc thay đổi nội dung của website

Như tên gọi của cuộc tấn công này, mục tiêu của tin tặc nhằm thay đổi nội dung trình bày tại trang chủ của website để thông báo hệ thống đã bị xâm nhập. Đây là cơ chế giám sát website từ xa, hướng từ bên ngoài:

Bước 1 - Khai báo thông tin của website.

Bước 2 - Lấy tài liệu CSS của website.

Bước 3 - So sánh nội dung của website với nội dung đã lưu trước đó.

4.2 Biện pháp giám sát tình trạng hoạt động của website

Để biết tình trạng hoạt động hay không hoạt động của một Website, ta dựa vào ba tham số để kiểm tra là DNS, HTTP Request, và IP:

- Bước 1 - Kiểm tra DNS.

- Bước 2 - Kiểm tra HTTP Request.

- Bước 3 - Kiểm tra IP Public.

4.3 Biện pháp phát hiện sự thay đổi tính toàn vẹn

Biện pháp này nhằm phát hiện sự thay đổi mã nguồn, giám sát thư mục hoặc giám sát tập tin của website: Một Website muốn hoạt động tốt và an toàn thì cần phải bảo vệ thư mục chứa mã nguồn của website trên máy chủ web.

4.4 Biện pháp phát hiện cuộc tấn công làm tê liệt website

Để phát hiện trường hợp kẻ tấn công phá hủy chương trình giám sát hay bị tấn công từ chối dịch vụ làm tê liệt máy chủ web. Học viên đề xuất sử dụng mô hình Agent – Controller. Agent chính là một ứng dụng được cài đặt trên các máy chủ. Controller là trung tâm giám sát được cài đặt trên một máy độc lập với máy chủ web, Controller này sẽ tiếp

nhận các thông tin từ Agent gửi về. Với mô hình này giữa Agent và Controller sẽ duy trì kênh liên lạc riêng sử dụng Advanced Encryption Standard (AES) để mã hóa các thông điệp trao đổi và định kỳ Controller sẽ gửi thông tin liên lạc nếu mất thông tin liên lạc trong một khoảng thời gian nhất định thì sẽ cảnh báo.

CHƯƠNG 5: XÂY DỰNG HỆ THỐNG GIÁM SÁT VÀ CẢNH BÁO CUỘC TẤN CÔNG AN NINH MẠNG DEFACE

5.1 Các yêu cầu đối với hệ thống đề xuất

Như đã tìm hiểu nêu trên, hậu quả của cuộc tấn công an ninh mạng deface gây ra hậu quả rất nghiêm trọng. Học viên đề xuất xây dựng một hệ thống có khả năng giám sát và cảnh báo khi có cuộc tấn công an ninh mạng deface với các yêu cầu như sau:

- Kịp thời phát hiện và cảnh báo khi có sự thay đổi bất thường của nội dung của website: Có thể kiểm tra sự thay đổi toàn bộ hoặc một phần của website.

- Kịp thời phát hiện và cảnh báo: Khi có sự thay đổi tính toàn vẹn; khi website không hoạt động và khi website phá hủy hoặc bị làm tê liệt.

- Giám sát liên tục hoặc giám sát trong một khung giờ cố định.

- Có thể giám sát nhiều website cùng lúc.

- Có lưu trữ kết quả giám sát để phục vụ việc tra cứu, đánh giá tình hình.

- Thực hiện cảnh báo đến máy tính, gửi tin nhắn và thư điện tử được chỉ định.

- Hệ thống hoạt động ổn định, tin cậy.

- Có giao diện đơn giản, thân thiện.

5.2 Mô tả hệ thống được đề xuất

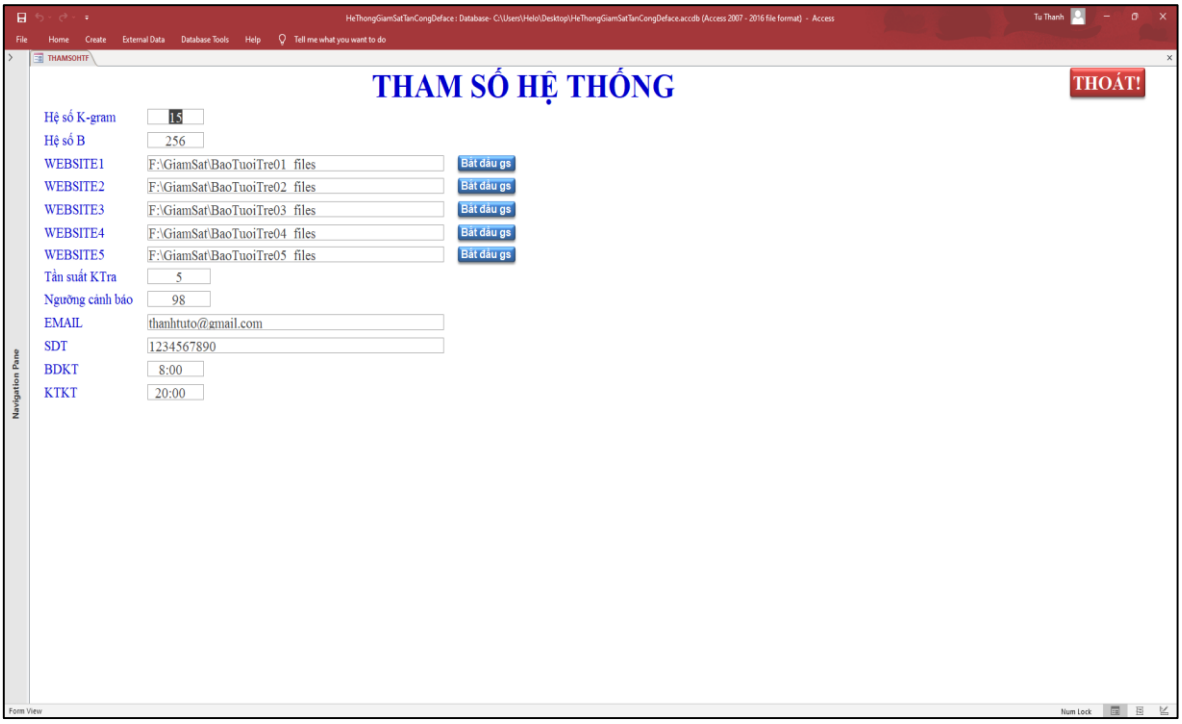
Hệ thống đề xuất sẽ gồm các giao diện như sau:

- Giao diện Chính: Đây là giao diện đầu tiên hiển thị và cho phép người sử dụng đi đến các giao diện khác của hệ thống.

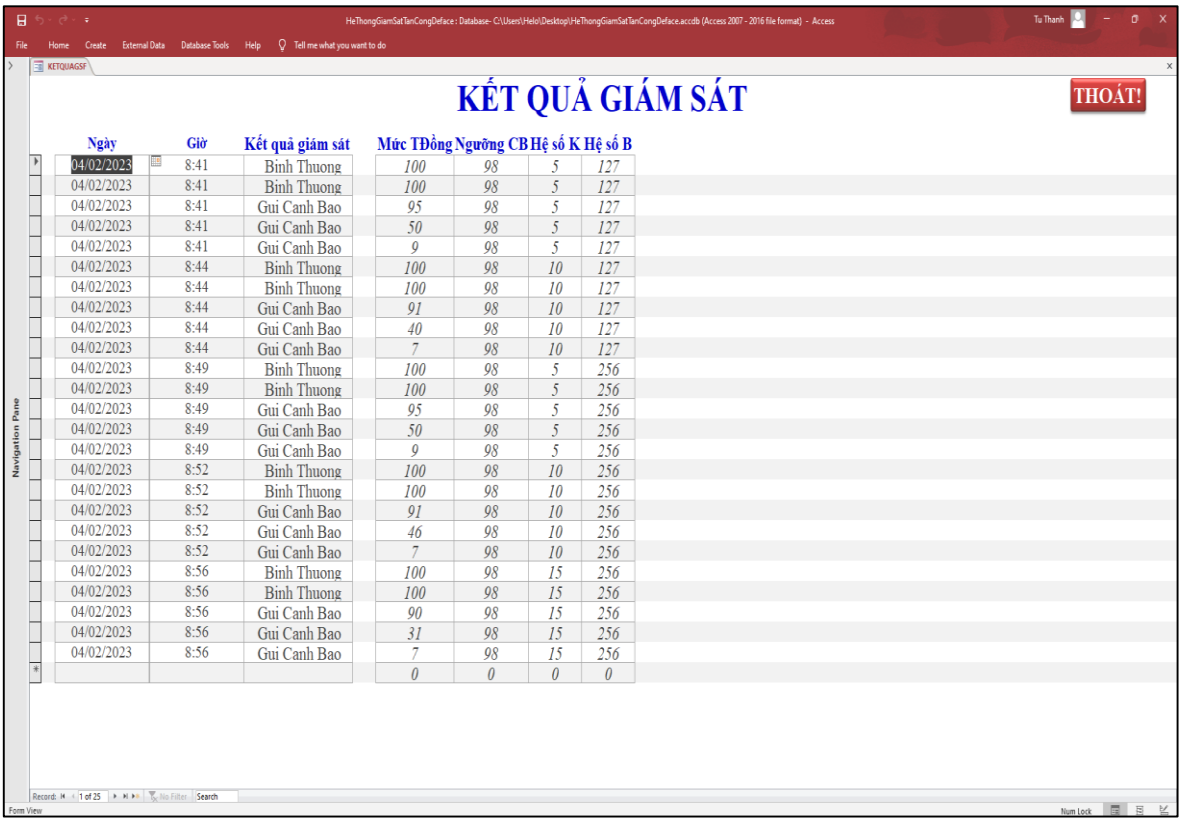
- Giao diện Tham số hệ thống: Giao diện này cho phép người sử dụng thiết lập các tham số của hệ thống như tần suất giám sát, số điện thoại, email để gửi cảnh báo...

- Giao diện Kết quả giám sát: Giao diện này cung cấp thông tin kết quả của việc giám sát theo thời gian của hệ thống.

Hình 5.1 Giao diện chính của hệ thống



Hình 5.2 Giao diện tham số của hệ thống



Ngày	Giờ	Kết quả giám sát	Mức TĐồng	Ngưỡng CB	Hệ số K	Hệ số B
04/02/2023	8:41	Binh Thuong	100	98	5	127
04/02/2023	8:41	Binh Thuong	100	98	5	127
04/02/2023	8:41	Gui Canh Bao	95	98	5	127
04/02/2023	8:41	Gui Canh Bao	50	98	5	127
04/02/2023	8:41	Gui Canh Bao	9	98	5	127
04/02/2023	8:44	Binh Thuong	100	98	10	127
04/02/2023	8:44	Binh Thuong	100	98	10	127
04/02/2023	8:44	Gui Canh Bao	91	98	10	127
04/02/2023	8:44	Gui Canh Bao	40	98	10	127
04/02/2023	8:44	Gui Canh Bao	7	98	10	127
04/02/2023	8:49	Binh Thuong	100	98	5	256
04/02/2023	8:49	Binh Thuong	100	98	5	256
04/02/2023	8:49	Gui Canh Bao	95	98	5	256
04/02/2023	8:49	Gui Canh Bao	50	98	5	256
04/02/2023	8:49	Gui Canh Bao	9	98	5	256
04/02/2023	8:52	Binh Thuong	100	98	10	256
04/02/2023	8:52	Binh Thuong	100	98	10	256
04/02/2023	8:52	Gui Canh Bao	91	98	10	256
04/02/2023	8:52	Gui Canh Bao	46	98	10	256
04/02/2023	8:52	Gui Canh Bao	7	98	10	256
04/02/2023	8:56	Binh Thuong	100	98	15	256
04/02/2023	8:56	Binh Thuong	100	98	15	256
04/02/2023	8:56	Gui Canh Bao	90	98	15	256
04/02/2023	8:56	Gui Canh Bao	31	98	15	256
04/02/2023	8:56	Gui Canh Bao	7	98	15	256
			0	0	0	0

Hình 5.3 Giao diện kết quả giám sát

5.3 Xây dựng hệ thống

Nhằm thực hiện hệ thống đã đề xuất, học viên sử dụng ngôn ngữ VBA(Visual Basic For Applications) trong phần mềm Microsoft Access để hiện thực hóa thuật toán Hàm băm Rabin-Karp và việc sử dụng hàm băm để phát hiện cuộc tấn công an ninh mạng deface đã nêu trên.

Học viên giả sử hệ thống sẽ thực hiện giám sát file “font.css” trên websibe của Báo Tuổi trẻ điện tử ở 05 trạng thái được lưu tại 05 thư mục “Giamsat” tại ổ F: của máy tính như hình:

- Trạng thái 1: Đây là trạng thái đầu tiên của websibe của Báo Tuổi trẻ.

- Trạng thái 2: Giả sử trạng thái của Báo Tuổi trẻ vẫn không đổi (file “font.css” không đổi nội dung so với trạng thái 1).

- Trạng thái 3: Giả sử trạng thái của Báo Tuổi trẻ có thay đổi do file “font.css” thay đổi nội dung là các kiểu định dạng từ “font-style: **normal**;” chuyển thành “font-style: **italic**;”

- Trạng thái 4: Giả sử trạng thái của Báo Tuổi trẻ có thay đổi do file “font.css” do nội dung bị xóa một phần.

- Trạng thái 5: Giả sử trạng thái của Báo Tuổi trẻ có thay đổi do file “font.css” do nội dung là các phần định dạng chữ font-face bị xóa hết.

5.3.1 Hàm tính giá trị băm của chuỗi ký tự

```
Public Function TINHGIATRIBAMCHUOIKYTU(chuoi As String, b As Integer) As Long
'Tinh_gia_tri_bam_chuoi_ky_tu
'Hàm tính giá tri bam cua chuoi ky tu
'b là he so co so (base)=256
'songuyento = 997

    Dim i As Integer
    Dim sokytu As Integer
    Dim p As Long
    Dim songuyento As Integer

    p = 0
    songuyento = 997
    sokytu = Len(chuoi)

    For i = 1 To sokytu
        p = (b * p + Asc(Mid(chuoi, i, 1))) Mod songuyento
    Next i

    TINHGIATRIBAMCHUOIKYTU = p
End Function
```

5.3.2 Hàm tính bảng băm của một file text

```
Public Function TINHBANGBAMFILETEXT(FText As String, BANGGTB As Integer)
'Tinh_bang_bam_file_text
'Hàm tính cac gia tri bam cua file text
'HSOK (K-gram) là so ky tu cua chuoi muon tinh
'BANGGTB:bang gia tri bam = 1 hoac 2
'-------------------------------------
    Dim db As DAO.Database
    Dim rs1 As DAO.Recordset
    Dim strSQL As String

    Dim i As Integer
    Dim sokytu As Integer
    Dim p As Long
    Dim chuoiK As String
    Dim sochuoi As Integer
    Dim HSOB As Integer
    Dim HSOK As Integer

'---------------------------------
'1.Lay gia tri hsoB
    Set db = CurrentDb()
    strSQL = " SELECT THAMSOHT.IDTHS, THAMSOHT.HSOK, THAMSOHT.HSOB, THAMSOHT.WEBSITE1, THAMSOHT.WEBSITE2 " & _
             " FROM THAMSOHT " & _
             " WHERE (((THAMSOHT.IDTHS)=1)); "
    Set rs1 = db.OpenRecordset(strSQL, dbOpenDynaset)

    HSOB = rs1!HSOB
    HSOK = rs1!HSOK
    rs1.Close
'---------------------------------
'2.Xoa cac record trong Bang gia tri bam 1 hoac 2
    If (BANGGTB = 1) Then
        strSQL = " Delete BANGGTB1.GTB " & _
                 " FROM BANGGTB1;"
    Else
        strSQL = " Delete BANGGTB2.GTB " & _
                 " FROM BANGGTB2;"
    End If
    DoCmd.SetWarnings False
    DoCmd.RunSQL strSQL
```

```
'---------------------------------
'3.Them moi cac record trong Bang gia tri bam 1 hoac 2
    If (BANGGTB = 1) Then
        strSQL = " SELECT BANGGTB1.GTB " & _
                 " FROM BANGGTB1;"
    Else
        strSQL = " SELECT BANGGTB2.GTB " & _
                 " FROM BANGGTB2;"
    End If
    Set rs1 = db.OpenRecordset(strSQL, dbOpenDynaset)
'---------------------------------
'4.Tinh cac gia tri bam va luu vao bang
    p = 0
    sokytu = Len(FText)
    sochuoi = ((sokytu - (sokytu Mod HSOK)) / HSOK) 'tinh so chuoi cua file

    For i = 1 To sochuoi
        chuoiK = Mid(FText, (i - 1) * HSOK + 1, HSOK)
        p = TINHGIATRIBAMCHUOIKYTU(chuoiK, HSOB)

        rs1.AddNew
        rs1!GTB = p
        rs1.Update
    Next i

rs1.Close
db.Close
End Function
```

5.3.3 Hàm tính mức độ tương đồng của hai tài liệu

```
Public Function TinhSuTuongDong02BangGTB() As Integer
'Tinh_su_tuong_02_bang_Gia_Tri_Bam
    Dim db As DAO.Database
    Dim rs1, rs2, rs3 As DAO.Recordset
    Dim strSQL1, strSQL2, strSQL3 As String
    Dim SH As Integer       'SH (Identical Hash): so giá tri bam giong nhau.
    Dim TyleTD As Integer   'ty le tuong dong
    Dim a, b As Integer
    Dim sophantuBangGTB1, sophantuBangGTB2 As Integer

    SH = 0
    TyleTD = 0
    sophantuBangGTB1 = 0
    sophantuBangGTB2 = 0

    strSQL1 = " SELECT BANGGTB1.GTB " & _
              " FROM BANGGTB1 " & _
              " GROUP BY BANGGTB1.GTB; "

    strSQL2 = " SELECT BANGGTB2.GTB " & _
              " FROM BANGGTB2 " & _
              " GROUP BY BANGGTB2.GTB; "

    Set db = CurrentDb()
    Set rs1 = db.OpenRecordset(strSQL1, dbOpenDynaset)
    Set rs2 = db.OpenRecordset(strSQL2, dbOpenDynaset)
```

```
'---------------------------
'1. Kiem tra bang bam 1 va bang bam 2 co phan tu nao ko
    If (rs1.BOF = True) And (rs1.EOF = True) Then
        MsgBox "BANG GIA TRI BAM 1 KHONG CO GIA TRI NAO!"
        TinhSuTuongDong02BangGTB = TyleTD
        Exit Function
    End If
    If (rs2.BOF = True) And (rs2.EOF = True) Then
        MsgBox "BANG GIA TRI BAM 2 KHONG CO GIA TRI NAO!"
        TinhSuTuongDong02BangGTB = TyleTD
        Exit Function
    End If
rs1.Close
rs2.Close
'---------------------------
'2. Lay tong phan tu cua tung bang
    strSQL1 = " SELECT Count(BANGGTB1.GTB) AS CountOfGTB " & _
              " FROM BANGGTB1; "
    strSQL2 = " SELECT Count(BANGGTB2.GTB) AS CountOfGTB " & _
              " FROM BANGGTB2; "

    Set rs1 = db.OpenRecordset(strSQL1, dbOpenDynaset)
    Set rs2 = db.OpenRecordset(strSQL2, dbOpenDynaset)

    sophantuBangGTB1 = rs1!CountOfGTB
    sophantuBangGTB2 = rs2!CountOfGTB
rs1.Close
rs2.Close
'---------------------------
    strSQL1 = " SELECT BANGGTB1.GTB " & _
              " FROM BANGGTB1 " & _
              " GROUP BY BANGGTB1.GTB; "

    strSQL2 = " SELECT BANGGTB2.GTB " & _
              " FROM BANGGTB2 " & _
              " GROUP BY BANGGTB2.GTB; "

    Set rs1 = db.OpenRecordset(strSQL1, dbOpenDynaset)
    Set rs2 = db.OpenRecordset(strSQL2, dbOpenDynaset)
```

```
'3. Xet tung gia tri bam cua bang bam 1 voi tat ca gia tri bam cua bang bam 2
    rs1.MoveFirst
    Do Until rs1.EOF
        rs2.MoveFirst
        Do Until rs2.EOF
            If rs1!GTB = rs2!GTB Then
                strSQL3 = " SELECT BANGGTB1.GTB, Count(BANGGTB1.GTB) AS CountOfGTB " & _
                          " FROM BANGGTB1 " & _
                          " GROUP BY BANGGTB1.GTB " & _
                          " HAVING (((BANGGTB1.GTB)= " & rs1!GTB & "  )); "

                Set rs3 = db.OpenRecordset(strSQL3, dbOpenDynaset)
                If (rs3.BOF = True) And (rs3.EOF = True) Then
                    a = 0
                Else
                    a = rs3!CountOfGTB
                End If
                rs3.Close

                strSQL3 = " SELECT BANGGTB2.GTB, Count(BANGGTB2.GTB) AS CountOfGTB " & _
                          " FROM BANGGTB2 " & _
                          " GROUP BY BANGGTB2.GTB " & _
                          " HAVING (((BANGGTB2.GTB)= " & rs2!GTB & "  )); "

                Set rs3 = db.OpenRecordset(strSQL3, dbOpenDynaset)

                If (rs3.BOF = True) And (rs3.EOF = True) Then
                    b = 0
                Else
                    b = rs3!CountOfGTB
                End If
                rs3.Close

                If (a < b) Then
                    SH = SH + a
                Else
                    SH = SH + b
                End If
            End If
        rs2.MoveNext
        Loop
    rs1.MoveNext
    Loop
rs1.Close
rs2.Close
db.Close
'---------------------------
'3. Tinh ty le tuong dong
    TyleTD = (2 * SH) / (sophantuBangGTB1 + sophantuBangGTB2) * 100
    TinhSuTuongDong02BangGTB = TyleTD
'    MsgBox SH & "--" & sophantuBangGTB1 & "--" & sophantuBangGTB2 & "--" & TyleTD
End Function
```

5.4 Kết quả thực nghiệm hệ thống và nhận xét

Để thực nghiệm hệ thống đã xây dựng, học viên đã giám sát 05 trạng thái giả thuyết với các tham số của thuật toán Rabin-Karp như sau:

- K=5 và B=127.

- K=10 và B=127.

- K=5 và B=256.

- K=10 và B=256.

- K=15 và B=256.

Bảng 4.1 Bảng kết quả thực nghiệm

NGAY	GIO	KQGS	HESOK	HESOB	NGUONGCB	P
04/02/2023	8:41	Binh Thuong	5	127	98	100
04/02/2023	8:41	Binh Thuong	5	127	98	100
04/02/2023	8:41	Gui Canh Bao	5	127	98	95
04/02/2023	8:41	Gui Canh Bao	5	127	98	50
04/02/2023	8:41	Gui Canh Bao	5	127	98	9
04/02/2023	8:44	Binh Thuong	10	127	98	100
04/02/2023	8:44	Binh Thuong	10	127	98	100
04/02/2023	8:44	Gui Canh Bao	10	127	98	91
04/02/2023	8:44	Gui Canh Bao	10	127	98	40
04/02/2023	8:44	Gui Canh Bao	10	127	98	7
04/02/2023	8:49	Binh Thuong	5	256	98	100
04/02/2023	8:49	Binh Thuong	5	256	98	100
04/02/2023	8:49	Gui Canh Bao	5	256	98	95
04/02/2023	8:49	Gui Canh Bao	5	256	98	50
04/02/2023	8:49	Gui Canh Bao	5	256	98	9
04/02/2023	8:52	Binh Thuong	10	256	98	100
04/02/2023	8:52	Binh Thuong	10	256	98	100
04/02/2023	8:52	Gui Canh Bao	10	256	98	91
04/02/2023	8:52	Gui Canh Bao	10	256	98	46
04/02/2023	8:52	Gui Canh Bao	10	256	98	7
04/02/2023	8:56	Binh Thuong	15	256	98	100
04/02/2023	8:56	Binh Thuong	15	256	98	100
04/02/2023	8:56	Gui Canh Bao	15	256	98	90
04/02/2023	8:56	Gui Canh Bao	15	256	98	31
04/02/2023	8:56	Gui Canh Bao	15	256	98	7

Dựa vào kết quả thực nghiệm trên, ta nhận thấy:

- Khi hai trạng thái giám sát của website không đổi thì hệ thống đề xuất sẽ không đưa ra cảnh báo: Mức tương đồng của 02 trạng thái =100%.

- Khi trạng thái giám sát của website có thay đổi thì hệ thống có phát hiện sự thay đổi: Mức tương đồng của hai trạng thái <100%.

- Khi cố định hệ số cơ sở B của thuật toán Rabin-Karp thì mức tương đồng của hai trạng thái sẽ giảm khi K càng lớn.

CHƯƠNG 6: KẾT LUẬN

1. Kết quả nghiên cứu của đề tài

Tấn công an ninh mạng deface là một kiểu tấn công phổ biến và gây ra nhiều hậu quả nghiêm trọng đối với bất kỳ tổ chức, cá nhân nhất là đối với các doanh nghiệp. Do vậy việc kịp thời phát hiện website bị tấn công để đưa ra các biện pháp phòng chống là cần thiết để hạn chế tối đa các hậu quả không mong muốn.

- Thông qua đề tài, học viên hiểu được cách mà trình duyệt web trình bày một trang web; cách thức một website hoạt động và các vấn đề về an ninh mạng, những lỗ hổng bảo mật, những kỹ thuật tấn công để xâm nhập, khống chế website và cũng như cách phòng chống.

- Hiểu được thuật toán Rabin-Karp và ứng dụng của thuật toán nhằm phát hiện cuộc tấn công an ninh mạng deface.

- Giải quyết một phần các yêu cầu đặt ra.

2. Hạn chế của luận văn

Tuy nhiên do giới hạn về thời gian và có nhiều kiến thức cần phải nghiên cứu nên luận văn còn tồn tại các hạn chế sau:

- Chưa xây dựng được hệ thống giám sát đối với các website trên môi trường thực tế mà chỉ dừng ở môi trường giả thuyết.

- Chưa xây dựng được hệ thống giám sát đặt ngoài website cần giám sát để đánh giá website có còn hoạt động hay không.

3. Kiến nghị và hướng nghiên cứu tiếp theo

Đề tài cần được nghiên cứu, cải tiến trong thời gian tới như sau:

- Xây dựng hệ thống giám sát đối với các website trên môi trường thực tế và hệ thống giám sát đặt ngoài website.

- Xây dựng nhiều kích bản về cuộc tấn công an ninh mạng deface để xác định các tham số hệ thống nhằm làm cho hệ thống hoạt động hiệu quả, đáng tin cậy.

- Nghiên cứu, ứng dụng các kỹ thuật khác nhằm tăng cường khả năng phát hiện cuộc tấn công an mạng deface.