(12) **BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ**

(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ

(11) **1-0020428**

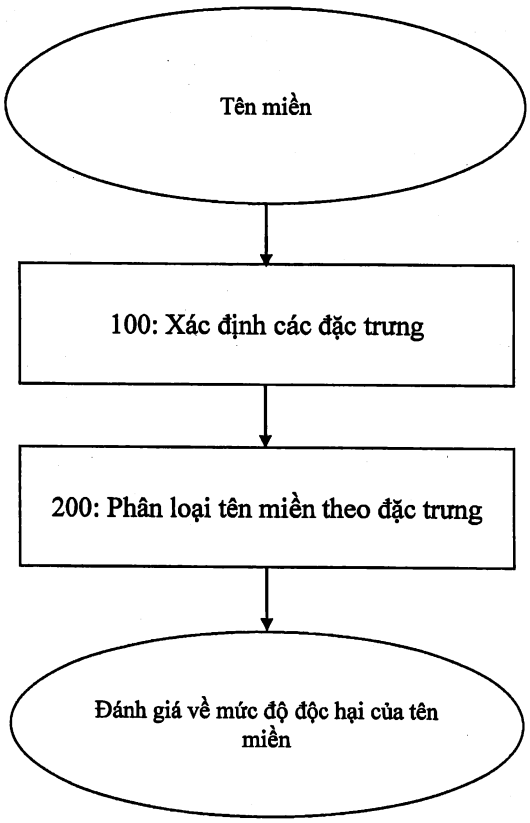
(51)[7] **H04L 29/06**

(13) **B**

(21) 1-2015-04495
(22) 24.11.2015
(45) 25.02.2019 371
(43) 25.02.2016 335
(73) VIỆN NGHIÊN CỨU CÔNG NGHỆ FPT - TRƯỜNG ĐẠI HỌC FPT (VN)
Số 8 Tôn Thất Thuyết, Mỹ Đình 2, Nam Từ Liêm, thành phố Hà Nội
(72) Vũ Công Thành (VN), Lê Hồng Phương (VN), Nguyễn Minh Đức (VN)

(54) QUY TRÌNH XÁC ĐỊNH TÊN MIỀN ĐỘC HẠI

(57) Sáng chế đề xuất quy trình nhận đầu vào là tên miền và đưa ra đánh giá xem tên miền đó bình thường hay là có liên quan đến các phần mềm độc hại. Quy trình này xác định bốn con số đặc trưng của tên miền, gồm độ dài tên miền, entropy của tên miền, và hai tích vô hướng của véc tơ n-gram của tên miền với hai véc tơ n-gram tham chiếu. Sau khi có được bốn con số đặc trưng cho tên miền, các con số này được đưa vào phân loại bằng kỹ thuật học máy thống kê có giám sát như phương pháp phân loại Rừng Ngẫu nhiên, để phân loại nó thuộc về loại bình thường hay là có liên hệ với phần mềm độc hại.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến quy trình để từ đầu vào là tên miền, đưa ra đánh giá rằng tên miền này có liên hệ với các phần mềm độc hại hay không.

Tình trạng kỹ thuật của sáng chế

Việc xác định liệu máy tính có chứa các phần mềm độc hại hay không là một trong các hoạt động quan trọng trong đảm bảo an ninh an toàn cho các hệ thống máy tính và hệ thống thông tin. Các phần mềm độc hại rất đa dạng và luôn biến đổi theo chiều hướng hoạt động ngày càng tinh vi để thoát khỏi sự phát hiện của các hệ thống an ninh an toàn thông tin, và có một khi không bị phát hiện thì có thể gây ra nhiều loại thiệt hại khác nhau cho những người sở hữu hệ thống thông tin hay dữ liệu trên đó.

Các phương pháp truyền thống để nhận diện phần mềm độc hại thường là xác định các đặc điểm của các phần mềm này, sau khi các phần mềm này đã phát tác, rồi đưa các đặc điểm vào một bảng tra. Mỗi khi có một phần mềm mới cần kiểm tra, các hệ thống an ninh an toàn thông tin truyền thống so sánh các đặc điểm của phần mềm mới với các đặc điểm đã có trong bảng tra, để xác định xem phần mềm mới có độc hại hay không. Ví dụ, có thể đơn giản là đưa mã máy quan trọng của phần mềm độc hại đã biết qua một hàm băm, ví dụ hàm MD5, và lưu kết quả trong một cơ sở dữ liệu. Khi có phần mềm mới cần kiểm tra, có thể đưa mã máy của phần mềm mới qua cùng hàm băm, và so sánh kết quả với cơ sở dữ liệu, nếu có sự trùng lặp thì đưa ra cảnh báo phần mềm mới là độc hại. Hoặc một cách áp dụng phức tạp hơn là ghi nhận các đặc điểm trong hành vi, chứ không phải trong mã máy, của phần mềm độc hại, ví dụ như hành vi truy cập vào những khu vực lưu trữ trọng yếu trên máy tính, chụp lại màn hình, ghi lại nhật trình gõ phím, vân vân, và lưu trữ các hành vi này vào trong cơ sở dữ liệu. Khi có phần mềm mới cần kiểm tra, xem xét lại các hành vi của phần mềm mới, và so sánh chúng với cơ sở dữ liệu hành vi độc hại đã biết, để đưa ra cảnh báo.

Các phương pháp trên có hạn chế là chỉ sau khi phần mềm độc hại đã phát tác ở một số máy tính nào đó rồi, thì bảng tra mới có được thông tin về đặc điểm của phần mềm độc hại đó; nghĩa là phương pháp này không phát hiện và ngăn chặn được những phần mềm độc hại có đặc điểm mới, chưa biết.

Hầu hết các phần mềm độc hại hiện đại đều có liên lạc thông tin đến những máy tính ở trên mạng Internet. Cách hoạt động phổ biến của chúng là, bằng một cách nào đó, có mặt trên máy tính nạn nhân, nhưng chưa có bất cứ hoạt động gì, để chờ thời cơ trong một thời gian dài. Khi có thông tin điều khiển từ một máy tính nhất định trên mạng Internet gửi tới chúng, chúng mới phát tác. Chẳng hạn, phần mềm độc hại có thể ở dạng các '*bot net*' tấn công từ chối dịch vụ vào một dịch vụ Web nào đó. Chúng có thể nằm rải rác trong rất nhiều máy tính ở nhiều nơi khác nhau trên mạng Internet, và không phát tác trong môt thời gian dài. Đến khi có lệnh điều khiển gửi tới chúng từ một máy tính nhất định của kẻ tấn công, chúng đồng loạt truy cập vào dịch vụ Web nạn nhân, khiến cho dịch vụ Web nạn nhân bị quá tải.

Như vậy có thể phát hiện phần mềm độc hại, ngay cả khi các đặc điểm của chúng chưa được biết đến, nếu chúng có kết nối internet đến tên miền của những máy tính điều khiển chúng. Các tên miền này có thể có những đặc trưng có thể giúp nhận biết sự nguy hiểm của chúng. Nếu có thể tự động đánh giá mức độ nguy hiểm của một tên miền mà một phần mềm kết nối đến, thì có thể tự động cảnh báo sự độc hại của phần mềm đó.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất quy trình nhận đầu vào là một tên miền và đưa ra đánh giá rằng tên miền này có liên hệ với các phần mềm độc hại hay không.

Cụ thể, quy trình nhận đầu vào là tên miền và đưa ra đánh giá xem tên miền đó bình thường hay là có liên quan đến các phần mềm độc hại, gồm các bước:

bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.' khỏi tên miền, để thu được tên miền rút gọn;

xác định bốn con số:

số thứ nhất là độ dài của tên miền rút gọn, bằng với số ký tự có trong tên miền rút gọn;

số thứ hai là entropy của tên miền rút gọn, bằng với:

$$-\sum_i P(x_i) \ln P(x_i)$$

trong đó x_i là ký tự thứ i trong tên miền rút gọn, $i = 1, 2, \ldots$, $P(x_i)$ là thương số của số lần xuất hiện của ký tự bằng với ký tự x_i trong tên miền rút gọn chia cho độ dài của tên miền rút gọn;

số thứ ba là tích vô hướng của véc tơ *n*-gram của tên miền rút gọn với véc tơ *n*-gram tham chiếu thứ nhất, và số thứ tư là tích vô hướng của véc tơ *n*-gram của tên miền rút gọn với véc tơ *n*-gram tham chiếu thứ hai, trong đó *n* là một số nguyên dương chọn trước và:

véc tơ *n*-gram của tên miền rút gọn là véc tơ gồm có các thành phần được tính bằng số lần xuất hiện của một chuỗi độ dài *n* ký tự nhất định - chuỗi *n* ký tự này tương ứng với thành phần đang xét - xuất hiện trong tên miền rút gọn;

véc tơ *n*-gram tham chiếu thứ nhất là một véc tơ có độ dài bằng với véc tơ *n*-gram của tên miền rút gọn, và có các thành phần được đặt bằng những giá trị cho trước;

véc tơ *n*-gram tham chiếu thứ hai là một véc tơ có độ dài bằng với véc tơ *n*-gram của tên miền rút gọn, và có các thành phần được đặt bằng những giá trị cho trước.

đưa bốn con số trên vào phân loại bằng kỹ thuật học máy thống kê có giám sát như phương pháp phân loại Rừng Ngẫu nhiên, để phân loại nó thuộc về loại bình thường hay là có liên hệ với phần mềm độc hại, sử dụng một mô hình Rừng Ngẫu nhiên được xây dựng sẵn, từ tập hợp bốn con số

đặc trưng tính từ những tên miền rút gọn đã biết chính xác là bình thường hay là có liên hệ với phần mềm độc hại.

Mô tả vắn tắt các hình vẽ

Hình 1 là sơ đồ khối của các bước thực hiện quy trình đánh giá một tên miền có liên hệ với các phần mềm độc hại hay không.

Mô tả chi tiết sáng chế

Các quan sát về các tên miền, và so sánh các tên miền bình thường, có trong danh bạ các tên miền nổi tiếng, như danh bạ của dịch vụ alexa.com, với các tên miền mà các phần mềm độc hại đã biết thường liên lạc đến, được thống kê bởi các dịch vụ hỗ trợ an toàn thông tin trên thế giới, như malwaredomains.com, cho thấy có sự khác nhau nhất định giữa tên miền bình thường và tên miền độc hại.

Chẳng hạn, tần suất sử dụng các chữ cái, trong bảng chữ cái tiếng Anh (từ 'a' đến 'z'), và các chữ số (từ 0 đến 9) là khác nhau giữa hai loại tên miền:

tên miền bình thường thường ít dùng các chữ số hơn tên miền độc hại;

tần suất xuất hiện các chữ cái trong tên miền độc hại chia đều hơn cho các chữ cái khác nhau; đặc biệt các chữ cái 'c', 'd', 'g', 'h', 'l', 'u' xuất hiện nhiều hơn trong các tên miền độc hại, khi so sánh với các tên miền thường.

Như vậy có thể xác định những đặc trưng trong các tên miền để từ đó phân loại nó thành loại độc hại hoặc bình thường.

Hình 1 là sơ đồ khối của các bước thực hiện quy trình đánh giá một tên miền có liên hệ với các phần mềm độc hại hay không. Quy trình gồm hai bước:

bước 100, nhận đầu vào là tên miền và đưa ra các đặc trưng của tên miền;

bước 200, nhận đầu vào là các đặc trưng của tên miền và đưa ra phân loại tên miền thuộc vào loại độc hại hay không.

Ở bước 100, có ba đặc trưng được xác định:

thứ nhất, độ dài của tên miền;

thứ hai, entropy của tên miền - được định nghĩa ở bên dưới đây;

thứ ba, tích vô hướng véc tơ *n*-gram của tên miền với véc tơ *n*-gram tham chiếu thứ nhất, và tích vô hướng véc tơ *n*-gram của tên miền với véc tơ *n*-gram tham chiếu thứ hai; với các véc tơ *n*-gram được định nghĩa ở bên dưới đây và *n* có thể chọn bằng một số nguyên dương nhất định ví dụ $n = 3$.

Độ dài của tên miền bằng với số ký tự trong tên miền, sau khi đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' (ví dụ '.net', '.com' ...) và bỏ đi các ký tự '.'; ví dụ tên miền 'vnexpress.net' sẽ bị lược bỏ còn 'vnexpress' và có 9 ký tự. Tên miền sau khi đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.' được gọi là tên miền rút gọn hoặc tên miền bị lược bỏ.

Với một tên miền *X*, sau khi đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.', còn lại một chuỗi *N* ký tự $x_1x_2...x_N$ liền nhau, thì entropy, ký hiệu là *H*, của nó là:

$$H(X) = -\sum_i P(x_i) \ln P(x_i)$$

với $P(x_i)$ là thương số của số lần xuất hiện của ký tự bằng với x_i trong chuỗi $x_1x_2...x_N$ chia cho N, và $i = 1, 2, \ldots N$. Ví dụ, với tên miền 'sub.example.com', sau khi bị lược bỏ thì còn lại 'subexample' và $P(\text{'s'}) = P(\text{'u'}) = P(\text{'b'}) = P(\text{'x'}) = P(\text{'a'}) = P(\text{'m'}) = P(\text{'p'}) = P(\text{'l'}) = 1/10$ và $P(\text{'e'}) = 2/10$. Khi đó entropy của 'subexample' là 2,394.

Véc tơ *n*-gram của một tên miền là véc tơ gồm có các thành phần được tính bằng số lần xuất hiện của một chuỗi độ dài *n* ký tự nhất định - chuỗi *n* ký tự này tương ứng với thành phần đang xét - xuất hiện trong tên miền đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.'. Ví dụ nếu *n* bằng 1, có cả thảy 36 chuỗi có độ dài 1 ký tự (26 chữ cái tiếng Anh, và 10 chữ số):

['a', 'b', 'c', 'd', 'e', 'f', 'g', ... , '7', '8', '9']

Khi đó, lấy ví dụ, để tính véc tơ 1-gram của 'sub.example.com', trước hết bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.' để thu được 'subexample'. Sau đó với từng chuỗi 1-gram, đếm số lần xuất hiện của chuỗi này trong 'subexample', để thu được véc tơ 1-gram.

{1, 1, 0, 0 2, 0, 0, …, 0, 0, 0}

Tương tự, nếu $n = 2$, có 36^2 chuỗi có độ dài 2 ký tự:

['aa', 'ab', 'ac', … , 'a8', 'a9', 'ba', 'bb', … '99']

Tổng quát có 36^n chuỗi có độ dài n ký tự.

Véc tơ n-gram tham chiếu thứ nhất là véc tơ gồm có các thành phần được tính bằng tổng số lần xuất hiện của một chuỗi độ dài n ký tự nhất định - chuỗi n ký tự này tương ứng với thành phần đang xét - xuất hiện trong tất cả các tên miền, đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.', nằm trong một tập hợp tên miền bình thường. Tập hợp tên miền bình thường, ví dụ, có thể lấy 500 tên miền phổ biến nhất ở một số quốc gia quan trọng (Mỹ, Trung Quốc, Nga, Italy, Việt Nam, Ấn Độ ...) được thống kê trên trang web alexa.com.

Véc tơ n-gram tham chiếu thứ hai là véc tơ gồm có các thành phần được tính bằng tổng số lần xuất hiện của một chuỗi độ dài n ký tự nhất định - chuỗi n ký tự này tương ứng với thành phần đang xét - xuất hiện trong tất cả các tên miền, đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.', nằm trong một tập hợp tên miền có liên quan đến các phần mềm độc hại. Tập hợp tên miền có liên quan đến phần mềm độc hại, ví dụ, có thể lấy trên 9000 tên miền được thống kê trên trang web malwaredomains.com.

Các đặc trưng thu được ở bước 100, là 4 con số gồm độ dài, entropy, và hai tích vô hướng, được cho thành đầu vào ở bước 200. Trong bước 200, 4 con số đầu vào được phân loại, vào một trong hai loại độc hại hay bình thường, bằng kỹ thuật học máy thống kê có giám sát như phương pháp phân loại Rừng Ngẫu nhiên (*Random Forest*) trong lĩnh vực học máy, như được trình bày trong bài báo *Random Decision Forests* của tác giả Ho, Tin Kam in trong kỷ yếu hội nghị

International Conference on Document Analysis and Recognition, năm 1995. Phương pháp này sử dụng một mô hình thống kê Rừng Ngẫu nhiên được xây dựng sẵn, hay còn gọi là được huấn luyện sẵn, từ tập hợp 4 con số đặc trưng tính từ những tên miền đã biết chính xác là bình thường hay độc hại, gọi là tập hợp mẫu.

Trong một phương án thực thi mở rộng của quy trình, thay vì sử dụng véc tơ *n*-gram trong cách tích vô hướng, có thể sử dụng véc tơ ghép nối từ các véc tơ sau:

véc tơ 2-gram

véc tơ 3 -gram;

...

véc tơ *n*-gram, với *n* lớn hơn 2.

Yêu cầu bảo hộ

1. Quy trình nhận đầu vào là tên miền và đưa ra đánh giá xem tên miền đó bình thường hay là có liên quan đến các phần mềm độc hại, gồm các bước:

bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.' khỏi tên miền, để thu được tên miền rút gọn;

xác định bốn con số:

số thứ nhất là độ dài của tên miền rút gọn, bằng với số ký tự có trong tên miền rút gọn;

số thứ hai là entropy của tên miền rút gọn, bằng với:

$$-\sum_i P(x_i) \ln P(x_i)$$

trong đó x_i là ký tự thứ i trong tên miền rút gọn, $i = 1, 2, \ldots$, $P(x_i)$ là thương số của số lần xuất hiện của ký tự bằng với ký tự x_i trong tên miền rút gọn chia cho độ dài của tên miền rút gọn;

số thứ ba là tích vô hướng của véc tơ n-gram của tên miền rút gọn với véc tơ n-gram tham chiếu thứ nhất, và số thứ tư là tích vô hướng của véc tơ n-gram của tên miền rút gọn với véc tơ n-gram tham chiếu thứ hai, trong đó n là một số nguyên dương chọn trước và:

véc tơ n-gram của tên miền rút gọn là véc tơ gồm có các thành phần được tính bằng số lần xuất hiện của một chuỗi độ dài n ký tự nhất định - chuỗi n ký tự này tương ứng với thành phần đang xét - xuất hiện trong tên miền rút gọn;

véc tơ n-gram tham chiếu thứ nhất là một véc tơ có độ dài bằng với véc tơ n-gram của tên miền rút gọn, và có các thành phần được đặt bằng những giá trị cho trước;

véc tơ *n*-gram tham chiếu thứ hai là một véc tơ có độ dài bằng với véc tơ *n*-gram của tên miền rút gọn, và có các thành phần được đặt bằng những giá trị cho trước;

đưa bốn con số trên vào phân loại bằng kỹ thuật học máy thống kê có giám sát như phương pháp phân loại Rừng Ngẫu nhiên, để phân loại nó thuộc về loại bình thường hay là có liên hệ với phần mềm độc hại, sử dụng một mô hình Rừng Ngẫu nhiên được xây dựng sẵn, từ tập hợp bốn con số đặc trưng tính từ những tên miền rút gọn đã biết chính xác là bình thường hay là có liên hệ với phần mềm độc hại.

2. Quy trình theo điểm 1, trong đó véc tơ *n*-gram tham chiếu thứ nhất là véc tơ gồm có các thành phần được tính bằng tổng số lần xuất hiện của một chuỗi độ dài *n* ký tự nhất định - chuỗi *n* ký tự này tương ứng với thành phần đang xét - xuất hiện trong tất cả các tên miền, đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.', nằm trong một tập hợp tên miền bình thường.
3. Quy trình theo điểm bất kỳ trong các điểm 1 và 2, trong đó véc tơ *n*-gram tham chiếu thứ hai là véc tơ gồm có các thành phần được tính bằng tổng số lần xuất hiện của một chuỗi độ dài *n* ký tự nhất định - chuỗi *n* ký tự này tương ứng với thành phần đang xét - xuất hiện trong tất cả các tên miền, đã bỏ đuôi bên phải ngoài cùng sau ký tự '.' và bỏ đi các ký tự '.', nằm trong một tập hợp tên miền có liên hệ với các phần mềm độc hại.
4. Quy trình theo điểm bất kỳ trong các điểm 1, 2 và 3, khác biệt ở chỗ, thay thế việc sử dụng các véc tơ *n*-gram trong các tích vô hướng bằng véc tơ ghép nối từ các véc tơ sau:

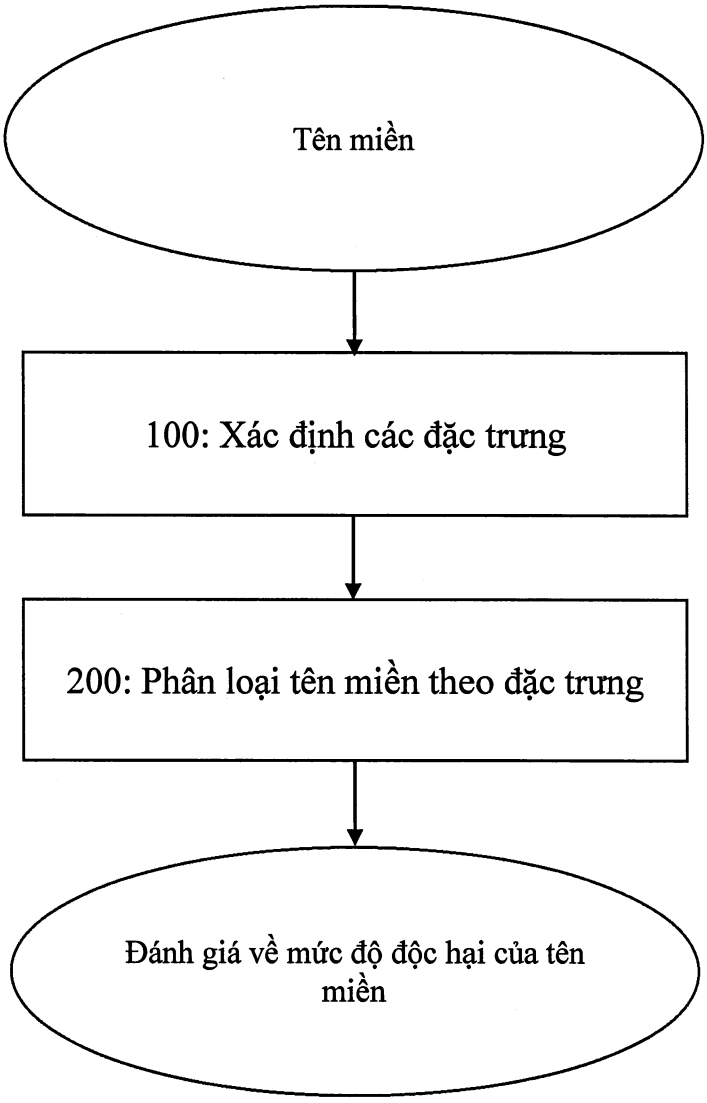
véc tơ 2-gram;

véc tơ 3 -gram;

...

véc tơ *n*-gram, với *n* lớn hơn 2.

Hình 1


Tên miền
100: Xác định các đặc trưng
200: Phân loại tên miền theo đặc trưng
Đánh giá về mức độ độc hại của tên miền