**BẢN MÔ TẢ GIẢI PHÁP DỰ THI**

**HỘI THI SÁNG TẠO KỸ THUẬT BÌNH DƯƠNG**

**LẦN VIII NĂM 2018 - 2019**

----------- oOo -----------

**A. THÔNG TIN CHUNG:**

1. Tên giải pháp: Giải pháp bảo mật IoT: “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học”

2. Thuộc lĩnh vực: Khoa học công nghệ, công nghệ thông tin.

3. Người dự thi:

Tên cơ quan:

Địa chỉ liên hệ:

Số điện thoại: , Fax: , E-mail:

**B. PHẦN TÓM TẮT NỘI DUNG GIẢI PHÁP :**

1. **Vấn đề mà giải pháp đã giải quyết :**

Hệ thống phát hiện lỗ hổng malware thiết lập trên máy chủ (server) phân tích dữ liệu gửi về bộ xử lý trung tâm (NodeMCU) từ các thiết bị cảm biến (sensor) trên nền tảng kết nối vạn vật (Internet of Thing - IoT).

Hình 1. Quy trình thực hiện của giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học”.



Các bước thực hiện của quy trình như sau:

Bước 1: Phát triển và cài đặt các chức năng giám sát (agent) trên thiết bị IoT để thu thập các dữ liệu cần thiết như: tình trạng bộ nhớ, tình trạng mạng, hiệu suất... Truyền các thông tin trên về thiết bị xử lý trung tâm.

Bước 2: Tại thiết bị trung tâm, phát triển và cài đặt các phần mềm nhằm trích xuất các đặc tính của dữ liệu.

Bước 3: Tiền xử lý các dữ liệu thô dự trên nền tảng (framework pandas) của ngôn ngữ lập trình python, xử lý các dữ liệu bị hỏng hoặc bị thiếu để đưa vào mô hình trong bước 4.

Bước 4: Dựa trên các đặc tính có sẵn từ bước 3, phát triển một mô hình dự đoán để đưa các cảnh báo có hay không một cuộc tấn công hay một loại mã độc đang lây nhiễm trên hệ thống IoT.

Sau khi thực hiện giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” giải quyết được các vấn đề như sau:

* Xây dựng ứng dụng “cảnh báo sớm các cuộc tấn công” dựa trên các ứng dụng kết nối vạn vật IoT (Internet of Things) trong việc triển khai thành phố thông minh. Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” sẽ thường xuyên thu thập, phân tích các thông tin sự kiện (gói tin, log file) và đưa ra các thông tin cảnh báo theo thời gian thực (real time) đối với các hành vi bất thường được ghi trên hệ thống.
* Đặc biệt, giải pháp sử dụng trí tuệ nhân tạo (AI - Artificial Intelligence) trong việc phân biệt các luồng dữ liệu bình thường và mã độc tự động. Cụ thể, sử dụng ML (Machine Learning) giúp hệ thống thông minh hơn qua khả năng tự học dựa trên các luồng dữ liệu đã phân tích mà không cần lập trình cụ thể.
* Ngoài ra, chức năng cảnh báo sớm là đưa ra các cảnh báo trong thời gian thực về các dữ liệu bất thường, các cuộc tấn công, các xâm hại vào hệ thống, các dấu hiệu vi phạm về chính sách an toàn thông tin trên môi trường ứng dụng IoT, để từ đó đưa ra các phương án xử lý nhanh chóng kịp thời. Hơn nữa, giải pháp còn tích hợp “trợ lý ảo” thông minh tự động trả lời các yêu cầu của người dùng về tình hình an toàn thông tin trên hệ thống IoT.

1. **Mô tả tóm tắt nội dung của giải pháp, kết quả thử nghiệm :**

* Chúng tôi đề xuất Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” trên môi trường IoT một phần nào giải quyết được bài toán bảo mật và cảnh báo sớm trên thiết bị IoT bằng cách phân tích các luồng dữ liệu gửi/nhận truy cập trên hệ thống.
* Hơn nữa giải pháp của chúng tôi tập trung vào thuật toán khai phá, phân tích mã độc dữ liệu trên giao thức TCP/IP.
* Ngoài ra giải pháp còn sử dụng cơ chế phát hiện xâm nhập IDS (Intrusion Detection System) để phát hiện các cuộc tấn công bằng cách phân tích phát hiện dựa trên sự bất thường của hệ thống. Khi được phát hiện giải pháp của chúng tôi sẽ ghi lại thông tin và gửi cảnh báo cho người dùng.

1. **Điểm mới - Điểm sáng tạo:**

* Bảo mật trong lĩnh vực IoT: Hệ thống phát hiện kiểm tra xâm nhập, phân tích lỗ hổng trên các thiết bị IoT.
* Ứng dụng trí tuệ nhân tạo: Ứng dụng thông minh có khả năng tự học bằng phương pháp gom cụm và chọn lọc các gói tin bất thường (các cuộc tấn công mới) trên môi trường IoT dựa trên thuật toán máy học (Machine Learning).
* Cảnh báo sớm: Dựa vào thuật toán Machine Learning hệ thống đưa ra cảnh báo sớm khi có dấu hiệu của cuộc tấn công xảy ra, giải pháp tự động gửi email cảnh báo đến quản trị viên thông báo có cuộc tấn công.
* Ứng dụng kết hợp với trợ lý ảo (Chatbot): Giúp người dùng, quản trị hệ thống, tương tác với trợ lý trực tuyến để truy xuất trạng thái, lịch sử hệ thống một cách nhanh chóng và ổn định.

1. **Hiệu quả kinh tế - xã hội :**

* Hiện nay, việc triển khai các giải pháp xây dựng thành phố thông minh dựa trên công nghệ 4.0 (IoT) giúp cuộc sống của chúng ta dễ dàng và tiện nghi hơn. Vì vậy giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” sẽ mang lại lợi ích to lớn, đảm bảo tính an toàn và bảo mật thông tin trong việc triển khai các giải pháp IoT.
* Chức năng cảnh báo sớm của hệ thống dựa trên trí tuệ nhân tạo, tính tự học phát hiện các cuộc tấn công, cung cấp biện pháp ngăn chặn kịp thời, làm giảm khả năng bị đánh cắp thông tin, bị chiếm quyền kiểm soát và giảm tổn thất về rủi ro kinh tế - xã hội.
* Bên cạnh đó, giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” còn cung cấp thông tin về lỗ hổng kịp thời, hỗ trợ cho việc cảnh báo và xử lý nhanh chóng.
* Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” ứng dụng trí tuệ nhân tạo giúp giảm chi phí nguồn nhân lực và thời gian để kiểm tra lại hệ thống. Tăng mức độ về an toàn thông tin trong IoT.

1. **Khả năng áp dụng :**

* Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” có thể áp dụng bảo mật IoT trong việc triển khai thành phố thông minh, nông nghiệp thông minh, y tế thông minh, dịch vụ vận tải thông minh...
* Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” cung cấp giải pháp cấp thiết, đảm bảo an toàn thông tin, hỗ trợ tính năng cảnh báo sớm cho các giải pháp, ứng dụng IoT khi tỉnh Bình Dương triển khai và áp dụng thành phố thông minh.

1. **Tình hình áp dụng, triển vọng áp dụng giải pháp:**

* Tình hình áp dụng: Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” trên môi trường IoT đang triển khai thử nghiệm đánh giá mức độ bảo mật IoT cho các giải pháp ứng dụng IoT phổ biến hiện nay.
* Triển vọng áp dụng giải pháp : Triển khai giải pháp bảo mật IoT là cần thiết trên tất cả các lĩnh vực khi ứng dụng IoT trong đời sống và xã hội.

**C. PHẦN MÔ TẢ NỘI DUNG CHÍNH CỦA GIẢI PHÁP :**

1. **Tên giải pháp:**

* Giải pháp bảo mật IoT: “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học”

“IoT security solution: "Detection of intrusion and early warning based on machine learning algorithm”

1. **Mô tả giải pháp kỹ thuật đã biết:**

* Về bản chất, các ứng dụng trên nền tảng IoT là việc kết nối nhiều thiết bị với nhiều ứng dụng theo các yêu cầu đa dạng khác nhau. Bản chất này cho thấy sự phức tạp của các ứng dụng IoT rất lớn từ các giải pháp xây dựng nhà thông minh (Smarthome), thành phố thông minh (smart city), đến chăm sóc sức khỏe (smart healthcare),... thông qua các trình điều khiển về các lĩnh vực IoT khác nhau sẽ là một thách thức lớn khi phát triển một chương trình bảo mật, phương pháp bảo mật hiệu quả được đề xuất cụ thể cho một ứng dụng sẽ có thể là không phù hợp cho một ứng dụng khác.
* Thực tế, các thiết bị cảm biến (sensors) có khắp mọi nơi, và có thể truy cập mặc định của nhà sản xuất, tự động thích ứng gửi/nhận dữ liệu đến các thiết bị khác,... khi triển khai hệ thống IoT hiện nay đa phần không đồng nhất tiêu chuẩn về giao thức, nền tảng truy cập [1], [2]. Như vậy bảo mật IoT hiệu quả không chỉ được đề xuất là bảo mật các thiết bị độc lập mà còn phải đảm bảo giải pháp bảo mật hiệu quả và tổng quát cho hệ thống (thiết bị đầu - cuối) [1a].
* Chúng tôi đề xuất Giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” trên môi trường IoT một phần nào giải quyết được bài toán bảo mật và cảnh báo sớm trên thiết bị IoT bằng cách phân tích các luồng dữ liệu gửi/nhận truy cập trên hệ thống theo các công nghệ giao thức phổ biến hiện nay như : WiFi, Bluetooth, ZigBee, 6LoWPAN, LoRaWAN, Web socket... trong phạm vi giải pháp này chúng tôi tập trung thuật toán khai phá, phân tích mã độc dữ liệu trên giao thức TCP/IP.
* Giải pháp sử dụng cơ chế phát hiện xâm nhập IDS (Intrusion Detection System) để phát hiện các cuộc tấn công bằng cách phân tích cơ chế phát hiện dựa trên sự bất thường của hệ thống hoặc dựa trên chữ ký. Khi một cuộc tấn công được phát hiện, giải pháp sẽ ghi lại thông tin hoạt động và gửi cảnh báo sớm cho người dùng.

1. **Mục đích của giải pháp dự thi:**

* Mục đích là xây dựng giải pháp “Phát hiện xâm nhập và cảnh báo sớm dựa trên thuật toán máy học” trên môi trường IoT
* Hiện nay phần mềm độc hại là nguyên nhân chính quan trọng trong hầu hết các sự cố xâm nhập và bảo mật.
* Những thách thức lớn mà các ứng dụng phân tích dữ liệu độc hại phải đối mặt hiện nay là số lượng lớn dữ liệu và dữ liệu cần được đánh giá cho mức độc hại tiềm ẩn.
* Ví dụ: các sản phẩm chống dữ liệu độc hại thời gian thực có mặt trên môi trường IoT.
* g một chức năng nhất định nào đó, gọi là một gia đình phần mềm, với cùng một dạng hành vi độc hại, liên tục được sửa đổi và xáo trộn bằng cách sử dụng các chiến thuật khác nhau, sao cho chúng trông giống nhau.
* Khi các thiết bị trong mô trường IOT bị tấn công, cần có công cụ để báo cho mọi người biết là mình đang bị tấn công và không còn được an toàn nữa.
* Hiện tại bây giờ thì việc bảo vệ, bảo mật về các thiết bị IOT chưa quan trọng, nhưng

1. **Mô tả giải pháp dự thi:**
2. **Nguyên lý của giải pháp:**

* Với sự phát triển nhanh về thiết bị Internet of Things như hiện nay và dự kiến đến năm cuối 2020 là 50 tỷ thiết bị thì thách thức cho vấn đề bảo mật cho lĩnh vực này được đặt ra yêu cầu cấp thiết cho nhiều quốc gia và toàn thế giới.
* Giải pháp triển khai thế mạnh của các kỹ thuật machine learning trong việc học các dữ liệu trong quá khứ để đưa ra các dự đoán cho các dữ liệu trong tương lai. Kỹ thuật này đã được chứng minh thành công trên các loại dữ liệu khác nhau như hình ảnh, ngôn ngữ tự nhiên. Các loại mã độc (malware) hiện nay bao gồm rất nhiều biến thể từ một tập nhỏ hơn các họ mã độc tồn tại, với một số thay đổi nhỏ để qua mặt các trình phân tích chúng. Sử dụng thuật toán máy học (machine learning) để phát hiện mã độc dựa trên các đặt tính cốt lõi của nó sẽ khắc phục được các trường hợp biến thể của mã độc.

1. **Các nội dung công nghệ chủ yếu:**

Bao gồm:

b1. Mô tả giải pháp, thuyết minh tính mới:

* Chúng ta thường phải phân tích phần mềm độc hại dựa vào tư duy của con người.
* Trong phân tích này, chúng tôi xác định thông tin nào là cần thiết để trích xuất các chỉ số. Một cách đơn giản là tìm kiếm các chuỗi trong phần mềm độc hại. Strings: Một chương trình chứa các chuỗi nếu nó kết nối với một đường dẫn (URL) hoặc sao chép một tệp vào một vị trí cụ thể. Tìm kiếm thông qua các chuỗi có thể là một cách đơn giản để có được gợi ý về tính liên quan đến chức năng của phần mềm độc hại.
* Để thực hiện phân tích động, chúng tôi xây dựng phần mềm phân tích mã độc trong môi trường an toàn mà không sợ làm hại hệ thống thực. Khi phần mềm độc hại chạy trong một hệ thống bị cô lập, chúng tôi sẽ theo dõi các nội dung sau: Các lệnh gọi giao thức lập trình ứng dụng (API) hệ thống và các đối số của chúng được thực hiện bởi tất cả các quy trình được sinh ra của phần mềm độc hại. Phần mềm độc hại gọi hệ thống đòn bẩy để thực hiện các hành vi độc hại trên hệ thống. Tương tác hệ thống tệp có thể hiển thị tất cả các tệp mà mã độc tạo hoặc cấu hình các tệp mà nó sử dụng. Các hoạt động mạng giúp chúng tôi xác định các kết nối mạng mà phần mềm độc hại giao tiếp với nhau. Thông tin này là các chỉ số của phần mềm độc hại và có thể được sử dụng theo nhiều cách khác nhau để mô tả gia đình của họ.

b2. Khả năng áp dụng:

* Giải pháp được thực hiện tại phòng thí nghiệm với 500 mẫu thực hiện và cho kết quả chính xác đến 96,3%. Vì vậy có thể áp dụng cho các hệ thống thông minh hiện nay tại Bình Dương và trong nước.

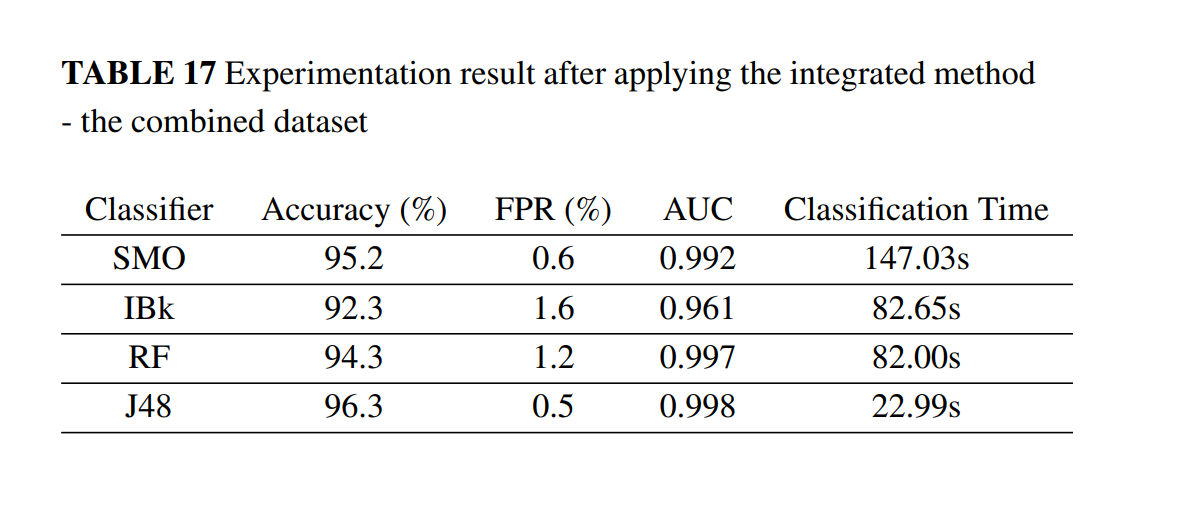
b3. Lợi ích kinh tế - xã hội:

* Giải pháp khi được tích hợp vào các thiết bị IoT sẽ đưa ra các cảnh báo bảo mật nhanh, giúp xử lý các sự cố bảo mật nhanh hơn. Vì vậy giải pháp phát hiện sớm mã độc và giúp cho nó tránh lây lan ra toàn mạng và ăn cắp dữ liệu người dùng.

1. **Kết quả của giải pháp:**

Kết quả thử nghiệm cho thấy giải pháp đạt được độ chính xác phân loại cao (96,3%) với không gian tính năng nhỏ (500 tính năng). Giải pháp cho thấy rằng phương pháp này tăng cường hiệu quả cho phương pháp phân tích mã độc trên môi trường IoT hiện nay.

Bảng 1. Bảng kết quả phân tích



1. **Đánh giá giải pháp:**
2. **Tính mới và tính sáng tạo:**
   * **Điểm mới:** *(trình bày những điểm mới của giải pháp tạo ra)*
   * **Điểm sáng tạo:** *(trình bày những điểm sáng tạo của giải pháp)*
   * **Đánh giá:***(đánh dấu vào những ô thích hợp hoặc viết riêng)*
   * Lần đầu tiên thực hiện tại khu vực phía Nam.
   * Đang có tính mới với khu vực phía Nam.
   * Lần đầu tiên thực hiện trong nước.
   * Đang có tính mới trong nước
   * Có tính mới với thế giới
   * Đã được cấp Bằng độc quyền Sáng chế hoặc Giải pháp hữu ích.
   * Được thiết kế mô phỏng theo công nghệ nước ngoài đã nhập vào Việt Nam.
   * Được thiết kế theo tài liệu nước ngoài mà công nghệ chưa vào Việt Nam.
   * Có tính sáng tạo về công nghệ.
   * Có tính sáng tạo trong kết cấu.
   * Mang tính kết hợp kỹ thuật của các ngành.

Ở việt nam có một số công cụ phân tích mã độc nhưng giải pháp này của chúng tôi áp dụng các công nghệ mới như deep learning trong phân tích mã độc.

1. **Khả năng áp dụng:**
   * **Trình bày về khả năng áp dụng vào thực tế của giải pháp tạo ra:**

Giải pháp được thực hiện tại phòng thí nghiệm với 500 mẫu thực hiện và cho kết quả chính xác đến 96,3%. Vì vậy có thể áp dụng cho các hệ thống thông minh hiện nay tại Bình Dương và trong nước.

* + **Đánh giá:**
  + Triển khai được ngay so với trình độ kỹ thuật và thị trường vật tư trong nước hiện nay.
  + Cần thêm một số chủng loại vật tư nhập không thông dụng (Có thể nêu rõ thêm bên dưới)
  + Cần có thêm một số điều kiện kỹ thuật khác (Có thể nêu rõ thêm bên dưới)
  + Có tính áp dụng đơn chiếc.
  + Có tính áp dụng trong sản xuất loại nhỏ hoặc theo thời vụ.
  + Có tính áp dụng ở quy mô công nghiệp.
  + Có khả năng áp dụng đại trà.

1. **Hiệu quả:**
   * **Kỹ thuật:**

Giải pháp dễ triển khai và vận hành trong hầu hết các thiệt bị IoT hiện nay.

* + **Kinh tế:**

Giải pháp phần mềm nên ít chi phí đầu tư ban đầu khi triển khai và vận hành .

* + **Xã hội:**

Giải pháp khi được tích hợp vào các thiết bị IoT sẽ đưa ra các cảnh báo bảo mật nhanh, giúp xử lý các sự cố bảo mật nhanh hơn. Vì vậy giải pháp phát hiện sớm mã độc và giúp cho nó tránh lây lan ra toàn mạng và ăn cắp dữ liệu người dùng.

1. **Mức độ triển khai:***(đánh dấu vào những ô thích hợp hoặc viết riêng)*
   * Nội dung thiết kế đã có đủ các thông số kỹ thuật và khả thi.
   * Đã thử nghiệm thành công (có phụ lục, tài liệu minh chứng đính kèm) từ ngày……tháng……năm………
   * Đã ứng dụng trong sản xuất ở quy mô nhỏ từ ngày……tháng……năm………
   * Đã ứng dụng và sản xuất ổn định, đại trà từ ngày……tháng……năm………

*(Lý giải thêm):*

1. **Phụ lục minh họa:** *(Đánh dấu vào các ô thích hợp hoặc viết riêng)*
   * Bằng độc quyền sáng chế, Giải pháp hữu ích, kiểu dáng công nghiệp.
   * Phiếu chứng nhận, kiểm nghiệm chất lượng hoặc độc tố.
   * Bản vẽ, sơ đồ.
   * Mô hình, vật mẫu.
   * Các kết quả đo, khảo sát thử nghiệm (có dấu xác nhận của cơ quan thử nghiệm, đo lường)
   * Bản nhận xét của các chuyên gia trong ngành.
   * Bản nhận xét của các chuyên gia trong các ngành chức năng liên quan.
   * Hợp đồng kinh tế đã ký kết.
   * Phiếu nhận xét của khách hàng.

*Phụ lục hay lý giải khác:*

1. **Các thuyết minh khác:**

Ngày ……tháng……năm 20

Tác giả

(hoặc đại diện tác giả)