التنبؤ بالبرمجيات الضارة في بيئة إنترنت الأشياء: نهج التنقيب عن البيانات

عبدالمحسن عطاالله سعود الحربي

+

إشراف

أ.د مد عبدال حميد

د. حسام لحظه

المستخلص

يُعد مجال الأمن السيبراني مجالاً بحثيًا مهمًا لأنه أكثر عرضة للهجمات على مستوى نظام الأجهزة الحاسوبية أ و على مستوى الشبكة من قبل مجرمي الإنترنت. يواجه الباحثون من التحديات في اكتشاف الثغرات الأمنية في الأجهزة الحاسوبية على سبيل المثال استقبال حزم ضارة على مستوى الشبكة أو تحليل البرمجيات الضارة لمعرفة هل هي ضارة أو غير ضارة على الأجهزة. إن الهدف الرئيسي لهذا البحث هو التنبؤ للبرمجيات الضارة في حركة مرور شبكة إنترنت الأشياء من أجل حماية أجهزة إنترنت الأشياء وتقليل نقاط الضعف باستخدام التعلم الآلة وتقنيات التنقيب عن البيانات. تركز الدراسات الحالية على اكتشاف التهديدات البرمجيات الضارة وتتجاهل أهمية التنبؤ بتهديدات البرمجيات الضارة التي تزيد من نقاط الضعف في أجهزة إنترنت الأشياء وتقليل ادائها. يستخدم مجرمي الإنترنت هجمات بالبرمجيات الضارة التي تستهدف ثغرات محددة بالأجهزة لتقوم بتعطيل النظام وسرقة البيانات والمعلومات للمستخدم بالإضافة إلى ذلك، تقوم أيضًا بحذف وتشفير الملفات من غير معرفة المستخدم أو إعطاء أي صلاحية. في هذا البحث استخدمنا خوارزميات التعلم الآلة و تقنيات التنقيب عن البيانات اثبت إمكانياتها في نموذج التنبؤ بمختلف خوارزميات التعلم الآلة للتنبؤ بالبرمجيات الضارة على الشبكة إنرنت الأشياء في مختلف أجهزة الإننرنت الأشياء. ومع ذلك قمنا بتحسين الأمان على الأجهزة التي تستهدف من قبل مجرمي الإنترنت، ايضًا قمنا بتنبؤ بأنواع الهجمات البرمجيات الضارة من خلال نموذج المقترح التنبؤية للبرمجيات الضارة. وعلاوة عن ذلك، حصلت احد خوارزيمات التعلم الآلة بنسبة دقة 97.14%وتنبأت ٨٧٥٤ عينة في مختلف البرمجيات الضارة في حركة مرور شبكة إنترنت الأشياء.

Predicting Malicious Software in IoT Environment: Data Mining Approach

by

Abdulmohsen Atallah Saud Alharbi

Advisor

Prof. Md Abdul Hamid

Co-Advisor

Dr. Husam Lahza

Abstract

The Internet of Things (IoT) enables devices to sense and respond by using the power of computing to autonomously come up with the best solutions for any industry today. However, IoT has vulnerabilities that make it easily hacked by cybercriminals. Cybercriminals know where IoT vulnerabilities are, such as unsecured update mechanisms and malicious software (malware) to attack IoT devices. The recently posted IoT-23 dataset based on several IoT devices such as Echo devices, Hue device and Somfy door lock device were used for machine learning classification algorithms and data mining techniques with training and testing for predictive modelling of a variety of malware attacks like Distributed Denial of Service (DDoS), Command and Control (C&C) and various IoT botnet like Mirai and Okiru. This research aims to develop predictive modeling that will predict malicious software in order to protect IoT environment and reduce vulnerabilities by using machine learning and data mining techniques. We collected, analyzed, and processed benign and several malware in IoT network traffic. Malware prediction is crucial in maintaining the safety of IoT devices from cybercriminals. Furthermore, Principal Component Analysis (PCA) method was applied to determine the important features of IoT-23 dataset. In addition, we have compared our study with previous studies that used the same dataset in terms of accuracy rate and other performance metrics. The results show that Random Forest classifier achieved a classification accuracy of 97.14% as well as predicted 8,754 samples various types of malware and obtained 96.44% of the Area Under Curve (AUC). Thus, Random Forest classifier outperforms several baseline machine learning classification models.