University of Mosul
College of Engineering

Securing Internet of Things (IoT) Infrastructure Using Cooperative, Hybrid & Lightweight Intelligent Intrusion Detection & Prevention System (LW-IDPS)

Sahar Lazim Qaddoori

Ph.D. Thesis in Computer Engineering\Computer Engineering

Supervised By
Prof. Dr. Qutaiba Ibrahim Ali

2022 A.C. 1444 A.H.

ABSTRACT

The Industrial Internet of Things (IIoT) is the integration of traditional Internet of Things (IoT) principles with industrial sectors and applications. The rapid evolution of edge-based IIoT had brought new difficulties in the construction of an efficient Intrusion Detection and Prevention System (IDPS) because edge devices have several restrictions, limiting their ability to perform higher-level processing tasks required by IDPSs.

Accordingly, the thesis proposes a Machine Learning (ML) model to identify the Message Queuing Telemetry Transport (MQTT)-based threats utilizing IDPS to guard the traffic network between edge node and IoT sensors. Also, it proposes two ML models to detect the abnormality in periodic and daily data metering. Because the ML model can't be trained directly on low-performance devices (like edge devices), so, a new methodology for updating ML models is proposed to detect new attacks and abnormalities.

Consequently, the Random Forest classifier has been used to eliminate the unrelated features from the MQTTset dataset to reduce the preprocessing time (decreased to more than 29%) and the complexity of the model based on the MQTT packet level. In addition, the performance comparison between three ML algorithms prove that Decision Tree (DT) performs comparatively better to detect attacks at the MQTT packet level because of its low prediction time (0.66304msec) and good accuracy with the MQTTset dataset (99.69%). As Density-based Spatial Clustering of Applications with Noise (DBSCAN) is suitable for clustering smart metering dataset because of the Silhouette Coefficient score of 0.663. Also, DT

performs well for detecting anomalies in daily power consumption. While One Class Support Vector Machine algorithm (OCSVM) is good to detect abnormalities in periodic power consumption due to its short prediction time (0.95319 msec) and small model size (11 KB).

Then, the security techniques (Transport Layer Security protocol (TLS), firewall, and Wi-Fi Protected Access II protocol (WPA2)) have been employed to verify that the exchanged data during the updating strategy is valid and undiscoverable.

Based on the results of network performance and resources usage, the structure of proposed edge device is lightweight and saves power. Finally, the proposed security paradigm for edge device is simply applied on a low-cost single-board computer (SBC) like the Raspberry Pi and is effective against a variety of internal and external threats.

جامعة الموصل
كلية الهندسة

تامين البنية التحتية لإنترنت الاشياء باستخدام نظام خفيف الوزن تعاوني هجين لكشف ومنع التطفل (LW-IDPS)

اطروحة دكتوراه تقدمت بها

سحر لازم قدوري

إلى

مجلس كلية الهندسة في جامعة الموصل كجزء من متطلبات

نيل شهادة دكتوراه فلسفة في هندسة الحاسوب / هندسة الحاسوب

بإشراف

الأستاذ الدكتور قتيبة ابراهيم علي

1444هـ 2022م

الخلاصة

إنترنت الأشياء الصناعي (IIoT) هو دمج مبادئ إنترنت الأشياء (IoT) التقليدية مع القطاعات والتطبيقات الصناعية الحديثة. أدى التطور السريع لإنترنت الأشياء المستندة إلى الحافة إلى ظهور صعوبات جديدة في بناء نظام فعال للكشف عن التسلل والوقاية (IDPS) لأن أجهزة الحافة لها قيود عديدة ، مما يحد من قدرتها على أداء مهام المعالجة ذات المستوى الأعلى التي تتطلبها IDPS.

وفقًا لذلك ، تقترح الأطروحة نموذجًا للتعلم الآلي (ML) لتحديد التهديدات القائمة على نقل القياس عن بُعد في قائمة انتظار الرسائل (MQTT) باستخدام IDPS لحماية شبكة المرور بين عقدة الحافة ومستشعرات إنترنت الأشياء. كما يقترح نموذجين للتعلم الآلي (ML) لاكتشاف الشذوذ في قياس البيانات الدورية واليومية. نظرًا لأنه لا يمكن تدريب نموذج التعلم الآلي مباشرة على الأجهزة منخفضة الأداء (مثل أجهزة الحافة) ، لذلك تم اقتراح منهجية جديدة لتحديث نماذج التعلم الآلي لاكتشاف الهجمات والتشوهات الجديدة.

وبالتالي ، تم استخدام مصنف الغابة العشوائية لإزالة الميزات غير ذات الصلة من مجموعة بيانات ال MQTTset لتقليل وقت المعالجة المسبقة (حيث انخفض إلى أكثر من 29٪) وتعقيد النموذج بناءً على مستوى حزمة ال MQTT. بالإضافة إلى ذلك ، تثبت مقارنة الأداء بين ثلاث خوارزميات التعلم الالي أن شجرة القرار (DT) تعمل بشكل أفضل نسبيًا للكشف عن الهجمات على مستوى حزمة MQTT نظرًا لوقت التنبؤ المنخفض (0.66304 مللي ثانية) والدقة الجيدة (99.69)٪ مع مجموعة بيانات الMQTTset. نظرًا لأن التجميع المكاني القائم على الكثافة للتطبيقات ذات الضوضاء (DBSCAN)

مناسب لتجميع مجموعة بيانات القياس الذكي بسبب درجة معامل Silhouette التي تبلغ 0.663. أيضًا ، يعمل DT جيدًا لاكتشاف الحالات الشاذة في استهلاك الطاقة اليومي. بينما تعد خوارزمية آلة المتجه للدعم من فئة واحدة (OCSVM) جيدة لاكتشاف التشوهات في الاستهلاك الدوري للطاقة نظرًا لقصر وقت التنبؤ (0.95319 مللي ثانية) وصغر حجم النموذج (11 كيلوبايت).

بعد ذلك ، تم استخدام تقنيات الأمان (بروتوكول أمان طبقة النقل (TLS) وجدار الحماية وبروتوكول الوصول المحمي بشبكة ال Wi-Fi (WPA2)) للتحقق من أن البيانات المتبادلة أثناء استراتيجية التحديث صالحة وغير قابلة للاكتشاف.

استنادًا إلى نتائج أداء الشبكة واستخدام الموارد ، فإن تصميم جهاز الحافة المقترح خفيف الوزن ويوفر الطاقة. أخيرًا ، يتم تطبيق نموذج الأمان المقترح لجهاز الحافة ببساطة على كمبيوتر منخفض التكلفة أحادي اللوحة (SBC) مثل Raspberry Pi وهو فعال ضد مجموعة متنوعة من التهديدات الداخلية والخارجية.