وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

بناء نموذج لكشف التسلل يعتمد على التعلم العميق

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
البرمجيات

من قبل
طيبة علي جاسم علي النصرالله

بإشراف
م.د. منى محمد طاهر جوهر

١٤٤٤ هـ
٢٠٢٢ م

الملخص

في الوقت الحاضر، تتزايد الأجهزة المتصلة بالشبكة مثل الهواتف المحمولة وأجهزة إنترنت الأشياء، لذلك من الضروري حماية هذه الاجهزة من الهجمات المعقدة التي تتعرض لها الشبكة ، يتم ذلك عن طريق بناء الأدوات والتطبيقات التي تكتشف الهجمات أو الحالات الشاذة. تعاني العديد من انظمة كشف الهجمات من نتائج إيجابية كاذبة عالية لذلك يجب تقليل هذا المعدل المرتفع بالإضافة الى ذلك تعاني من سوء تصنيف لنوع الهجوم لأنها تستخدم أنماط تصنيف هجوم تعتمد على التوقيع والتي تكتشف أنواع الهجمات الشائعة وليس لديها القدرة على اكتشاف أنواع جديدة، و للتغلب على هذه العيوب تستخدم خوارزميات التعلم العميق بوضعها كنموذج في هذا النوع من التطبيق.

في انظمة كشف التسلل على الشبكة تتم مراقبة نوع الاتصال العادي في حركة مرور الشبكة عن طريق تحديد حزم البيانات الحميدة المسجلة مسبقًا ، ويساعد في اكتشاف أنواع جديدة عند تحليل الانحرافات المرورية في الشبكة .

فيهذه الرسالة تم بناء نموذج يكشف (يصنف)الهجمات على الشبكة باستخدام تقنيات التعلم العميق اذ اعتمدنا الشبكةالعصبية التلافيفية (Convolutional Neural Network) وشبكة ذاكرة طويلة المدى(Long Short-Term Memory) ، واستنادًا إلى مجموعتي بيانات (KDD-CUP1999) وبيانات(CIC-2018) وهي بيانات شائعة الاستخدام في انظمة الكشف عن التسلل والهجمات المختلفة والتي تحتوي على انواع منالهجمات الشائعة والحديثة فضلا عن ذلك اعتمدنا نوعين من التصنيف الثنائي للفئات وتصنيف متعدد مع اظهار دقة اختبار النموذج في التدريب والخسارة .

حيث حصلنا على نسبة دقة كشف باستخدام الشبكةالعصبية التلافيفية مع بيانات (CIC-2018) تصنيف ثنائي(٠.٩٩٨) مع نسبة خسارة (٠.٠٠٠٤٦) وتصنيف متعددبلغت نسبة الدقة (٠.٩٩١) ونسبة الخسارة(٠.٠٠٢٤) ومع مجموعة بيانات (KDD-CUP1999) بلغت نسبة الدقة (٠.٩٧٣) ونسبة الخسارة(٠.٠٠٤٢) .

وحصلنا على نسبة دقة كشف باستخدام شبكة ذاكرة طويلة المدى(LSTM) مع بيانات (CIC-2018) تصنيف ثنائي(0.986) مع نسبة خسارة (٠.٠٧٩) ومع مجموعة بيانات (KDD-CUP1999) بلغت نسبة الدقة (٠.٩٥) ونسبة الخسارة(0.043) .

Ministry of Higher Education and
Scientific Research
University of MosulCollege of Computer
Science and MathematicsDepartment of Software

Building an Intrusion Detection Model Based on Deep Learning

Thesis Submitted to the Council of the College of Computer Science & Mathematics University of Mosul as a Partial Fulfillment of Requirements for the Degree of Master of Science in Software

By

Teba Ali Jasim Ali

Supervised by

Dr. Muna Mohammad Taher Jawhar

2022 A.D. **1444 A.H.**

Abstract

Nowadays, more and more devices are connected to the network, such as mobile phones and IoT devices, so it is necessary to protect these devices from complex network attacks. This is done by building tools and applications that detect attacks or anomalies. Many attack detection systems suffer from high false positive rates, so this high rate must be reduced. In addition, they suffer from misclassification of the type of attack because they use patterns A signature-based attack classification that detects common attack types but does not have the ability to detect common attack types. In order to overcome these shortcomings, deep learning algorithms are used by setting them as a model in this type of application. In network intrusion detection systems, the normal connection type in network traffic is monitored by identifying data packets. The benign is pre-recorded and helps in discovering new types when analyzing traffic deviations in the network. In this thesis, a model was built that detects (classifies) attacks on the network using deep learning techniques as we adopted the Convolutional Neural Network. and Long-Short-Term Memory, based on two datasets of (KDD -CUP1999) and (CIC-2018) data, which are commonly used data In the systems of detection of intrusion and various attacks, which contain types of common and modern attacks, we have adopted two types of binary classification: multiple classification with showing The accuracy of the model was tested in training and loss, Where we obtained a detection accuracy ratio using the Convolutional Neural Network with (CIC-2018) data, a binary classification (0.998) with a loss ratio (0.0046) and a multiple classification accuracy rate of (0.991) and the loss ratio (0.024) .

and with the (KDD -CUP1999) data set the accuracy was (0.973) and the loss ratio was (0.042).

We obtained a detection accuracy ratio using long-term memory with (CIC-2018) binary classification data (0.986) with a loss ratio of (0.079) .

and with the (KDD -CUP1999) data set where the accuracy ratio was (0.95) and the loss ratio was (0.043).