

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المدرسة العليا للإعلام الآلي - 08 ماي 1945 – بسبدي بلعباس
Ecole Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès



Mémoire de Fin d'étude

Pour l'obtention du diplôme d'ingénieur d'état

Filière : Informatique

Spécialité : Ingénierie des Systèmes Informatiques (ISI)

Thème

“AndroPred: A Deep Learning Based System for Android Malware Detection”

Présenté par :

- Mr Abdelhamid GARAH

Soutenu le : **23/09/2020**

Devant le jury composé de :

- Mr Zakaria GHEID
- Mr Alaa Eddine BELFEDHAL
- Mme Djihed ANANI

Président
Encadreur
Examineur

Année Universitaire : 2019 / 2020

Abstract

Android is the most used mobile operating system in the world. Due to its popularity, malware is increasing every year, posing many problems for users. As malware has increased, anti-malware solutions have as well. Researchers persistently devise counter-measures strategies to fight back malware, one of these strategies is the use of machine learning and deep learning methods to detect Android malware.

In this work in order to develop our framework "AndroPred", we propose a static detection method based on deep learning. We directly extract AndroidManifest.xml file from Android APK file, and convert the Manifest file into a bytecode image, then use the convolution neural network (CNN), to train a detection model and apply it to classify malware. CNN can automatically learn features of bytecode file which can be used to recognize malware.

key words : Android security, malware detection, convolution neural network (CNN) and bytecode image.

Résumé

Android est le système d'exploitation mobile le plus utilisé au monde. En raison de sa popularité, les logiciels malveillants augmentent chaque année, ce qui pose de nombreux problèmes aux utilisateurs. Les logiciels malveillants ont augmenté, tout comme les solutions anti logiciels malveillants. Les chercheurs élaborent sans cesse des stratégies de contre-mesure pour lutter contre les logiciels malveillants, l'une de ces stratégies est l'utilisation de méthodes d'apprentissage automatique et d'apprentissage profond pour détecter les logiciels malveillants Android.

Dans ce travail, afin de développer notre framework "AndroPred", nous proposons une méthode de détection statique basée sur un apprentissage profond. Nous extrayons directement le fichier AndroidManifest.xml à partir du fichier APK d'Android, et nous convertissons le fichier Manifest en une image bytecode, puis nous utilisons un réseau de neurones convolutifs (CNN), pour entraîner un modèle de détection et l'appliquer à la classification des logiciels malveillants. CNN peut apprendre automatiquement les caractéristiques du fichier bytecode qui peut être utilisé pour reconnaître les logiciels malveillants.

mots clés : sécurité Android, détection des logiciels malveillants, un réseau de neurones convolutifs (CNN) et image bytecode.

ملخص

أندرويد هو نظام التشغيل المحمول الأكثر استخداما في العالم. نظرا لشعبيته، تزداد البرامج الضارة كل عام، مما يطرح العديد من المشاكل للمستخدمين. نظرا لزيادة البرامج الضارة، زادت أيضا حلول مكافحة البرامج الضارة. ابتكر الباحثون عدة إستراتيجيات لمحاربة البرامج الضارة، واحدة من هذه الاستراتيجيات المطبقة هي استخدام التعلم الآلي وطرق التعلم العميق للكشف عن برامج أندرويد الضارة.

في هذا العمل من أجل تطوير أدواتنا "AndroPred"، نقترح طريقة اكتشاف ثابتة تعتمد على التعلم العميق. نستخرج مباشرة ملف "AndroidManifest.xml" من ملف الاندرويد "APK"، ونحول ملف "Manifest" إلى صورة الرمز الثنائي، ثم نستخدم الشبكات العصبونية الالتفافية، لتدريب نموذج الكشف وتطبيقه لتصنيف البرامج الضارة. يمكن للشبكات العصبونية الالتفافية التعرف تلقائيا على ميزات ملف الرمز الثنائي والتي يمكن استخدامها للتعرف على البرامج الضارة.

الكلمات الرئيسية: أمان أندرويد، اكتشاف البرامج الضارة، الشبكات العصبونية الالتفافية و صور الرموز الثنائية