CNNを用いたPE内関数の類似性によるマルウェア検知手法

A Malware Detection Method by Function Similarity in PE Files using Convolution Neural Network

田村壮世・法制倫理分科会・情報セキュリティ大学院大学

Many of the conventional malware detection methods are based on signature-based pattern matching, In the case of unknown malware, signature creation could not be made in time and infected with malware. Also, in recent years targeted attacks have made it possible to use dedicated malware that aims only at specific organizations and individuals, but detection is becoming more difficult.
In this research, we focused on Portable Executable (PE) format which is Windows executable file, extracted multiple functions from the section containing Entry Point (EP) and used convolutional neural network (CNN) Do deep learning. By this, we classify malignant and benign by function and propose a method to comprehensively judge whether file is malignant or benign from this classification result. Since this method does not use signatures, even unknown malware can be detected.

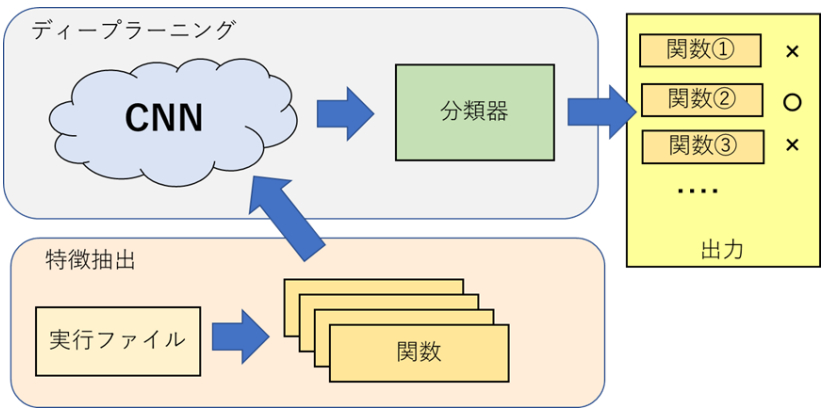
はじめに

従来手法では、シグネイチャ作成までのタイムラグがあり新種のマルウェアの検出が難しいため、ディープラーニングを用いた機械学習ベースの手法を提案

提案手法

実行ファイルから関数コードを抽出・学習し、関数ごとの判定結果から総合的にマルウェアかどうかを判定する。

※学習には、14,169個のマルウェアとvector等から取得した8,000個の実行ファイルを使用した。



結　果

学習及び検知テストに使用したデータセットにおいて98.5％の精度でマルウェアと通常のソフトウェアを分類できた。また、2018年１月に取得した最新のマルウェアの検知結果においてはK社製ソフトよりもマルウェアを検知した。

	条件	検知数	未検知数	検知率	実行時間
K社製ソフト	最新の定義ファイル	135	117	54%	数分
提案手法	半年前の学習データ	190	62	75%	２４秒