# 平成27年度 情報工学コース卒業研究報告要旨

| 村瀬・嶋田 研究室 | 氏　　名 | 飛 山 駿 |
|---|---|---|
| 卒業研究題目 | Deep Neural Network によるプロセスの振舞いに着目したマルウェア検出 | |

近年では、マルウェアの数は劇的に増加しており、また標的型攻撃などの経済的な利益や機密情報の奪取を目的とした高度で巧妙な攻撃が深刻な問題となっている。このような攻撃では、攻撃に未知のマルウェアが使用されることも多く、従来のマルウェア検知手法では攻撃を防ぎきることが難しくなっている。そのため、マルウェアに感染していることを前提とした対策が必要とされている。

本研究では、感染端末の検知のために、PC のプロセスログからそのプロセスのマルウェア度を推定することでマルウェアプロセスを検出する手法を提案する。本手法では Deep Neural Network による機械学習を用いて、マルウェアを感染させた PC のマルウェアプロセスのログファイルからマルウェアプロセスの特徴を学習し、学習した特徴を利用してマルウェアプロセスを検出する。提案手法の流れを図 1 に示す。提案手法の学習段階では、まず実行されている全プロセスの API コールのログを記録し、プロセスごとの API コールログファイルを作成する。次に、記録したログから API コールの言語モデルを学習する。学習は Long Short-Term Memory(LSTM) を使用した Recurrent Neural Network(RNN) により行い、プロセスの振舞いをよく表す特徴を学習する。そして、学習した RNN を使用してログから特徴を抽出し、抽出したプロセスの特徴を画像化する。画像化した特徴を Convolutional Neural Network(CNN) により学習し、マルウェア/正常プロセスの 2 クラス分類器を作成する。評価段階においては、評価対象 PC のログより画像化したプロセスの特徴を作成して分類器にかけ、分類結果をシグモイド関数により 0 から 1 の範囲に射影することでそのプロセスのマルウェア度を算出する。

提案手法の性能評価のため、ネットワーク的に遮断された仮想環境上の WindowsXP でマルウェア 26 種類を実行して記録したログを使用して実験を行った。実験では、44 個のマルウェアプロセス、39 個の正常プロセスのログを使用して RNN の学習を行い、RNN の学習に使用したログを含めた 150 個のログを使用して 5 分割交差検証により CNN の学習と評価を行った。RNN から抽出される特徴ベクトルの長さを 350/30/20 とした時の特徴画像のサイズをそれぞれ 350×350/30×30/20×20 と変化させ、結果を比較した。実験の結果、画像サイズが 350×350/30×30/20×20 のときの ROC 曲線下の面積の平均は、それぞれ 87.4/95.6/91.8 となった。

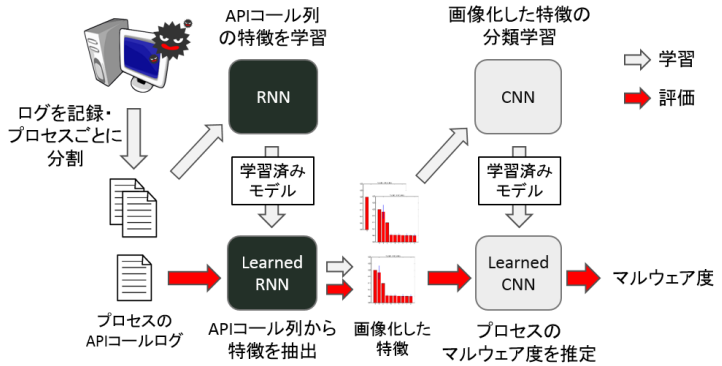


図 1 提案手法の流れ