講演番号：133

ニューラルネットワークを用いた悪性 PDF の検知

Malicious PDF Detection Using Neural Network

D−19

島部凌太朗　伊與田 光宏
Ryotaro SHIMABE　Mitsuhiro IYODA
千葉工業大学 情報工学科
Department of Computer Science, Chiba Institute of Technology

1. はじめに

文書ファイルの交換に用いられる形式として、PDF (Portable Document Format) が広く利用されている。しかし、PDF には、 JavaScript などのコードを実行する機能や、ファイルを添付する機能があり、これを悪用して PDF にマルウェアを埋め込んだ悪性 PDF をメールに添付して送信する標的型攻撃が増加している。実行形式のファイルに比べて、 PDF は単なる文書ファイルとして認識されているため、ユーザは何の疑いもなく悪性 PDF を開いてしまう可能性がある。

2. 目的

本研究の目的は、ニューラルネットワークを用いて悪性 PDF を検知することである。また、提案手法が未知の悪性 PDF 検知に対しても有効であるか評価する。

3. 既存研究

悪性 PDF の検知手法は、PDF に埋め込まれたコードを仮想環境内で実行して挙動を解析する動的解析手法と、コードを実行せずに解析する静的解析手法に分けられる。動的解析手法では、近年のマルウェアには仮想環境を検知する機能を持つものもあり、仮想環境内では正常に動作するため、マルウェアを検知できない場合がある。静的解析手法では、JavaScript のコードを直接解析することで、悪性 PDF の検知を行う手法が提案されている。しかし、圧縮されたコードの場合、コードが解析できず、悪性 PDF を検知できないという問題がある。近年の研究では、 PDF の圧縮されない内部構成要素に着目し、機械学習を用いて解析する手法が提案されている。

既存研究 [1] では、PDF のファイル構造から参照関係を表すパスを抽出し、これを学習させることによって良性 PDF と悪性 PDF を分類する手法が提案されている。このパスはキーワードと呼ばれる構成要素によって定義される。

4. 提案手法

本研究では、既存手法と同様に、良性 PDF と悪性 PDF の間にあるキーワードの出現頻度の違いに着目した検知を行う。学習用データセットに含まれる PDF を全て読み込み、抽出したキーワードからキーワードの集合を作成する。次に、個々の PDF に対して各キーワードの出現数をカウントし、 Bag-of-words を作成する。これを PDF のベクトルとして、ニューラルネットワークに入力し、学習させる。本システムの処理の流れを図 1 に示す。

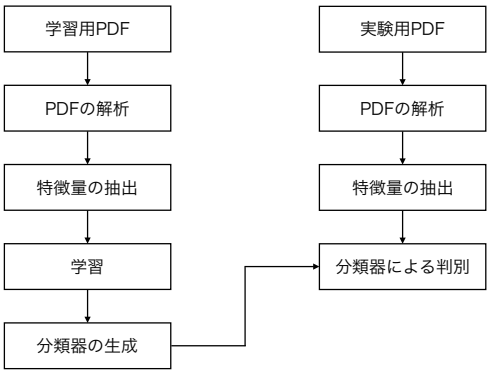

図 1. 処理の流れ

5. 実験

データセットを用いた実験によって既存手法と提案手法の比較を行う。 PDF の実験サンプルとして、Cotangio [2] で公開されている PDF のデータセットを使用する。このデータセットには、 19,982 個の PDF が含まれており、 9,000 個の良性 PDF と 10,982 個の悪性 PDF で構成されている。

また、未知の悪性 PDF への検知性能を評価するために、上の実験で作成した学習済みの分類器による、別のデータセットを用いた実験を行う。この実験では、VirusShare.com [3] から収集した 128 個の悪性 PDF を使用する。

6. おわりに

本稿では、 PDF を構成するキーワードから Bag-of-words で PDF をベクトル化し、ニューラルネットワークを用いて学習と分類を行うことで、悪性 PDF を検知する手法を提案した。

参考文献

[1] Nedim Šrndić and Pavel Laskov, “Hidost: a static machine-learning-based detector of malicious files,” EURASIP Journal on Information Security (2016), 2016
[2] “Contagio,” http://contagiodump.blogspot.jp
[3] “VirusShare.com,” https://virusshare.com