2E-04

One-Class SVM を用いたマルウェア PDF 検出手法

岩本舞 †　　小島俊輔 †　　中嶋卓雄 ‡
† 熊本高等専門学校　　‡ 東海大学

1 はじめに

PDF（Portable Document Format）は，現在世界中で広く利用されているファイル形式である．一般に PDF ファイルは静的文書とみなされており，実行ファイルやマクロファイルと比べ，危険性はあまり認知されていない．しかし PDF は様々なコードを実行可能な形式であり，攻撃にも利用されている．

そこで我々は，One-Class Support Vector Machine（以下 One-Class SVM）を用いた教師なし学習により PDF マルウェアを検出する手法を提案する．本手法では，あらかじめ 正常 PDF のみを学習した分類器を作成し，外れ値を検出することでマルウェア PDF を検出する．

2 従来研究

マルウェア解析手法は，大きく動的解析と静的解析に分類される．動的解析は複雑に難読化されたファイルも扱うことができるが，実行環境がマルウェアに感染する危険性がある．一方静的解析は，コードの実行を伴わない，比較的安全な解析手法である．従来研究では，JavaScript の解析を行う手法や，メタデータの特徴や文書構造によりマルウェアと正常な PDF ファイルを判別する手法が提案されており，分類には主に SVM や Random Forest Classifier といった教師あり機械学習手法が用いられている．

本手法は，静的解析に分類される．正常 PDF のみを用いた教師なし学習により正常 PDF の特徴ベクトルからはずれたものをマルウェアとして検出するため，新手のマルウェアや，標的型攻撃などサンプルの収集が難しいマルウェアの検出が期待できる．

3 提案手法

本稿では，12 次元の特徴ベクトルを用い，あらかじめ正常 PDF のベクトルを One-Class SVM で学習し，外れ値が検出された場合にマルウェアとみなす手法を提案する．SVM ではベクトルの要素を 0〜1 に正規化する必要があり，今回は Flag（特徴の有無を 0/1 の二値で正規化），nNumber（数値 N について 10 を上限とし $N/10$ で正規化），nLength（N byte の長さを 100MB を上限に $log_{10}N$ で正規化）の三種類の正規化を用いた．実験に用いた特徴の概要を以下に示す．

(a) 不正な `/Length 0`
ストリームオブジェクトでは，ストリームの長さを示す `/Length` の値を指定する必要がある．マルウェアではストリームが存在するにもかかわらず`/Length 0` となっているものが見られた．

(b) JavaScript の使用
マルウェア PDF は主に JavaScript を用いて攻撃を行う．JavaScript は `/S /JavaScript` または `application/x-javascript` で実行できる．正常な PDF ファイルにも JavaScript を用いているファイルは存在するが，正常 PDF では複雑なプログラムを利用することは少ないため，ソースコードの情報量により区別できる．ただし，一部のマルウェアは以下 (c) (d) の工夫によりソースコードを少なく見せている場合がある．

(c) `this.info` および `getAnnot` を用いた難読化
PDF では JavaScript コードを指定する名前 `/JS` が用いられており，簡単にコードを取り出せる．そこで，難読化のために `this.info` および `getAnnot` を用いたマルウェアが存在する．この手法では，別のコードや難読化したコードなどをプロパティ情報や注釈情報として埋め込み，JavaScript コードから参照する．また通常，プロパティ情報に大量のデータを埋め込むことはないため，プロパティ情報の長さもベクトル要素とした．

(d) `eval` および `Function` によるコード実行
JavaScript では，`eval` や `Function` を用いることで文字列をコードとして実行することができる．マルウェアでは，難読したコードをアンパックした文字列を引数として，`eval` や `Function` によりコードを実行するものが見られた．

(e) 誤った文法
一般に PDF はソフトウェアにより出力されるため，基本的な文法ミスは少ないと考えられる．一方マルウェアは正常なソフトウェアで作成されていないため，xref テーブルの不整合など文法が正しくないものが多く見られた．そのため，`%%EOF` なし，`%%EOF` 後のコード，Xpdf[1] の pdfinfo コマンドで解析した際のエラーをベクトルに加えた．

A Detection Technique of Malicious PDF Files using One-Class SVM
†Mai IWAMOTO　†Shunsuke OSHIMA　‡Takuo NAKASHIMA
†National Institute of Technology, Kumamoto College
‡Tokai University

表 1: 既存手法と提案手法の正答数および Accuracy の比較

	Benign (7189)			Malicious (221)			Total Accuracy	
Method	True	False	Error	True	False	Error	exclude error	all
PDFMS	6817	4	368	207	2	12	0.9991	0.9479
Propsed ($c = 2^{-2}$, $\nu = 2^{-13}$)	7182	7	0	221	0	0	0.9991	0.9991

4 実験

4.1 データセット

今回の実験では，正常 PDF として，熊本高等専門学校八代キャンパスに設置されたプロキシサーバで収集した PDF ファイル 7189 個を用いた．また，マルウェア PDF として，2010 年から 2015 年の間に収集された D3M データセット [2] に含まれる PDF ファイル 349 個のうち，VirusTotal にマルウェアとして登録されていた 221 個を用いた．実験に用いたファイルに重複はない．

4.2 実験手法

本稿では，提案手法と既存手法における正常 PDF およびマルウェア PDF 分類の正解率を，同じデータセットを用いて比較する．

比較対象の既存手法には，PDF Malware Slayer[3]（以下 PDFMS）を用いた．PDFMS は機械学習に基づくマルウェア検出ツールで，PDF ファイルに出現するキーワードを正常セットとマルウェアセットに分類してクラスタリングし，Random Forest Classifier により正常な PDF ファイルとマルウェア PDF ファイルを判別する．

提案手法の実験には LIBSVM [4] に実装された One-Class SVM を，カーネル関数には，一般的に利用される RBF カーネル $k(x_i, x_j) = e^{-\|x_i - x_j\|^2/c}$ を用いた．RBF カーネルおよび One-Class SVM を使用するにあたっては，パラメタ c および ν を設定する必要がある．今回は予備実験の結果より，$c = 2^{-2}$，$\nu = 2^{-13}$ を使用した．

評価には，SVM を評価する際の一般的な手法である K-fold cross-validation を用い，今回は $K = 4$ とした．なお，提案手法では正常 PDF の特徴ベクトルのみを学習するため，学習用マルウェア PDF は，既存手法での学習のみに用いる．

実験では，4 つのデータセットについて正解率を算出し，平均値で評価する．ここでいう正解率とは，正しいクラスに分類されたデータの数を評価データの総数で割った値である．

4.3 実験結果

実験の結果を 表 1 に示す．既存手法における正解率はエラーとなったものを除くと 99.91%であり，提案手法でも同じ 99.91%の正解率を示した．ただし既存手法では，解析エラーとなり分類ができなかったファイルが正常 PDF で 89 個，マルウェア PDF で 12 個あった．提案手法では，エラーとなったファイルはなく，また，すべてのマルウェアを検出できた．なお，提案手法で正しく分類されなかった正常 PDF（Benign-False）7 個については，JavaScript や `%%EOF` 後のコード，シンタックスエラーといったマルウェアに類似した特徴が含まれていたファイルがあった．またそれ以外にも，分類器の作成に使用した学習用の正常 PDF と比較すると外れ値の特徴を持つものがあり，誤ってマルウェアとして検出された．本手法は外れ値検知のため，マルウェアの特徴がなくても分類器を作成する正常 PDF に類似のファイルが含まれなければ外れ値として検出されてしまうが，より多くの正常 PDF を収集し学習することにより改善できると考える．

5 まとめ

本稿では，正常な PDF ファイルの特徴ベクトルのみを One-Class SVM で学習し，外れ値に分類されたファイルをマルウェアと判定する手法を提案した．

提案手法は，分類の正解率が 99.91% と，教師あり機械学習を用いた既存手法と並ぶ高い正解率を示し，また，すべてのマルウェアを検出できた．本手法は教師あり学習と異なり分類器の作成にマルウェアが必要なく，新手のマルウェアや収集の難しい標的型攻撃のマルウェアなどの検出も期待できる．以上より，本手法は有用であると考える．

参考文献

[1] Clyph & Cog, LLC. Xpdf. `http://www.foolabs.com/xpdf/home.html`.

[2] 神薗雅紀, 秋山満昭, 笠間貴弘, 村上純一, 畑田充弘, 寺田真敏. マルウェア対策のための研究用データセット〜mws datasets 2015〜. Technical Report 6, jun 2015.

[3] Pattern Recognition and Applications Lab. Slayer. `https://pralab.diee.unica.it/en/Slayer`.

[4] Chih-Chung Chang and Chih-Jen Lin. Libsvm – a library for support vector machines. `https://www.csie.ntu.edu.tw/~cjlin/libsvm/`.