Vol.2016-MPS-111 No.18
2016/12/12

Web アプリケーションに対する OS コマンドインジェクション攻撃の検知についての考察

秋本　悠一朗[†1]　松田　健[†2]　園田　道夫[†3]　趙　晋輝[†1]

概要：我々が日常的に利用する Web アプリケーションには，多様なサイバー攻撃による脅威が存在している．その中の１つである OS コマンドインジェクション攻撃は，Web アプリケーションが OS コマンドを実行する箇所の脆弱性を狙い，OS コマンドを注入することで不正な動作を引き起こす攻撃である．従来から入力文字列を検査することが基本的な対策となっているが，攻撃と同じ特徴を持つ正常文を精度よく判別する手法の開発は重要な課題である．本研究では Web アプリケーションのフォームへ入力される OS コマンドインジェクション攻撃の攻撃文と正常文に含まれる特徴文字の頻度から特徴量を生成した．この特徴量を用いて機械学習アルゴリズムを適用し攻撃と正常の検知を試みた．

キーワード：OS コマンドインジェクション，機械学習，SVM

1. はじめに

サイバー攻撃の手法は日々進化し続けている．一方，攻撃を受ける側については基本的な対策がなされていないケースも散見される．最近では，標的型攻撃と呼ばれる人間の心理的なミスにつけ込む手法を取り入れたサイバー攻撃が大きな問題となっているが，よく知られた古い攻撃手法が通用するアプリケーションは未だに存在していると考えられる．特に，インジェクション系の攻撃と言われる Web アプリケーションに対する攻撃のうち，SQL インジェクション攻撃やクロスサイトスクリプティング攻撃は重大な被害をもたらすことで知られている．本研究で対象とする OS コマンドインジェクション攻撃は，その根本的対策が比較的容易なこともあり，サイバー攻撃に詳しい技術者や研究者にとっては古い攻撃手法として知られている．しかしながら，そのような知識を持たずに Web アプリケーションが作られる場合や，何らかの理由で対策がなされていない Web アプリケーションが存在する可能性もあるため，OS コマンドインジェクション攻撃を検知する技法の開発は不必要であるとは言い難い．本研究の目的は OS コマンドインジェクション攻撃の検知だけでなく，ユーザからの入力に悪意のある要素があるかどうかを判別することでもある．このような技術は他のサイバー攻撃の対策にも重要な技術であると考えられる．従来研究では，SQL インジェクション攻撃の特徴抽出を行い，機械学習の手法を用いて攻撃を検出する手法が研究されている[1][2]．本研究においても OS コマンドインジェクション攻撃に頻出するセミコロンやスラッシュなどを攻撃特徴文字として，機械学習による OS コマンドインジェクション攻撃の検知を試みた．

2. OS コマンドインジェクション攻撃

OS コマンドインジェクション攻撃は，Web アプリケーションの入力フォーム等に不正な OS コマンドを注入することで成立する．

例えば、以下がプログラムに組まれているとする．

```
system("sendmail $to_address <$message_file");
```

攻撃者が

```
evil@site.com < /etc/passwd; rm -rf /
```

という文字列を入力した場合，evil@site.com へ/etc/passwd ファイルが流出し，セミコロンで連結されている rm コマンドによりルートディレクトリ以下のすべてのディレクトリ・ファイルが削除される．

Web アプリケーション開発に用いられる言語の多くはシェル経由での OS コマンドの実行が可能である．Perl や PHP の sysytem()，exec()などが該当する．

3. 既存の対策手法

攻撃を検知する手法として，WAF のブラックリスト方式，ホワイトリスト方式がある[3]．ブラックリスト方式は，攻撃のパターンをブラックリストに登録し，一致する入力を拒否する手法である．ホワイトリスト方式は，あらかじめ指定した入力のみを許可し，それ以外を拒否する手法である．どちらの手法も事前にリストへ登録しておかなければならず，リストのみでは未知の攻撃に対処することが難しい欠点がある．

別の対策手法として，シェルを動かさないプログラムを組む，特殊な記号を別の文字列に置き換えるエスケープ処理などが挙げられる．前者はプログラマに高度な知識と技術が求められ，対策が浸透するまでに時間がかかってしまう．後者は OS ごとにエスケープするコマンド名・オプション名・オプション値が異なり，対策が困難である．

4. Support Vector Machine(SVM)[4]

SVM はパターン認識を行う教師あり学習の一つであり，完全な分離を目指すハードマージン SVM と多少の識別誤りを許すソフトマージン SVM がある．SVM の識別関数は以

†1 中央大学理工学部情報工学科
†2 長崎県立大学情報システム学部情報セキュリティ学科
†3 サイバー大学

下の式で定義される．

$$y(x) = \mathbf{w}^t\varphi(\mathbf{x}) + b$$

$$= \sum_{n=1}^{N} a_n t_n k(x, x_n) + b$$

ここで **w** は重みベクトル，**x** は入力ベクトル，ϕ は特徴空間変換関数，a_nはラグランジュ乗数，t_nは目標値，$k(x, x_n)$はカーネル関数，b はバイアスパラメタである．

SVM は「マージン最大化」という明確な目標と，「カーネルトリック」による非線形な問題を線形分離できる点で優れている．マージンとは分離超平面と最も近いデータとの距離であり，ソフトマージン SVM でのマージン最大化は以下の目的関数を最小化することで得られる．

$$C\sum_{n=1}^{N} \zeta_n + \frac{1}{2}||w||^2$$

$$\text{s.t. } t_n(y(x_n)) \geq 1 - \zeta_n \quad n = 1, \ldots, N$$

$$\zeta_n \geq 0$$

ζ は誤分類に対するペナルティ，C>0 はペナルティとマージンの大きさ間のトレードオフを制御するパラメタである．本研究では，Python のライブラリ scikit-learn0.17 の sklearn.svm.SVC()を使用して実験を行うことにする．使用したカーネル関数は多項式カーネル，ガウスカーネルである．なお，通常のベクトルを利用する場合を線形カーネルと呼ぶことにする．

5. 実験用データの生成

実験に用いるデータとして，OS コマンドインジェクション攻撃として有効な攻撃文を 500 個，攻撃特徴記号を多く含む正常文 500 個を人工的に生成した．正常文にはスラッシュや空白などが多く含まれる顔文字や URL を使用している．これらのデータに対して特徴抽出を行い，ランダムに各 300 個を学習データに，各 200 個をテストデータとして用いる．

攻撃文における記号の出現頻度を調べ，頻度の高いものを調べた(表 1). 上位 8 個の記号を選択して攻撃特徴文字とする．各攻撃文の攻撃特徴文字の含有率を計算し，8 次元特徴ベクトルを作成した．

表 1　記号の出現頻度

	文字	出現頻度
1	SP	0．269
2	/	0．180
3	.	0．114
4	|	0．073
5	-	0．057
6	:	0．049
7	"	0．049
8	;	0．045

6. 実験結果・考察

表 2 は SVM による OS コマンドインジェクション攻撃の検知実験の結果をまとめたものである．表 2 を見て分かる通り，いずれのカーネル関数を用いた場合でも概ね良い結果が出ているが，線形カーネルの場合の結果が最も良いことが分かる．これは，カーネル関数を利用しない場合が最も検知精度が良いということを示している．その理由について簡単に考察を行う．考えられる理由の 1 つとして，OS コマンドインジェクション攻撃の文字列の構造が比較的単純であることが挙げられる．本研究で用意した正常文は多くの記号が含まれるものを選んだが，実験に用いた正常文の方が多種多様な記号が含まれている．このような正常文の特徴に多項式カーネルやガウスカーネルは影響を受けたものと考えられる．もう 1 つ別の理由としては，本研究で用意できたデータの個数はまだ十分ではないとも考えられるため，攻撃文・正常文ともにデータを増やして実験を行う必要性があるとも考えられる．

今回の結果から OS コマンドインジェクション攻撃は機械学習アルゴリズムによる検知はある程度有効的に働くものと考えられる．今後は，未知のデータに対して精度の高い検知が可能になるよう, SVM のパラメタのチューニングや特徴抽出の改善をしていきたい．

表 2　SVM での検知結果(C=1000，γ=0.1)

		Accuracy	Precision	Recall	F-measure
線形カーネル	攻撃	98%	99%	99%	99%
	正常		99%	98%	99%
多項式カーネル	攻撃	96%	93%	100%	97%
	正常		100%	93%	96%
ガウスカーネル	攻撃	96%	95%	98%	96%
	正常		98%	96%	96%

Accuracy (正解率):予測結果と答えが一致する割合.

Precision (精度):真であると予測したデータのうち実際に真であるものの割合.

Recall (再現率):実際に真であり，予測結果が死んであるものの割合.

F-measure :Precision と Recall の調和平均.

参考文献

[1] 園田道夫 松田健 小泉大城 趙晋輝. “主成分分析を用いた分類器による SQL インジェクション攻撃の自動検出法” FIT2013

[2] Ryohei Komiya Incheon Paik Masayuki Hisada, “Classification of Malicious Web Code by Machine Learning”

[3] IPA “安全なウェブサイト運営のための WAF”(2016/11/12)

[4] 栗田 多喜夫 “サポートベクターマシン入門” http://home.hiroshima-u.ac.jp/tkurita/lecture/svm.pdf (2016/11/12)

[5] 梅原 章宏 松田 健 園田 道夫 水野 信也 趙 晋輝, “クロスサイトスクリプティング(XSS)攻撃における バッチ学習とオンライン学習の比較実験” 情報処理学会第 78 回全国大会