A-020

線形分類器によるクロスサイトスクリプティング(XSS)攻撃の検知に関する考察

Consideration on the Cross-Site Scripting Attacks Detection Using Linear Classifiers

梅原 章宏 [1]　松田 健 [2]　園田 道夫 [1]　水野 信也 [2]　趙 晋輝 [1,3]
Akihiro Umehara　Takeshi Matsuda　Michio Sonoda　Shinya Mizuno　Jinhui Chao

1. はじめに

近年インターネットの普及により,Web サイトを日常的に閲覧することが多くなった.それに伴い,Web サイトに入力された個人情報を狙うサイバー攻撃も増加している.

クロスサイトスクリプティング(XSS)攻撃は,Web サイトの入力部分の脆弱性に対して攻撃を行うサイバー攻撃の一種である.従来の対策として構文解析によるフィルタが提案・実現されている[1]が,XSS 攻撃に用いられる入力は,正常な入力との区別が容易でなく,機械的な攻撃検知が難しい.

先行研究[2]では入力に特徴抽出を行い,生成した 3 次元ベクトルを用いて攻撃の検知を試みた.その結果ある程度の攻撃検知が可能であったが,より攻撃と正常の入力を分離できるような特徴抽出法を用いることで性能の向上が期待された.

本研究では,XSS 攻撃に含まれる記号に着目して,攻撃と正常の入力に対して新たな特徴抽出の手法を適用した.また生成したデータを線形分類器であるサポートベクターマシン(SVM),Soft-Confidence Weighted Learning(SCW)を用いて分類を行い,攻撃検出への影響を考察した.

2. XSS 攻撃

XSS 攻撃による主な被害として,Cookie 値を盗まれることによる成りすまし被害や,Web ページの改ざんが挙げられる.例えば通信販売サイトのマイページに XSS 攻撃に対する脆弱性が存在する場合,スクリプトを用いてセッション ID を奪うことで,第三者が不正にログインすることが可能となる.これにより,クレジットカードや住所などの個人情報が流出し,悪用される恐れがある.

既存の検知手法として,特定の入力を拒否するブラックリスト方式,特定の入力のみ許可するホワイトリスト方式,'<'や'&'といった HTML における特別な記号を別の文字記号に置き換えるエスケープ処理がある.

3. 線形分類器

3.1.1 SVM

SVM は線形分類器の一つであり,分類は以下の式で定義される[3].

$$y(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) + b = \sum_{n=1}^{N} a_n t_n k(\mathbf{x}, \mathbf{x}_n) + b$$

ここで$\mathbf{x}$は入力ベクトル,$\mathbf{w}$は重みベクトル,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,$k(\mathbf{x}, \mathbf{x}_n)$はカーネル関数,bはバイアスパラメータである.SVM は分類境界と最も近いデータとの距離(マージン)を最大化することで,汎化誤差を最小化するような分類境界を求める.

マージンを最適化する解は以下の目的関数を最小化することで得られる.

$$C\sum_{n=1}^{N} \xi_n + \frac{1}{2}\|\mathbf{w}\|^2$$
$$\text{s.t.} \quad t_n(y(\mathbf{x}_n)) \geq 1 - \xi_n, \quad n = 1, \ldots, N$$
$$\xi_n \geq 0$$

ここでCは誤分類に対するペナルティの大きさを制御するパラメータであり,大きいほど誤分類を許さない分類境界を求める.またξ_nはスラック変数であり,$0 \leq \xi_n \leq 1$となるデータは正しく分類,$\xi_n > 1$となるデータは誤分類されている.

3.1.2 SCW[4]

SCW は逐次学習型の線形分類器の一つであり,分類は以下の式で定義される.

$$y(\mathbf{x}) = \mathrm{sgn}(\boldsymbol{\mu}_{t-1} \cdot \mathbf{x}_t)$$
$$\text{if } \boldsymbol{\mu}\mathbf{x} \geq \mathbf{0}: \; y(\mathbf{x}) = 1$$
$$\text{else} \quad : \; y(\mathbf{x}) = -1$$

ここで$\mathbf{x}$は入力ベクトル,$\boldsymbol{\mu}$は重みの平均ベクトルである.SVM と異なりバイアスパラメータは存在しない.また,損失関数l^{ϕ}は以下の式であらわされる.

$$l^{\phi} = \max\left(0, \phi\sqrt{\mathbf{x}_t^{\mathrm{T}}\Sigma\mathbf{x}_t} - y_t\boldsymbol{\mu}\cdot\mathbf{x}\right)$$

ここでΣは共分散行列,$\phi = \Phi^{-1}(\eta)$である(Φは正規分布の累積密度関数,ηは誤差を許容する程度をあらわすパラメータである).

$\boldsymbol{\mu}, \Sigma$の更新式は,以下の最適化問題であらわされる.

$$(\boldsymbol{\mu}_{t+1}, \Sigma_{t+1}) = \arg\min_{\boldsymbol{\mu},\Sigma} D_{\mathrm{KL}}(\mathcal{N}(\boldsymbol{\mu}, \Sigma)||\mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)) + Cl^{\phi}\left(\mathcal{N}(\boldsymbol{\mu}, \Sigma); (\mathbf{x}_t, y_t)\right)$$

ここでD_{KL}はカルバック情報量,$\mathcal{N}$は平均ベクトルμ,共分散行列Σの多変量正規分布,Cは重みの更新を制御するパラメータである.

SCW は同じ逐次学習型線形分類器である CW[5]を改良したものであり,特徴としてデータの信頼度による重み付けを行うことが挙げられる.また,マージンの最大化を行うことで,CW の弱点であるノイズに弱い部分を克服している.

4. 線形分類器を用いた実験

4.1.1 使用するデータの作成

実験を行うにあたり,文献[6]や Web ページ[7]などを利用し,攻撃入力 915 個,正常入力 1290 個を用意した.この中から攻撃入力 500 個,正常入力 500 個をランダムに抽出し,攻撃と正常の入力の入り混じったデータ数 1000 個のデータセットの生成を行った.

4.1.2 特徴抽出

生成したデータセットに対して特徴抽出を行い,4 次元の特徴ベクトルに変換し,それを入力として用いた.特徴抽出は,ASCII コードに基づいて表 1 の通りに行った.

1 中央大学大学院理工学研究科情報工学専攻
Departments of Information and System Engineering, Graduate School of Science and Engineering, Chuo University
2 静岡理工科大学総合情報学部コンピュータシステム学科
Department of Computer Science, Faculty of Comprehensive informatics, Shizuoka Institute of Science and Technology
3 中央大学理工学部情報工学科
Departments of Information and System Engineering, Faculty of Science and Engineering, Chuo University

表 1 特徴ベクトルの生成

特徴ベクトル$\mathbf{x} = (x_1, x_2, x_3, x_4)$

x_1									x_2		x_3	x_4
SP	!	"	#	$	%	&	'	(	:	;	数字 0~9	文字 a~z, A~Z
)	*	+	,	-	.	/	[	\\	<	=		
]	^	_	`	{	|	}	~		>	?		
その他制御文字									@			

4.1.3 線形分類器

実験に用いた線形分類器は SVM,SCW の 2 種類である.SVM の実行プログラムには,Python の機械学習ライブラリ scikit-learn0.15.2 の関数 SVC を使用した.以下に示した SVM のカーネル関数もこれに準拠するものである.

線形カーネル:$< x, x' >$

ガウスカーネル:$\exp\left(\frac{1}{2\sigma^2}|x - x'|^2\right)$

SCW の実行プログラムは図 1 のアルゴリズムを参考に Python で実装した.Python のバージョンは 2.7.8 である.

4.1.4 評価

評価を行うにあたって,特徴抽出を行ったデータセットを 5 分割し,交差確認を行った.SVM の評価項目として,以下のものを用いた.

1. データ全体に対して予測が正しかったものの割合(正解率,Accuracy)
2. 予測が実際に正しいものの割合(精度,Precision)
3. 真の結果に対してその結果であると予測されたものの割合(再現率,Recall)
4. 精度と再現率の調和平均(F 値,F-measure)
5. ROC 曲線の下の面積(AUC)

SCW の評価項目としては,上記項目の 1~4 を用いた.

5. 結果・考察

2 つのデータセットに対して SVM を用いた分類を行った結果を表 2 に示す.ここでペナルティとマージンのトレードオフを制御するパラメータ$C = 100$(Cが大きいほど攻撃と正常を厳密に分ける分類境界を求める),ガウスカーネルのパラメータ$\sigma = 10$である.

表 2 SVM による分類結果

	線形カーネル		ガウスカーネル	
	データ 1	データ 2	データ 1	データ 2
正解率	0.926	0.939	0.950	0.953
攻撃精度	0.962	0.971	0.961	0.946
正常精度	0.896	0.912	0.959	0.962
攻撃再現率	0.888	0.906	0.960	0.962
正常再現率	0.965	0.972	0.941	0.944
攻撃 F 値	0.923	0.937	0.950	0.953
正常 F 値	0.929	0.941	0.950	0.953
AUC	0.926	0.939	0.951	0.953

また,データセット 1 について特徴ベクトルを 4.1.2 で示した方法で生成した時の線形・ガウスカーネルを用いた場合と,先行研究[2]で生成した$\mathbf{x} = (x_1 + x_2, x_3, x_4)$であらわされる特徴ベクトルを用いた場合の SVM の分類結果を比較したグラフを図 1 に示す.

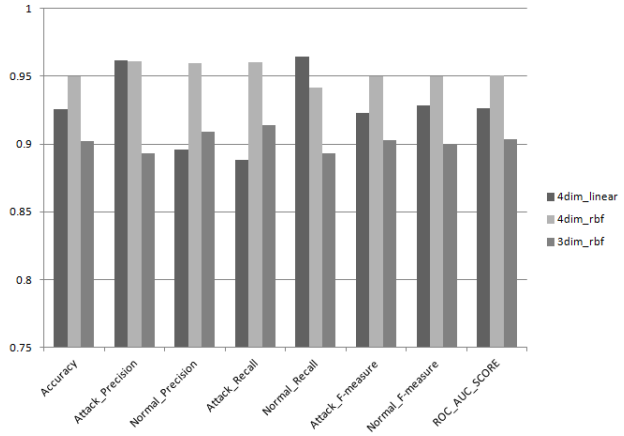

図 1 SVM による分類結果

表 2 の正解率をみると,線形カーネルで約 92%,ガウスカーネルで約 95%となり,高い精度でデータを分離できていることがわかる.また,両カーネルでの正常精度の差から,ガウスカーネルの方が正常入力を攻撃入力と誤検知する回数が少ないといえる.さらに図 1 より,3 次元特徴ベクトルと 4 次元特徴ベクトルを比べた際に,ほとんどの値について 4 次元ベクトルの結果が 3 次元ベクトルの結果を上回っていることから,提案した特徴抽出法は有効であると考えられる.

次に,2 つの 4 次元特徴ベクトルからなるデータセット 1,2 と,先行研究[2]で生成した 3 次元特徴ベクトルからなるデータセット 1 に対して,SCW を用いた分類を行った結果を表 3 に示す.ここでパラメータ$C = 100, \eta = 0.5$とする.

表 3 SCW による分類結果

	データ 1	データ 2	3 次元
正解率	0.690	0.764	0.53
攻撃精度	0.783	0.790	0.496
正常精度	0.593	0.738	0.557
攻撃再現率	0.665	0.751	0.529
正常再現率	0.738	0.759	0.526
攻撃 F 値	0.715	0.769	0.511
正常 F 値	0.652	0.757	0.54

こちらの場合も,3 次元特徴ベクトルと比較して 4 次元ベクトルの方が全体的に結果が向上していることから,提案した特徴抽出により,攻撃と正常の入力を分離可能な特徴空間へ写像することができたといえる.ただし,SVM と比較した際には結果が劣っているため,より効果的な特徴抽出の方法や,アルゴリズムの改良が課題に挙げられる.

参考文献

[1] "IE8 Security Part Ⅳ: The XSS Filter - IEBlog - SiteHome - MSDN Blogs",http://blogs.msdn.com/b/ie/archive/2008/07/02/ie8-security-part-ⅳ-the-xss-filter.aspx,最終閲覧日:2015/1/7

[2] 梅原 章宏 松田 健 園田 道夫 趙 晋輝,"線形分類器によるクロスサイトスクリプティング(XSS)の検知に関する考察",情報処理学会第 77 回全国大会 学生セッション 6S-05

[3] Christopher M.Bishop,"PATTERN RECOGNITION AND MACHINE LEARNING",Springer(2006)

[4] Jialei Wang,Peilin Zhao,Steven C.H. Hoi,"Exact Soft Confidence-Weighted Learning",ICML(2012)

[5] Mark Dredze,Koby Crammer,Fernando Pereira,"Confidence-weighted Linear Classification",ICML,pp.264-pp.271(2008)

[6] Jeremiah Grossman,Robert "Rsnake" hansen,Petko "pdp" D.petkov,Anton Rager,Seth Fogie,"XSS ATTACKS",SYNGRESS(2007)

[7] "ページ一覧取得",http://tshinobu.com/lab/get-page-link/,最終閲覧日:2015/1/4