[Work in Progress] 研究報告

ページ内リンクを用いたフィッシングサイト検知手法の提案

北山 直洋[1]　今泉 貴史[2]

Phishing site detection based on links in a web page

1. はじめに

昨今フィッシング行為が増加，深刻化している．一般的な対策として，ブラックリスト方式や，コンテンツベース方式 [1] があるが，前者はリストの管理コストが大きく，後者は検知時間が長いという欠点があり，十分な対策とは言い難い．そこで本研究では，既存手法の欠点を解消したフィッシングサイト検知手法を提案する．

2. 提案手法

本研究では，フィッシングサイトは SSL 通信を使用しないという特徴と，同ドメインの web ページが少ないという特徴を利用して判定する．以下に示す条件のどちらかに当てはまれば正規サイトと判定し，どちらも当てはまらない場合フィッシングサイトと判定する．

- **SSL 通信を用いた判定**
  検知対象ページが SSL 通信使用の場合，あるいは，検知対象ページのリンク先に SSL 通信使用ページが存在し，ドメイン比較判定に成功した場合
- **総リンクを用いた判定**
  検知対象ページのリンク先のすべてのページに対しドメイン比較判定を行い，成功率，成功数が予め設定していた閾値より共に大きい場合

図 1　ドメイン比較

表 1　実験 1 結果

	検知数	検知時間		SSL	総リンク
		平均	最大	判定数	判定数
正規	20	2.68	7.3	18	2
フィッシング	19	2.9	6.8	1	0

ドメイン比較は，ドメインレベル数が 2 以下の場合はリンクページの全ドメイン，3 以上の場合は最下位レベルドメインを省いた残りのドメインが検知対象ページのドメインに含まれていれば成功とする．また，判定順は検知時間削減のため，先に SSL 通信を用いた判定から行う．

3. 実験

フィッシングサイトとして，PhishTank[2] から 2015 年 7 月 8 日時点で最新の 20 件，正規サイトとして，それらの模倣元のサイトのトップページ 20 件を実験対象とした．また，検知対象ページにアクセスしてから検知結果が出るまでの時間を検知時間とした．閾値は以前の実験データから成功数:7，成功率:0.3 が適切だと判断したので，本実験ではこの閾値を使用する．

実験結果を表 1 に示す．SSL 判定数は SSL 通信の調査のみで，総リンク判定数は総リンクの調査で正規サイトと検知した数である．

4. 考察

SSL 通信を使用していることが原因でフィッシングサイトを 1 件誤検知しているが，本件は非常に稀である (1%未満)．今後の課題として，検知率，検知時間の改善，より実環境に即した大規模な実験が挙げられる．

参考文献

[1] 中山 心太, 吉浦 裕. 模倣コンテンツの特性に基づくフィッシング検知方式. 電気通信大学電気通信学部人間コミュニケーション学科. 2007.
[2] ”PhishTank — Join the fight against phishing.” http://www.phishtank.com/. (参照 2015-02-12)

1 千葉大学大学院融合科学研究科
Graduate School of Advanced Integration Science, ChibaUniversity
2 千葉大学統合情報センター
Institute of Management and Information Technologies, ChibaUniversity