4X-02

通信遷移と PageRank を用いた悪性リダイレクト防止手法

佐藤 祐磨† 中村 嘉隆† 高橋 修†

†公立はこだて未来大学 システム情報科学部

1 はじめに

ドライブバイダウンロード攻撃は，マルウェア感染攻撃の総称である．ユーザが Web ページにアクセスした際，ユーザの意図にかかわらず，悪意あるソフトウェアをダウンロードさせる攻撃である[1]．ドライブバイダウンロード攻撃の流れは，攻撃者が，良性 Web サイトの Web ページを改ざんする．改ざんの目的は，良性 Web ページから攻撃者が用意する攻撃 Web サイトへのリダイレクトである．改ざんされた Web ページにアクセスしたユーザは，改ざんのリダイレクト処理により，攻撃者が用意した攻撃サイトへ誘導させられる．攻撃サイトでは，ユーザの使用する OS，ブラウザ，ブラウザのアドオンの脆弱性を突く攻撃が行われ，ユーザの制御が攻撃者に奪われる．その後，ユーザはマルウェア配布サイトへ誘導され，悪意あるソフトウェアをダウンロードさせられる．

本研究では，ドライブバイダウンロード攻撃で発生するリダイレクトを防止することで既存手法よりも正確にマルウェアのダウンロードを防ぐことが目的である．

2 関連研究

(1)難読化スクリプトコード解析

ドライブバイダウンロード攻撃では，難読化スクリプトコードを利用し，第三者がスクリプトコードの挙動を簡単に解析できないようにされていることが多い．この難読化スクリプトコードを静的，動的に解析する手法が提案されている[2]．しかし，未知の難読化パターンに対応できない場合，攻撃を正確に検知できないという問題がある．

(2)ブラックリスト

ユーザがブラックリストに載る URL にアクセスした際，通信を遮断，警告を出し攻撃を防ぐ手法がある[3]．しかし，近年のドライブバイダウンロード攻撃では，Web ページにアクセスしたユーザのブラウザ情報を識別し，その情報によってリダイレクトする URL やスクリプトコードの挙動を変更する．このような攻撃に対して，ブラックリストでは，攻撃を検知，対策することは難しい．

(3)HTTP 通信解析

HTTP ヘッダ，IP アドレス，ドメイン情報を利用して，悪性と考えられる HTTP 通信を検知し，攻撃を防ぐ手法がある．HTTP ヘッダに含まれる Referer，Location 情報を利用し，ユーザがアクセスした Web ページの階層をカウントし，攻撃と考えられる通信を遮断する[4]．しかし，HTTP ヘッダ情報の一部は改ざん可能であり，現在の手法では，攻撃を正確に検知することができない．

(1)～(3)までで挙げた既存手法では，未知の難読化スクリプトコードパターンに対応できない場合に攻撃を正確に検知できない問題や，ブラックリストに登録されていない Web ページにアクセスした場合，攻撃を防ぐことができない場合が考えられる．また，現在の HTTP 通信解析手法は，リダイレクトによって取得される実行ファイルが良性であっても，誤検知によって通信が遮断されてしまう問題がある．

3 提案手法

本研究では通信遷移と PageRank を用いて，悪性リダイレクトを防止する手法を提案する．

3.1 基本的な考え方

PageRank は，ある論文の重要性は他の論文からの引用数によって評価されるという学術論文の考えを Web に適用したものである[5]．PageRank は Google から見た Web ページの重要度であり，0 から 10 の 11 段階でランク付けされる．PageRank の値が高ければ高いほど，Google から見た Web ページの重要度は高い．本研究では，この PageRank を利用し，ユーザがアクセスする Web ページの信頼度の指標とし，PageRank で評価されない Web サイトとの通信を遮断して，攻撃サイト・マルウェア配布サイトへの悪性リダイレクトを防止する．

3.2 提案手法アルゴリズム

ユーザのクリックまたは URL バーに入力した Web ページの階層を 1 とし，その Web ページが読み込む Web ページの階層を Referer，Location 情報を利用してカウントする．階層が 4 以上の通信に対して，リクエスト URL の完全修飾ドメイン名 FQDN の PageRank を取得する． PageRank が 0 または存在しない場合は，

“A method of preventing the malicious redirection of Web sites by transition of HTTP communications and PageRank status”
Yuma Sato†, Yoshitaka Nakamura†, Osamu Takahashi†
†School of Systems Information Science, Future University Hakodate

悪性リダイレクトとみなし，通信を遮断する．また，HTTP ヘッダは書き換えが可能なので，Referer 情報が存在しない場合もある．このときはリクエスト URL に含まれる FQDN の PageRank を取得し PageRank が 0 または存在しない場合は同様に悪性リダイレクトとみなす．さらに階層が 2 以上で Content-Type が pdf を含む Web ページ，Web ブラウザが自動で読み込む Web ページの Content-Type が octet-stream，x-msdownload，x-msdos-program，x-download のいずれかを含みかつ階層 1 の FQDN と異なる Web ページの FQDN の PageRank を取得し悪性リダイレクトの判別を行う．悪性 Web サイトの生存期間は 1 日のものが多いとわかっている[6]．生存期間が短い Web サイトに PageRank は存在しないため，PageRank に評価されない Web サイトとの通信を遮断することで悪性リダイレクトを防ぐ．

表 1 に各手法の特徴を示す．提案手法は，既存手法の問題をすべて解決できると考えられる．

表 1. 提案手法と既存手法の比較

	未知の攻撃	ブラウザ情報識別によるリダイレクト	検知精度
関連研究(1)	×	○	△*1
関連研究(2)	×	×	△*1
関連研究(3)	○	△*2	×
提案手法	○	○	○

1 未知の攻撃に非対応，2 検知率が低い

4 実験と評価

実験では，HTTP 通信をキャプチャした通信データに提案手法を適用する．

4.1 実験用通信データ

良性データとして Alexa が提供する HTTP 通信の上位 100 件を利用する[7]．また悪性データは NTT セキュアプラットフォーム研究所が提供する D3M2014 データセットを利用する[8]．

4.2 評価

本研究の評価項目は以下で行う．

- 真陽性率：攻撃が発生したセッションに対して，リダイレクト先の攻撃を未然に防いだ割合
- 偽陰性率：攻撃が発生したセッションに対して，リダイレクト先の攻撃を防げなかった割合
- 真陰性率：良性 Web ページを良性 Web ページとみなした場合
- 偽陽性率：良性 Web ページを悪性リダイレクトとみなした場合

5 実験結果と考察

実験結果を表 2 に示す．提案手法の真陽性率は 98.57%，偽陰性率は 1.43%，真陰性率は 98.38%，偽陽性率は 1.62%となった．提案手法は，関連研究(3)の既存手法よりも真陽性と偽陰性の評価項目で攻撃検知率が 22.35%上回った．Referer 情報が存在しない Web ページの FQDN の PageRank を取得し攻撃を判別することで既存手法よりも攻撃検知が向上したと考えられる．

表 2. 提案手法と既存手法の検知率

	真陽性	偽陰性	真陰性	偽陽性
提案手法	98.57%	1.43%	98.38%	1.62%
既存手法	76.22%	23.78%	98.05%	1.95%

6 おわりに

本研究では，ドライブバイダウンロード攻撃の既存検知手法よりも検知率の向上を目的とし，通信遷移の特徴と Google の PageRank を利用して，悪性リダイレクトを防止する手法を提案した．提案手法は，既存手法よりも，正確な検知率であり，提案手法の有効性を示した．

今後は，ユーザが直接攻撃 Web ページにアクセスした際，攻撃を防ぐ手法を提案したいと考えている．

参考文献

[1] “IPA 独立行政法人 情報処理推進機構：コンピュータウイルス・不正アクセスの届出状況[2010 年 11 月分]について”, https://www.ipa.go.jp/security/txt/2010/12outline.html

[2] 神薗雅紀, 西田雅太, 星澤裕二, "動的解析を利用した難読化 JavaScript コード解析システムの実装と評価." 情報処理学会シンポジウム論文集, Vol.2010, No.9, pp.453-458, 2010.

[3] Luca Invernizzi, Stefano Benvenuti, Marco Cova, Paolo Milani Comparetti, Christopher Kruegel, Giovanni Vigna “EvilSeed: A Guided Approach to Finding Malicious Web Pages, ”Proceedings of the 2012 IEEE Symposium on Security and Privacy (SP'12), pp.428-442, 2012.

[4] 安藤慎悟, 寺田真敏, 菊池浩明, 趙晋輝, "通信の遷移に着目した不正リダイレクトの検出による悪性 Web サイト検知システムの提案, " 情報処理学会研究報告, Vol.2011-CSEC-54, No.32, pp.1-6, 2011.

[5] Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd, "The PageRank Citation Ranking: Bringing Order to the Web," Manuscript in progress.

[6] 秋山満昭, 八木毅, 針生剛男, "改ざん Web サイトのリダイレクトに基づく悪性 Web サイトの生存期間測定," 電子情報通信学会技術報告, Vol.113, No.502, pp.53-58, 2014.

[7] "Alexa - Actionable Analytics for the Web", http://www.alexa.com/

[8] 秋山満昭, 神薗雅紀, 松木隆宏, 畑田充弘, “マルウェア対策のための研究用データセット~MWS Datasets 2014~,” 情報処理学会研究報告, Vol.2014, No.19, pp.1-7, 2014.