2W-05

HTTP リクエストシーケンスに注目した Drive-by Download 検知手法

工藤　聖†　　トラン・コン・マン†　　中村　康弘†

†防衛大学校

1　はじめに

不正リダイレクトによる Drive-by Download（以下「DbD」という.）等，Web アクセスを用いたサイバー攻撃が活発に行われている [3]．組織ネットワークに対し DbD のようなマルウェア送付攻撃が行われる場合，攻撃サイトへのアクセス，マルウェアのダウンロード及び活動という一連の動作によって，特徴的な HTTP リクエストシーケンスが発生する．本研究では，このシーケンスを検出することにより，膨大なペイロード・データを解析することなく，ヘッダ情報のみで攻撃を検出する手法を提案する．D3M Dataset 及び組織の proxy ログ等を対象に，手法の妥当性と性能について検証する．

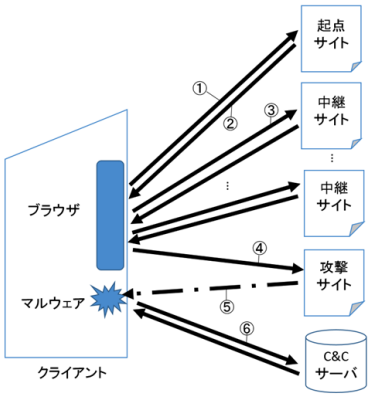

図 1: Drive-by Download 概要

2　Drive-by Download

DbD は，改竄された Web サイト等から不正なリダイレクトを経てマルウェアをダウンロードさせる攻撃である．概要を図 1 に示す．DbD 攻撃者は，既存サイトを改竄し，リダイレクトコードを挿入する．リダイレクト手法は，サーバの機能による 301 及び 302 リダイレクトの他，meta タグ，JavaScript コード,iframe タグ等が用いられている．クライアントは，記述されたタグやコードに従って中継サイトに誘導され（①〜③），複数の中継サイトを経て，攻撃サイトに誘導される（④）．攻撃サイトは, ブラウザや Flash Player 等のプラグインの脆弱性を利用し，ユーザが意図しない形でマルウェアをダウンロードさせる（⑤）．不正なコードを含んだ pdf 形式等のダウンローダが先行的にダウンロードされ，これがマルウェアの実行ファイルをダウンロードする場合もある．

ダウンロードされたマルウェアは自動実行され，窃取した情報の送信や，指示の取得を目的として，C&C サーバと通信を行う（⑥）．

3　先行研究

著者らは既報 [1] において，組織ネットワークの proxy サーバを監視し，DbD を検出する手法を提案した．提案手法を実装し，組織の proxy ログに適用した結果，DbD 候補を検出できることを確認した．しかし，実際の DbD 発生時のデータを保有していなかったため，実攻撃への適用可能性については検証できなかった．

4　提案手法

本章では，先行研究で提案した手法について述べる．

特定のクライアントの発する HTTP リクエストを時系列的に監視し，以下の 3 条件を満たした場合，一連のシーケンスが DbD の可能性ありと判断する．

条件 1　リダイレクトの発生

リダイレクトの発生をステータスコードにより判定する．レスポンスのステータスコードが 300 番台であれば，条件 1 が成立したと判定する．

条件 2　リダイレクト発生直後の特定タイプのファイルのダウンロード

リダイレクト発生直後に，実行可能形式の exe ファイルや，脆弱性が頻繁に悪用される swf ファイル，pdf ファイル等がダウンロードされた場合，攻撃の可能性があると判断する．攻撃サイトにリダイレクトされてから，マルウェアをダウンロードするリクエストが発生するまでの時間は極めて短時間であると考えられる．よって，条件 1 の成立から 1 秒以内に，URL の末尾拡張子が上記 3 種類に該当するリクエストが発生した場合，条件 2 が成立したと判定する．

Detection of Drive-by Download using HTTP request sequence
†SEI KUDO　†TRAN CONG MANH　†YASUHIRO NAKAMURA
†National Defense Academy

条件 3 **ダウンロードから一定時間以内に，ブラウザと異なるユーザエージェントからのリクエスト発生**

条件 1 及び条件 2 が成立してから，一定時間（本研究では 120 秒とした）以内に異なるユーザエージェントからのリクエストが発生した場合，条件 3 が成立したと判定する．

5 検証実験

実験は，txt 形式で記録されたログファイルを，提案手法を実装した perl スクリプトで読み出す方式とした．データセット等の pcap 形式のデータはログに準拠した txt ファイルに変換した．

5.1 D3M を用いた検証

D3M Dataset 2015[2] を用いて検証を行ったが，攻撃を検出できなかった．

D3M Dataset はハニーポットによる巡回とサンドボックスによるマルウェア通信を別個に実行している．このように，形態が実際の DbD と異なるため，提案手法における条件 2 と条件 3 の関連付けができず，検出できなかった．

5.2 模擬 DbD による検証

模擬 DbD サイト及び模擬マルウェアを作成し，これらが動作する際の通信ログを記録して検証を行った．概要を図 2 に示す．

模擬 DbD サイトは，攻撃サイトへの転送をディレクトリ間の 302 リダイレクトにより模擬する．攻撃サイトにおける，脆弱性を悪用した意図しないダウンロードは，exe ファイルダウンロード時に表示される警告を承認することにより模擬した．

模擬マルウェアは，C 言語で作成した exe ファイル形式の HTTP クライアントであり，あらかじめ指定された Web サーバに HTTP リクエストを送信する．ダウンロードされた模擬マルウェアの自動実行は，手動で実行することにより模擬した．

実験の結果，提案手法によりシーケンスを検出できることが確認できた．

5.3 組織 proxy ログを用いた検証

組織の proxy ログを用いて検証を行った結果，3108 件の DbD 候補が検出された．ログ期間は 2015 年 10 月 1 日〜3 日，ユーザ数は約 2000 人，ログ行数は約 300 万行である．

検出結果を確認したところ，正規のリダイレクト及び特定ファイルのダウンロードの後，OS やセキュリティ

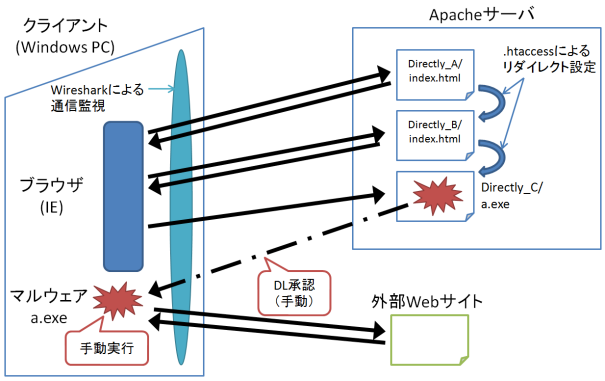

図 2: 模擬 DbD による検証実験

ソフトのアップデートプログラムによる通信が生起したことにより，検出されたものが多数見られた．検出結果に実際の DbD の履歴は含まれておらず，検出されたシーケンスはすべて False Positive であった．

6 まとめ

リダイレクトを利用した DbD を検出できることが確認できた．

False Positive は平均して 83 秒に 1 件発生しており，実用性を考えた場合さらなる低減を図ることが望ましい．今後は，ユーザエージェント文字列の特徴解析等，正常通信を識別する方法について検討する．

D3M については，データセットの実用性をさらに高めるため，収集方法や保存形式について改善を行っていく必要があると考える．

また，提案手法ではリダイレクトの検出にステータスコードを用いたが，この方法では iframe による外部ページ読み込み等，ステータスコードが 300 番台にならない場合は検出できない．様々なリダイレクトに対応できる手法を開発し，提案手法に組み込んでいく必要がある．

参考文献

[1] 工藤 聖, 他. HTTP リクエストシーケンスに注目した不正リダイレクトの検出． Computer Security Symposium 2015,2A1-1 (2015)

[2] 神薗 雅紀, 他． マルウェア対策のための研究用データセット〜MWS Datasets 2015〜. 研究報告セキュリティ心理学とトラスト（SPT）,2015-SPT-14(6),1-8 (2015)

[3] Internet Infrastructure Review, http://www.iij.ad.jp/company/development/report/iir/