5X-2

動的解析と連携する通信可視化による ドライブ・バイ・ダウンロード攻撃の解析支援

義則 隆之† 佐藤 両† 松井 拓也† 廣友 雅徳‡ 毛利 公美†† 神薗 雅紀‡‡ 白石 善明†

†名古屋工業大学 ‡佐賀大学 ††岐阜大学 ‡‡（株）セキュアブレイン

1. はじめに

ユーザによる Web サイトの閲覧を契機としてマルウェアを秘密裏にダウンロードさせるドライブ・バイ・ダウンロード攻撃（Drive-by-Download attack：以下 DBD 攻撃）の被害が拡大している．ユーザは攻撃の起点となる Web サイトにアクセスすると，一段もしくは複数段のリダイレクトを経て悪性 Web サイトへ誘導され，ブラウザやプラグインの脆弱性を突く攻撃コードによってマルウェアを自動で実行される．攻撃を受けたユーザはベンダーに対応を依頼するが，ベンダーが受理したブラウザやプラグインの脆弱性に関する届出のうち，65%は修正パッチを公開するまでに 45 日以上要するといわれている[1]．ベンダーの対応を待つ間にゼロデイ攻撃により被害が拡大する恐れがあるので，組織でも速やかに対策を講じることが望ましい．ブラウザやプラグインを無効にすると業務に支障を来すので，組織で保管する通信データから悪性 Web サイトを探してアクセスを制御する対策が有効である．したがって，DBD 攻撃の特徴をもとに通信データから悪性 Web サイトを特定することを考える．

解析者は以下のような手順で通信データを解析する．

Step1. 通信データから疑わしい通信を探す
Step2. 通信に含まれるファイルを抽出する
Step3. 抽出したファイルを解析する

ユーザが通信データを解析するにあたり，いくつかの問題点がある．Step1 では，通信データから疑わしい通信を探し出すのに相応の時間を要するという問題がある．Step1 を支援するために，サーバ型のハニーポットが収集したログに基づいた攻撃者の地理的分布を世界地図に可視化する手法[2]や，Gumblar に感染した端末を設置し，全ての通信データの送信元アドレスを Google Earth 上にマッピングし可視化を行う手法[3]が提案されている．Step2,3 では，攻撃コードには解析を妨害するために難読化されているものが多く，コードを静的に解析するだけでは攻撃の全容が掴めないという問題がある．Step2,3 を支援するために PDF 形式の未知のマルウェアを検知するための静的解析と動的解析を統合した検知手法[4]や，動的解析を利用し難読化された JavaScript を解析するシステム[5]が提案されている．通信データの可視化と攻撃コードの解析を一連のシステムとして結合すれば，より速やかな解析を実現し，有用な支援となると考えられる．

本稿では攻撃コードの動的解析と通信可視化を連携させた DBD 攻撃の解析支援システムを提案する．通信データを解析し，提案システムが通信データの解析支援に有用であることを確認する．

2. DBD 攻撃解析支援システム

提案システムのコンセプトを図 1 に示す．提案システムの通信可視化と動的解析の概要を以下に示す．

通信可視化：送信元アドレスと送信先アドレスを世界地図上にマッピングし，リダイレクトがあった場合には送信元 IP アドレスから送信先 IP アドレスに向かって有向辺を引くことで，ユーザは通信フロー（通信の一連の流れ）を直感的に把握できる．しかし，正規 Web サイトにも別の Web サイトへリダイレクトするものがあり，通信フローを世界地図上に可視化するだけでは，DBD 攻撃がどの通信フローに該当するのかわからない．DBD 攻撃に用いられる攻撃コードは解析を妨害する目的で難読化されていることが多いので，難読化されたコードを含む通信フローの有向辺とそうでない有向辺を異なる色で描画することで，ユーザが DBD 攻撃の通信フローに見当をつけることができると考える．そこで，有向辺を描画する際に，通信データに含まれるファイルを抽出し各ファイルに難読化されたコードが含まれていないか判定し，難読化されたコードが含まれていれば通信フローの有向辺とそうでない有向辺を色分けする．

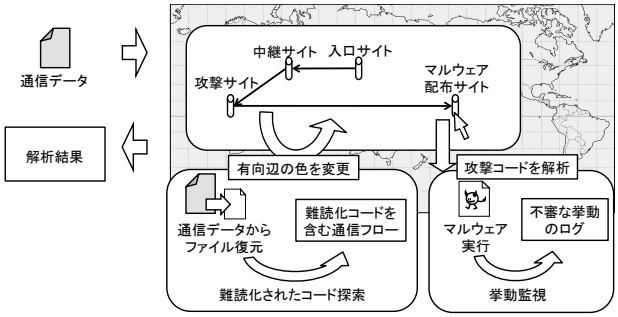

図 1 DBD 攻撃解析支援システムのコンセプト

攻撃コード動的解析：難読化されたコードを含む通信フローの有向辺をクリックすると当該通信フローに含まれるファイルを抽出し，自動で動的解析を行う．攻撃コードは一般に外部サーバからプログラムをダウンロードして，端末の管理者権限を奪取して任意のプログラムを実行する．そこで，外部との通信やプロセス生成などの挙動を監視し，それらの挙動を示したファイル名と Web サイトの IP アドレスをユーザに提示する．

3. DBD 攻撃解析支援システムの設計

3.1 構成モジュール

提案システムの構成を図 2 に示す．提案システムは以下の 3 つのモジュールで構成される．

［攻撃フロー特定支援部］ 通信データを解析し可視化インタフェースに有向辺を描画する．

［動的解析部］ 解析対象のファイルを実行し，アクセス先ドメイン名を取得する．プロセス生成やネットワーク通信の挙動を監視し，ファイルに悪意のあるコードが含まれているかを判定する．

［システム制御部］ 可視化インタフェースを通してユーザ入力を受け付け，攻撃フロー特定支援部と動的解析部を制御する．

3.2 解析の流れ

提案システムによる通信データの解析は以下の順に行われる．

通信可視化

Step1. 可視化インタフェース制御部が通信データの入力を受け付け，通信データを IP 国別割当部に渡す
Step2. IP 国別割当部は通信データに存在する IP アドレスを抽出し，国別 IPv4 割当リストと照合し各 IP アドレスがどの国に位置づけられるかを対応付ける
Step3. IP 国別割当部は Step2 の結果を可視化情報管理部に渡す
Step4. 可視化情報管理部は送信先 IP アドレスや送信元 IP アドレスに基づき一連のリダイレクトの流れをリダイレクト経由登録リストに記入し，通信フロー可視化部に渡す
Step5. 通信フロー可視化部はリダイレクト経由登録リストを参

A Drive-by-Download Attack Analysis Support System by Visualization Combined Dynamic Analysis

† Takayuki YOSHINORI, Ryo SATO, Takuya MATSUI and Yoshiaki SHIRAISHI・Nagoya Institute of Technology
‡ Masanori HIROTOMO・Saga University
†† Masami MOHRI・Gifu University
‡‡ Masaki KAMIZONO・Secure Brain Corp.

照し，可視化インタフェースに IP アドレスごとにポールを描画し，リダイレクトがあれば送信元 IP アドレスから送信先 IP アドレスに向かって有向辺を描画する

有向辺の色の変更

Step1. ファイル抽出部は通信データに含まれるすべてのファイルを抽出し，ファイル格納部に渡す

Step2. ファイル格納部は受け取ったファイルを対応する IP アドレスごとに分類し，送信元サーバのフォルダ構成と同一になるようにストレージに格納する

Step3. 難読化コード判定部はストレージ内の各ファイルに難読化されたコードが含まれていないか判定し，可視化情報管理部に判定結果を渡す

Step4. 可視化情報管理部は難読化されたコードを含む通信フローの有向辺をそうでない有向辺と異なる色に設定するよう通信フロー設定情報リストを修正し，通信フロー可視化部に渡す

Step5. 通信フロー可視化部は通信フロー設定情報リストを参照し，難読化されたコードを含む通信フローの色を変更する

攻撃コード解析

Step1. ユーザが可視化インタフェースに描画されたポールをクリックすると，可視化インタフェース制御部が動的解析制御部に攻撃コード解析を依頼する

Step2. 動的解析制御部はファイル実行部にファイルの実行を依頼し，挙動監視部に挙動の監視を依頼する

Step3. ファイル実行部はポールと対応する IP アドレスを取得し，ストレージから対応するフォルダを探索する．攻撃に用いられる PDF 形式のファイルなどが当該フォルダ内に存在すれば動的解析環境内にてリーダーやビューワーで開く

Step4. 挙動監視部はリーダーやビューワーの動作を監視し，監視結果を攻撃判定部に渡す

Step5. 攻撃判定部は監視結果からリーダーやビューワーが外部との通信やプロセスを生成していれば，ファイルに悪性コードが含まれていると判定し，判定結果を可視化インタフェース制御部に渡す

Step6. 可視化インタフェース制御部は判定結果を可視化インタフェースに出力する

4. プロトタイプの実装と評価

提案システムのプロトタイプを実装した．使用したライブラリとソフトウェアを表 1 に示す．動的解析環境を仮想マシンで構築し，外部への影響を考慮してネットワーク接続を切断する．ブラウザとプラグインには古いバージョンを適用し，脆弱性を有する環境を構築する．挙動監視部は Adobe Reader を監視し，ネットワーク通信やプロセス生成に関する API の呼び出しをフックする．攻撃判定部はポップアップで判定結果をユーザに提示する．

提案システムの有用性を確認するために，研究室の学生 11 名を被験者として，通信データの解析を実施した．通信データは D3M 2012（Drive-by-Download Data by Marionette 2012）[8]から抽出した難読化されたコードを含む通信に正規の通信を混在させて作成し，データサイズを約 10MB とした．被験者は(A)よく用いられる解析手法，(B)提案システムをそれぞれ用いて，悪性 Web サイトの IP アドレスを特定するのに要する時間を測定した．(A)には通信データ解析に Wireshark[9]，PDF タイプの攻撃コード解析には Wepawet[10]を用いた．被験者は各ツールの使用方法を事前に理解しているものとする．また，実験後にアンケートを行い，(A)と(B)の解析のしやすさについてそれぞれ 5 段階で評価した．

(A)，(B)について解析時間を測定した結果，(A)に要した時間の平均は 7 分 2 秒，(B)に要した時間の平均は 2 分 1 秒であった．(B)を利用した場合は(A)と比較して解析時間を 5 分程度短縮できた．(A)では，被験者の大半が疑わしい通信に見当を付けられず，通信データから PDF ファイルを一つずつ抽出し解析した．(B)では，被験者が難読化されたコードを含む通信フローとそうでない通信フローを判別でき，短時間で解析できた．

アンケートの集計結果を表 2 に示す．(B)は(A)と比較して解析しやすいと答えた被験者が多かった．これらの結果から，提案システムが通信データの解析支援に有用であるといえる．

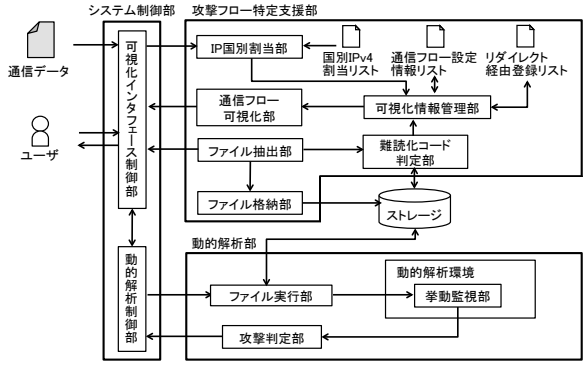

図 2　DBD 攻撃解析支援システムの構成

表 1　使用したライブラリとソフトウェア

実装環境		動的解析環境	
通信フロー可視化部	Java 3D 1.5.1	仮想環境	Vmware Workstation 8.0
IP国別割当部 ファイル復元部 ファイル格納部	jNetPcap 1.3.b4[6] （WinPcap[7]の Javaのラッパー）	OS	Windows XP SP2
挙動監視部	C++	ブラウザ	Internet Explorer 6.0
上記以外のモジュール	JDK 1.7.0_05	プラグイン	Adobe Reader，Flash Player，WinZip，QuickTime，JRE （全てセキュリティパッチ未適用）

表 2　アンケートの集計結果

アンケート項目	(A)の平均点	(B)の平均点
インタフェースが使いやすい	2.45	4.45
操作が簡単である	2.09	4.82
悪性サイトを特定するまでが簡単である	2.00	5.00
スムーズに解析できる	1.91	4.73

5.そう思う　4.ややそう思う　3.どちらともいえない
2.あまりそう思わない　1.そう思わない

5. まとめ

本稿では通信データの可視化と攻撃コードの動的解析を連携させた DBD 攻撃解析支援システムを提案した．提案システムは疑わしい通信の見当を付け，通信に含まれるファイルを動的解析する機能を持っており，悪性 Web サイトの特定を支援するものである．

提案システムのプロトタイプを用いたユーザ実験により，よく用いられている解析手法と比べて通信データの解析に要する時間を短縮できることを確認した．また，アンケートを実施したところ，解析のしやすさの観点で提案システムがより高い評価であった．

参考文献

[1] 情報処理推進機構：ソフトウェア等の脆弱性に関する届出の処理状況[2012 年第 3 四半期（7 月~9 月）]，独立行政法人 情報処理推進機構（2012）．

[2] Visoottiviseth, V., Jaralrungroj, U., Phoomrungraungsuk, E., and Kultanon, P.,: Distributed Honeypot Log Management and Visualization of Attacker Geographical Distribution, Proc. Computer Science and Software Engineering(JCSSE), IEEE Xplore DIGITAL LIBRARY, pp.23-28(2011).

[3] 金子 博一，松木 隆宏，新井 悠：通信トラフィックの分析による Gumblar 感染 PC の可視化，電子情報通信学会技術研究報告，情報通信システムセキュリティ（ICSS） 110（79），pp.1-6（2010）．

[4] Zacharias, T., Giorgos, S., Michalis, P., Evangelos, P.,: Combining static and dynamic analysis for the detection of malicious documents, Proc. the 4th European Workshop on System Security (EuroSec'11), No.4, DOI: 10.1145/1972551.1972555 (2011).

[5] 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，マルウェア対策人材育成ワークショップ （MWS2010），2A3-1（2010）．

[6] Riverbed Technology: WinPcap, WinPcap(online), available from<http://www.winpcap.org/>(accessed 2012-08-20).

[7] Sly Technologies: Sly Technologies jNetPcap, jNetpcap(online), available from<http://jnetpcap.com/>(accessed 2012-08-20).

[8] MWS2012 実行委員会：研究用データセット，MWS 2012 Datasets について（オンライン），入手先<http://www.iwsec.org/mws/2012/about.html#datasets>（参照 2012-12-10）．

[9] Riverbed Technology: Wireshark, Wireshark(online), available from<http://www.wireshark.org/>(accessed 2012-12-18).

[10] University of California: Wepawet, Wepawet(online), available from<http://wepawet.iseclab.org/>(accessed 2012-12-18).