セッショングラフ構造を用いた不審アクセスログの分析に関する検討

林　直樹[†1]　関口　竜也[†2]　仲小路　博史[†1]

概要：近年増加しているサイバー攻撃の一つとして DbD 攻撃がある．DbD 攻撃は標的を多段のリダイレクトを経て悪性コードの配布サイトまで誘導する攻撃フローが特徴である．攻撃者は，経路上のサーバや悪性コードの配布サイトを短期間の間に変更する場合があり，セキュリティの対策や監視を行う側がそのような全てのサイトの URL を知ることは困難であるため，DBD 攻撃の痕跡の調査を行うことが困難である問題がある．そこで我々は，プロキシサーバのログ内のリファラ情報を用いて通信セッションのグラフ構造を構築し，そのグラフ構造を元に不審通信を分析・調査する方式を検討した．

キーワード：サイバー攻撃，DbD 攻撃，プロキシサーバ

Analysis of Malicious Access Logs Using the Session Graph

Naoki Hayashi[†1]　Tatsuya Sekiguchi[†2]　Hirofumi Nakakoji[†1]

Abstract: . Recently, the number of DbD attack is rapidly increasing. One of the features of DbD attack is the attack flow where the attacker induces multiple redirects of the target's web browser and conduct the target to the malware distribution server. Attacker may change malicious servers in short span. So, it is difficult for the security operators to know them all and thus, it's hard for them to trace the whole attack. So, we developed a method of constructing session graph from proxy server's logs and analyzing it.

Keywords: Cyber Attack, DbD Attack, Proxy Server

1. はじめに

標的型攻撃等のサイバー攻撃の高度化により，組織の機密情報や顧客情報の漏洩といったセキュリティ事件や事故が増加・深刻化している．これらに対応するため，セキュリティ運用を専門的に行う組織である SOC(Security Operation Center)や CSIRT(Computer Security Incident Response Team)の分析業務も増加しているが，一方で，そのような分析を行うためのスキルを備えた人材の確保・育成が間に合っていない状況にあり，自動化等を行うことで分析業務の属人性を軽減させる必要が出てきている．

近年特に増加しているサイバー攻撃の一つとして，DbD (Drive-by-Download) 攻撃が挙げられる．DbD 攻撃は，攻撃対象のアクセス先をリダイレクトにより多段に遷移させてマルウェア配布サイトまで誘導する攻撃フローを特徴としている[1]．DbD 攻撃では，同一のマルウェアが様々な悪性サイト（マルウェア配布サーバ）を用いて配布されることがあるため，ある一つの悪性サイトへのアクセスをブロックするだけでは対策が不十分になることがある．例えば，保護対象組織の従業員がよく利用するようなウェブサイトが改ざんされ，攻撃者に準備された複数の悪性サイトにランダムに誘導される場合には，既知の悪性サイトをブロックするだけでなく，当該入り口サイトの管理者に改ざんの旨を通知することで復旧を促すことなども重要である．

すなわち，DbD 攻撃による侵害のログを調査する際には，どのような通信を経てマルウェア配布サイトまで誘導されたのか，全体像を把握することが，有効な対策を打つために必要である．

現状，多くの SIEM 等のログ分析システムにおいては，ログに対して単純な検索しかできないため，悪性サイトへの通信を記録したコネクションログ単体を見つけ出すことは容易であるが，それがどのような通信経路を辿ってきたものかを分析する機能が無く，多段のリダイレクトを含むような攻撃の全容を把握することが困難である課題がある．

そこで我々は，DbD 攻撃による不正通信ログが見つかった際の対応の分析支援を目的に，複数の通信先を遷移する攻撃の痕跡を可視化・分析する方式を検討し，アナリスト向けの可視化画面を設計した．

2 章では，まずセッショングラフの定義を述べる．3 章では，我々が設計した可視化方式を示す．4 章は評価であり，5 章と 6 章はそれぞれ考察とまとめである．

†1 株式会社日立製作所
Hitachi Ltd.
†1 株式会社日立システムズ
Hitachi Systems Ltd.

2. DbD 攻撃の特徴と既存研究

2.1 DbD 攻撃の特徴

DbD 攻撃において発生する典型的な通信フローは，図 1 の通りである．

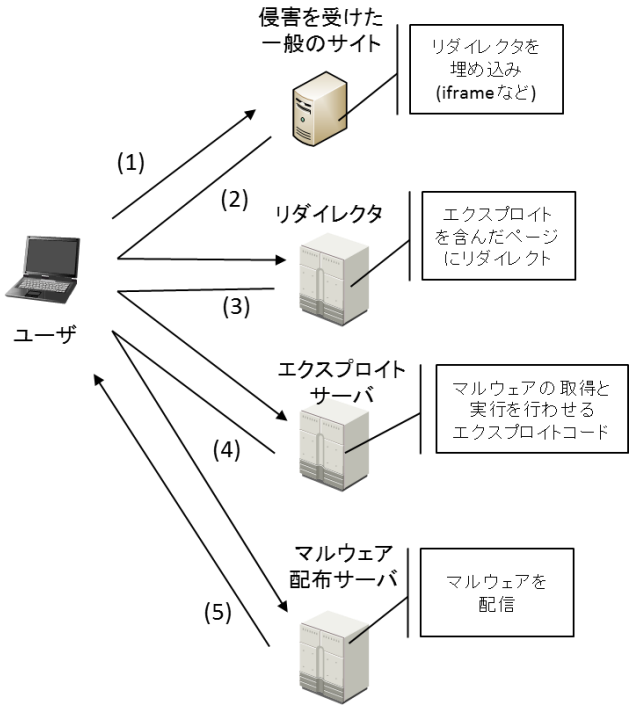

図 1　DbD 攻撃の典型的な通信フロー

Figure 1 Typical flow of DbD attack

(1) 攻撃者が一般のサイトに対し，エクスプロイトコードを配備したサーバへのリダイレクト命令を含んだコードを不可視の iframe 等で埋め込む．対象サイトが侵害を受けたことを知らないユーザは当該サイトにアクセスする．

(2) 対象サイトに埋め込まれたリダイレクト命令を含むコードをリダイレクタから読み込む．

(3) リダイレクタからエクスプロイトを実行するサーバへ転送される．

(4) エクスプロイトが実行されマルウェアがダウンロードされる．

(5) マルウェアが実行される．

リダイレクトは 1 回ではなく，複数回発生場合があることも知られている．

また，各サーバは，攻撃者によって複数準備されることもあり，時間の経過等で別のサーバを経由するように切り替わることもある．

SOC では，このような攻撃を検知して対策を行うため，保護対象組織にプロキシサーバや，IDS/IPS，アンチウィルスゲートウェイなどの機構を導入するとともに，それらの機器のログを取得して分析を行っている．

既存の代表的な DbD 攻撃対策手法としては，URL ブラックリスト方式が挙げられる．URL ブラックリストの代表的なものとしては，例えば，Google Safe Browsing [2]や Mcfee Site Advisor [3]のリストが現在運用されている．これらのブラックリストにマッチする通信を検知，あるいは遮断することで DbD 攻撃に対処することが可能である．しかしながら，これらのブラックリストには図 1 に示した DbD 攻撃において発生する全ての通信フローに利用されたサーバを含んでいるとは限らない．すなわち，図 1 に示した通信のうち，いずれかのみの警告が発せられることとなる．

DbD 攻撃では，攻撃者が図 1 に示した各役割のサーバを複数準備してそれぞれを頻繁に切り替えることがあるため，ある悪性な通信を検知した際，当該アクセス先を遮断したのみでは十分とは言えず，当該通信を発生させるにいたった一連の通信を調査して各サーバを遮断することが再発防止のために重要である．

そのような対処を行うためには，通信フローを把握するために多段の通信の相互の関係性を分析する必要がある．そのためには，セキュリティオペレータが通信ログを再帰的に調査することとなるためｓ，分析に掛かる工数が大きい．近年 DbD 攻撃の件数が増加していることも鑑みると，そのようなログ分析を効率化・自動化する技術が必要であると言える．

2.2 DbD 攻撃の分析と可視化に関する既存研究

DbD 攻撃通信の分析と可視化に関しては既存研究が存在する．

義則ら[4]が提案しているシステム "Flow Visualizer" は DvD 攻撃の通信を可視化するものである(図 2)．ハニーポットで収集した DbD 攻撃サイトの通信データを国別に分類し，世界地図上にプロットしている．また，通信頻度を元に通信データの色分けが行われている．

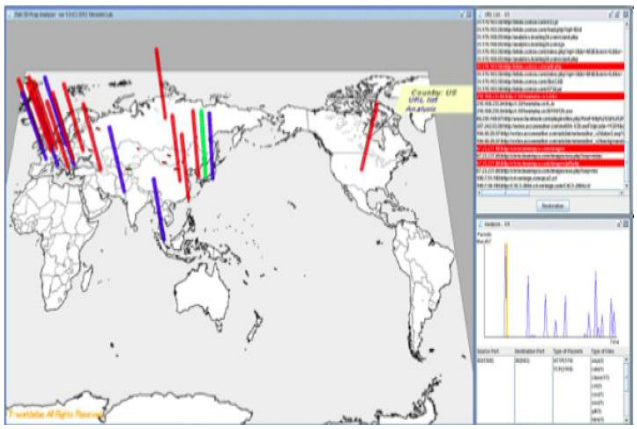

図 2　"Flow Visualizer"[4]

Figure 2 "Flow Visualizer"[4]

また，金子ら[5]は，Gumbler に感染した PC の挙動を可視化する可視化システムを提案している．当該提案では，動的解析を用いて Gumbler に感染した PC から発生する通信を観測し，観測結果を世界地図上にプロットする可視化を行っている．

尼子ら[6]は，オペレータに対して，DbD 攻撃が発生していることを認知させることを主眼に，DbD 攻撃に特徴的なリダイレクト通信を，横軸に発生時間，縦軸に IP アドレスを設定した平面状にプロットするシステムを提案している(図 3).

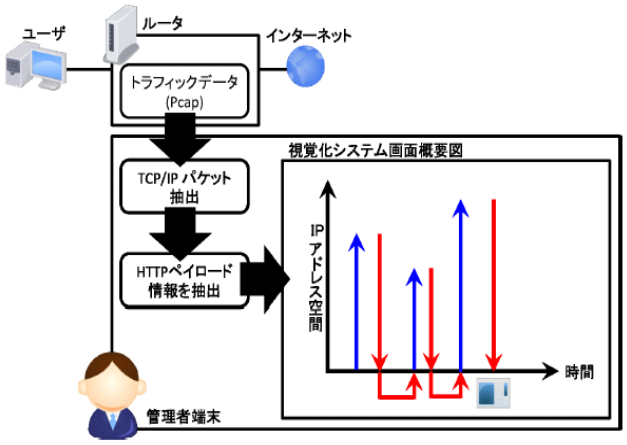

図 3 尼子ら提案の攻撃認知支援システム[6]

Figure 3 Attack detection tool by Amako et.al.[6]

また，松本ら[7]は，GeoLocation サービスを利用して IP アドレスの国名を解決して，発生した通信を世界地図上に描画し，かつ，リダイレクトが発生した通信については色を変更する可視化システムを提案している(図 4).

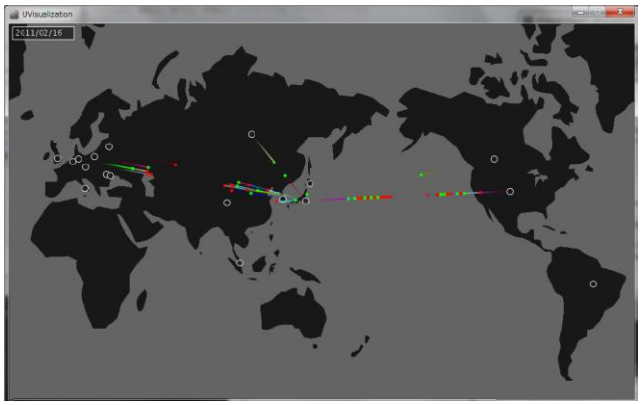

図 4 DbD 攻撃の可視化実行例[7]

Figure 4 Visualization System of DbD Attack[7]

3. 提案の可視化方式

前章で述べたように，DbD 攻撃のような多段のリダイレクトを含む攻撃を検知した場合，単一の通信ログだけではなく，当該通信が発生するにいたった通信フローを把握することが対策のために重要である．

本検討では，不審通信の警告が発生した際，当該通信がどのような遷移を経て発生したのかをオペレータが認識し，かつ当該通信群に対するセキュリティオペレーションを支援するシステムを提案することが主眼である．

3.1 セッショングラフ

HTTP 通信でリダイレクトが発生する場合，遷移元の情報は HTTP リクエストヘッダのリファラに記載される（なお，リファラはリダイレクトだけではなく，リソースファイルの取得時や，例えばリンクを辿るなどの遷移の際にも記載される)．リファラは，プロキシサーバを配している環境であれば，プロキシサーバののログとして保存することが可能な情報である．図 5 に，リクエストヘッダの例を示す．

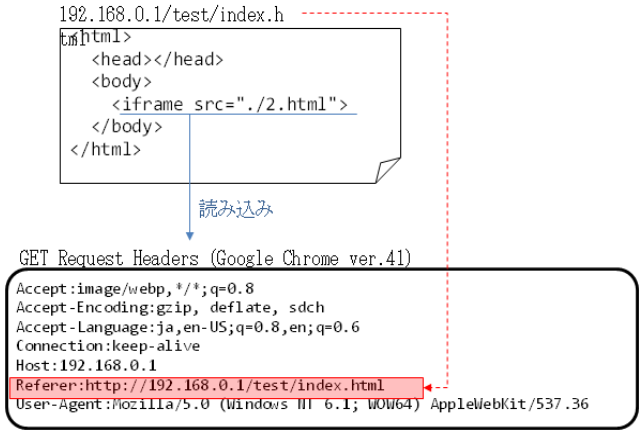

図 5 リファラヘッダ

Figure 5 Referer Header

また，図 6 に，対応するプロキシサーバのログの例を示す．

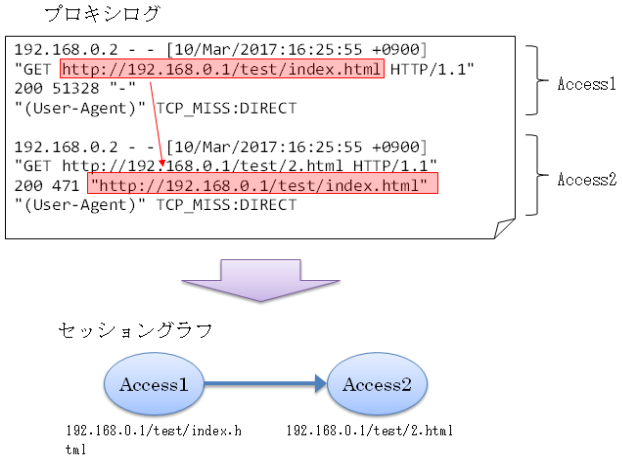

図 6 プロキシログと対応するセッショングラフの例

Figure 6 An Example of Proxy Server's Log and its corresponding Session Graph

プロキシログの中で，リファラ記載の URL と，アクセス先の URL が一致するログを関連する通信として紐付けていくことで，通信フローががどのような HTTP コネクションで構成されていたかを有向きグラフとして復元することができる．図 6 の下段の有向グラフは，Access2 が Access1 の通信が元となって発生したことを表している．以降，このように HTTP 通信間の関係をグラフで表したものを「セッショングラフ」と呼ぶ．

3.2 提案システム

悪性サイトへのアクセスの警告をオペレータがセキュリティ機構から受け取った際，下記に挙げた点を実現できることが対策の迅速化・効率化のために有用である．

- 警告の対象となった悪性サイトへの通信が，どのようなサイトを経由して遷移したか，高度な専門スキルを有さないオペレータであっても容易に理解できること
- 当該悪性サイトへの遷移に関わる通信のログのみを抽出できること

上記を実現するため，我々は図 7 に示すようなシステムを検討した．

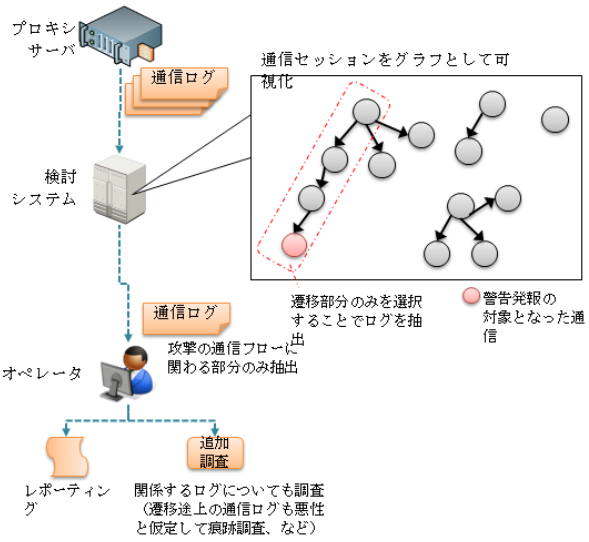

図 7 提案の可視化システム

Figure 7 Our Proposed Visualization System

本システムは，プロキシログを入力としてセッショングラフを構築し，オペレータに対して警告対象の通信の遷移状況を有向グラフとして提示するものである．また，オペレータのその後の対応としては，遷移系路上のアクセス先の調査とレポーティング等を行うことが想定されるため，当該対応に必要な情報を抽出して出力できる機能を備えている．

4. 評価

3 章に述べた設計を元に，実際にシステムを実装し，社内のプロキシサーバのログを可視化する実験を行った．なお，処理に用いた計算機のスペックは下記の通りである．

- CPU　Core i7 3.2GHZ
- メモリ 4GB
- DB ミドル MongoDB2.6

4.1 可視化画面の例

可視化の事例を図 8 に示す．

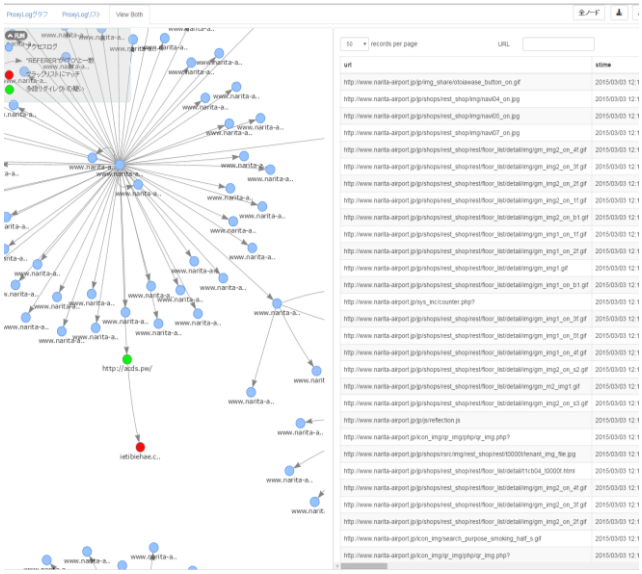

図 8 実装したシステムによる多段リダイレクトの可視化の例

Figure 8 An Example of Our Visualization of Multipre Redirection

図 8 の赤で示されているノードは，既知のブラックリスト URL への通信ログである．また，緑で示したノードは出次数が 1 のノードである．近年の Web ページの多くは画像やスクリプト，スタイルシートなどのリソースファイルをインクルードしていることが多い．そのような Web ページにアクセスした場合，セッショングラフ上では，骨組みとなる HTML ファイルへのアクセスログのノードの出次数は大きく，逆に，リソースへのアクセスログのノードは出次数が 0 となる．一方，多段のリダイレクトが発生したような場合は，出次数が 1 のノードが連結し，紐状の形状となる．図 8 を見るに，既知のブラックリスト URL にマッチした通信は，緑ノードで示したリダイレクタへの通信を契機に発生しており，かつ，リダイレクタへの通信がどのサイトへのアクセスから発生したかが可視化されている．また本試作は，GUI 上で選択した通信のログを抽出して出力できるよう設計しており，オペレータが対応に利用することが出来る．

4.2 処理時間

セッショングラフを構築するために要する処理時間と，処理対象のログのレコード数の関係を図 9 に示す．

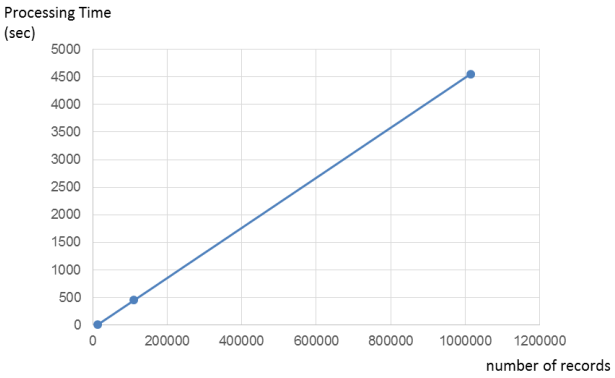

図 9 セッショングラフを構築するために要する処理時間

Figure 9 Processing Time for Constructing the Session Graph

本試作では，100 万レコードの処理におよそ 4500 秒を要する結果となった．また，レコード数の増加に対し，処理時間は線形に増加する結果となった．すなわち，我々が実験に用いた計算機環境では一日あたりおよそ 2000 万レコードまでなら対応可能である．

5. 考察と今後の課題

本提案システムでは，次の点を実現した．

- 警告の対象となった悪性サイトへの通信が，どのようなサイトを経由して遷移したかを可視化すること
- 当該悪性サイトへの遷移に関わる通信のログのみを抽出できること

本システムを利用することで，高度な専門スキルを有さないオペレータであっても，多段リダイレクトを伴う DbD 攻撃の遷移を容易に把握し，対応につなげることが出来るようになる．

一方で，Javascript などでリダイレクト通信を発生させる場合，リファラをヘッダに付与しないことも可能であるため，そのような場合には単純にはセッショングラフを構築できない課題がある．そのようなものについては，例えば，通信ヘッダのみではなく，通信ペイロードも保存しておき，当該通信の内容を分析する等により通信の遷移を調査することは可能である．ただし，通信ペイロードの保存に掛かるコストや，分析の難易度などの課題があるため，本システムに組み入れるためには，それらを解決する必要がある．

6. おわりに

攻撃者が DbD 攻撃に使用する悪性サイトはインフラ化してきており，攻撃件数，悪用されるサーバの数は増加しているのに対し，ログを元に攻撃の全容を把握できるスキルを備えたセキュリティオペレータの育成は間に合っていない状況にある．

そこで，本稿では多段リダイレクトの痕跡を抽出して可

視化することで，オペレータによる対策を迅速化するシステムについて検討した．既知のブラックリスト URL にマッチした悪性な通信を検知した場合，現状は，当該通信がどのような経緯で発生したかを調査するためにオペレータが手作業でログを再帰的に検索する必要があるが，本システムを利用することで，個々のオペレータのスキルに拠らず容易に悪性通信が発生するに至った通信フローを把握することができる．

今後は，5 章に課題として述べた，ペイロードの分析結果を元に，リファラに残らない遷移情報も復元する方式の検討，および，セッショングラフの構造の特徴に着目した検知手法の検討を行うことで本システムの改良を行う予定である．

参考文献

[1] 藤田, "侵入検知に関する誤検知低減の研究動向", 信学論(B) Vol.J98-B No.4, pp.402-411, 2006

[2] "Google Safe Browsing", [online] https://www.google.com/transparencyreport/safebrowsing

[3] "Mcfee Site Advisor", [online] http://siteadvisor.com

[4] 義則隆之，他，"通信可視化と動的解析の連携による攻撃解析支援"，コンピュータセキュリティシンポジウム 2012 論文集, p.224-231, 2012.

[5] 金子博一，他，"通信トラフィックの分析による Gumbler 感染 PC の可視化", IEICE Technical Report, IA2010-1 ICSS2010-1, 2010

[6] 尼子，他，"情報資格化による Drive-by Download 攻撃対策の一検討", 2014-CSEC-64 No.41, 2014

[7] 松本，他，"Drive-by-download 攻撃通信の可視化システム", CSS2014, 2014