企業での実環境を考慮したサイバー攻撃検知システムの有効性評価

重田 真義† 大谷 尚通† 大嶋 真一†

† 株式会社 NTT データ
135-8671 東京都江東区 豊洲 3-3-9
{shigetam,ootanihs,ooshimasni}@nttdata.co.jp

あらまし 我々は，ネットワーク機器のログを監視してサイバー攻撃を検知するシステムを開発し，運用している．本システムは定期的にログを取得し，攻撃を調査するため，検知までにタイムラグが発生するおそれがあった．また，ログをリアルタイムに転送すると，装置の追加や既存環境の大幅な変更が必要になったり，ネットワーク機器へ負荷がかかったりするおそれがある．そこで我々は，ネットワーク上に流れるパケットをキャプチャし，それを解析することで，感染した端末をリアルタイムに検知する手法の設計と実装を行った．本発表では，実環境における検知実績をもとに，リアルタイム検知方式の有効性を評価する．

The assessment of the effectiveness of cyber attack detection system for enterprise use

Masayoshi Shigeta† Hisamichi Ohtani† Shinichi Ohshima†

†NTT DATA Corporation
Toyosu 3-3-9 Koto-ku Tokyo 135-8671, JAPAN
{shigetam,ootanihs,ooshimasni}@nttdata.co.jp

Abstract We develope and use the system to detect cyber attacks by monitoring network devices. This system cannot detect before serious information leakage because the system regularly gets and uses the network devices' log to detect attacks. If it is transferred to the system in real time, it may be necessary for us to add some devices and change current network environment drastically and become overloaded with network devices. Therefore, we design and implement the real time detection system to capture the packet on the network and analyze it.

In this paper, we describe the evaluation of the real time detection system by the detection results in enterprise use.

1 はじめに

企業・官公庁・教育機関等に対するサイバー攻撃は，日々その激しさを増しており，それにより発生するセキュリティインシデントは，JPCERT コーディネーションセンターに報告があったものだけでも，年間約 19000 件 [1] 存在する．また，これらのサイバー攻撃により，深刻な機密情報の漏えいが発生したり，ビジネスの継続が危ぶまれたりするケースも増加しており，新聞・ニュースでもこれらの攻撃による被害等が頻繁に報道されている状況である．このような状況を受け，各企業では，より一層，それらのサイバー攻撃への対策を強化する必要性が増している．

また，近年のサイバー攻撃は多様化・複雑化

が進んでおり，セキュリティ対策が難しくなってきている．特に侵入されて半年以上経って被害が大きくなってきてから，サイバー攻撃が発見された事件がいくつも存在する [2]．なぜなら，危険な Web サイトのブラックリスト情報やウイルス定義ファイル情報の更新が追い付かず，URL フィルタ，IPS 等で最新のサイバー攻撃を未然に防止することや，ウイルス対策ソフト，IDS 等で最新のサイバー攻撃を検知することが難しくなってきているためである．これを踏まえ，昨今，サイバー攻撃を未然に防止するだけではなく，インシデントの発生を早期に検知して被害を最小化する対策 [3][4] も求められている．このような状況を受けて，本研究では，ネットワーク機器のログを分析して，最新のサイバー攻撃によってマルウェアに感染した端末を早期に検知する「サイバー攻撃検知システム」を開発してきた [5][6][7].

2 ネットワーク機器の通信ログを活用したサイバー攻撃検知システム

2.1 本システムの概要

我々が提案し，開発および運用を進めてきたネットワーク機器の通信ログを活用したサイバー攻撃検知システム [5] は,（ウイルス対策ソフトや IPS/IDS 等のセキュリティ製品では未然の検知と防止が難しい）サイバー攻撃によって，マルウェアに感染した端末を検知できる．特に本システムは，Gumblar 攻撃 [8] に代表される，ユーザが Web サイトへアクセスしたときにマルウェアをダウンロードさせられる Drive by Download 攻撃（以下,「DbD 攻撃」という）の検知を得意としている．

本研究では，サイバー攻撃の事例情報を多数収集して，それらの攻撃の一連の流れをモデル化している．DbD 攻撃の複数の感染事例を収集・分析してモデル化したところ，マルウェア感染時（以下,「感染フェーズ」という）に特徴的な振る舞いがあることがわかった．この感染フェーズの特徴的な振る舞いを redirect ステップ，pre-exploit ステップ，exploit ステップ，pre-DL ステップ，malware DL ステップの 5 つのステップに分類した [5]．5 つのステップの流れを図 1 に示す.

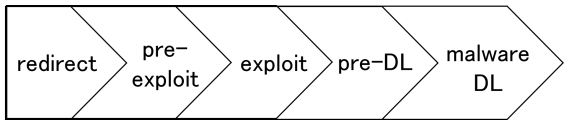

図 1: 感染フェーズの特徴的な 5 つのステップ

本システムは，できる限り早期にマルウェアに感染した端末を検知することが目的である．そのため，感染フェーズの振る舞いから定性的な特徴 [5] を見つけて検知パターンを開発し，運用することを優先している．また，感染フェーズで検知できなかった場合のために，マルウェア感染後の C & C フェーズに見られる特徴でも検知できるよう，それらの特徴を捉える検知パターンも開発・運用している．

2.2 検知タイムラグの問題

本研究では，前述のサイバー攻撃検知システム (以下,「日次版検知システム」という) を用いて，ある実環境の通信ログを 1 日 1 回まとめて収集・分析して，DbD 攻撃の監視を日々続けており，年間 100 件以上のマルウェア感染が疑われる端末を検知している．しかしながら，1 日分の通信ログをその翌日に分析する運用では，マルウェアに感染してから検知するまでに最大 24 時間経過してしまうおそれがある．

また，ベライゾンの調査報告書 [2] では，ネットワーク侵入によるデータ漏えい/侵害に関わるセキュリティインシデントの 33 %は，1 時間以内に機密情報が漏えいすると報告されており，企業などの機密情報を狙う最新のセキュリティインシデントに対応するためには，DbD 攻撃をリアルタイムに検知できなければならない．

そこで本研究では，ネットワーク機器の通信ログをリアルタイムに分析することで，DbD 攻撃によりマルウェアに感染した端末を 10 分以内に検知するという目標を新たに立て，その問題の解決に取り組んだ．

3　通信ログのリアルタイム取得方式の検討

通信ログをリアルタイムに取得する方式について，以下の 2 種類の実現方式を考え，それらを比較・検討した．

3.1　リアルタイム取得方式の比較

1. リアルタイムログ転送方式

   各種ネットワーク機器からログ分析システムへ通信ログを随時，転送する方式．

   メリット

   通信ログを転送する方式であるため，複雑な処理が不要．

   デメリット

   リアルタイムに大容量の通信ログを転送する必要があるため，ネットワークに大きな負荷がかかる場合がある．これを解決するために，ネットワーク機器の増強や，大幅なネットワーク構成の見直しが必要になるおそれがある．

2. パケットキャプチャリング方式

   ユーザ端末 - ネットワーク機器間の通信路に流れる通信パケットをキャプチャし，その通信パケットをネットワーク機器の通信ログ形式に変換する方式

   メリット

   大幅なネットワーク構成の変更を行うことなく，通信パケットをキャプチャする機器をネットワークへ接続するだけで，実現できる．既に構築済みのネットワークに導入しやすい．

   デメリット

   通信パケットを通信ログ形式へ変換する処理が必要なため，システム構成が複雑である．通信パケットのキャプチャ漏れが発生しないよう，トラフィックのピークに対応可能な性能のキャプチャ装置を採用する必要がある．また，システム運用後のネットワーク環境の変化に伴うトラフィック量の増加を予想した設計が難しい．

3.2　実現方式の決定

本システムは，既に構築済みのネットワーク環境でリアルタイム検知を実現しなければならない．よって，既存のネットワーク構成を大幅に見直す必要がない，パケットキャプチャリング方式を採用した．

4　リアルタイム検知処理の課題

4.1　通信ログの蓄積時間と検知漏れ

日次版検知システム [5] は，1 日 1 回，1 日分の通信ログに対して検知パターンを使って検索し，マルウェアに感染した端末を発見する．したがって，日次版検知システムは，1 日分の通信ログのどこかに不審な通信の遷移パターンが存在すれば，検知できる．

これに対して，通信ログをリアルタイムに検索する場合は，通信ログを一定時間蓄積し，検知パターンを使ってその蓄積した通信ログを検索する．このとき，通信ログの蓄積時間はできる限り短いほうが，リアルタイム性が高い．DbD 攻撃の感染フェーズの検知パターンとは，感染フェーズ (図 1) の複数ステップを遷移したときの特徴的な通信ログのパターンを定義したものである．この複数ステップの遷移には時間がかかる．

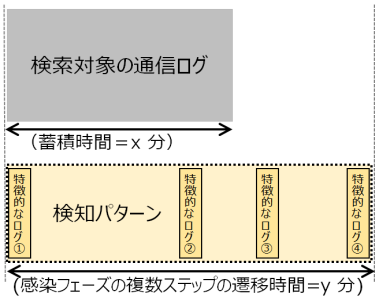

図 2: 通信ログの蓄積時間と検知漏れ

図 2 のように，通信ログの蓄積時間 x 分が，定性的な特徴のある通信ログの遷移時間 y 分よりも短い場合，定性的な特徴のある通信ログが検索対象の通信ログにおさまらない．そのため，不審な通信の遷移を捉える検知パターンを使っ

て蓄積された通信ログを検索したとしても，何も検知できない．したがって，検知パターンによる検知漏れがなく，かつリアルタイム性が高い適切な通信ログの蓄積時間 x 分を決定しなければならない．

4.2　通信ログの検索方式による検知漏れ

蓄積された x 分間の通信ログを x 分間隔で順次検索した場合を図 3 に示す．感染フェーズの通信ログが，検索対象ログ A と検索対象ログ B にまたがって存在した場合，これを検知できない．このような検知漏れが発生しないよう，検索方式を工夫しなければならない．

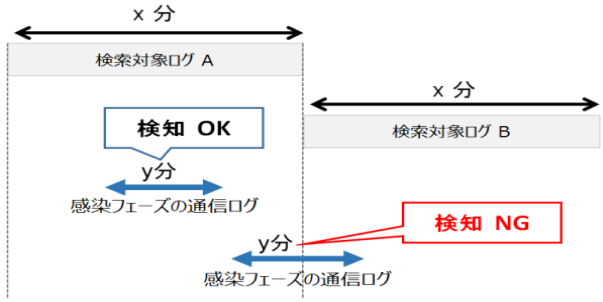

図 3: 通信ログの検索方式による検知漏れ

4.3　検索処理の遅延

本システムをリアルタイム化する場合は，感染フェーズの検知パターン，C & C フェーズの検知パターンを同時に短い間隔で検索しなければならない．トラフィックのビーク時間帯でも，最低限，通信ログの蓄積時間以内，つまり実時間以内で検索処理を終了しなければならない．

5　各課題に対する対応方針

本章では，4 章で述べた各課題に対する対応方針を記載する．5.1 節で通信ログの蓄積時間に関わる検知漏れの対応方針を，5.2 節で通信ログの検索方式による検知漏れの対応方針を，5.3 節で検索処理の遅延に対する対応方針を記載する．

5.1　通信ログの蓄積時間の決定

検知パターンによる検知漏れがなく，かつリアルタイム性が高い適切な通信ログの蓄積時間 x 分を決定するために，本システムで検知した約 2 年間分の DbD 攻撃 約 200 件の通信ログを分析し，感染フェーズにおける定性的な特徴のある通信ログの遷移時間を調査した．遷移時間とは，図 1 の redirect ステップから malware DL ステップが終了するまでにかかる時間とする．ただし，分析した通信ログの中には，Exploit Kit による攻撃が途中で失敗し，malware DL ステップまで至らなかったログも存在する．その場合は，exploit ステップや pre-DL ステップまでの時間を使用した．

表 1: 感染フェーズの遷移時間

Exploit Kit 名	最小 [秒]	最大 [秒]
Blackhole EK	4	63
Angler EK	18	50
Cool EK	30	30
Fiesta EK	18	28
Gongda EK	8	34
Goon EK	18	41
Neutrino EK	4	64
Nuclear EK	26	47
Redkit EK	13	13
RIG EK	21	52
SweetOrange EK	13	36
不明	4	69

表 1 にその調査結果を示す．Exploit Kit ごとに最小遷移時間と最大遷移時間をまとめた．上記の調査の結果，redirect ステップから malware DL ステップが終了するまでにかかる時間は，最長でも 69 秒であることが判明した．よって，定性的な特徴のある通信ログの遷移時間 y 分を包含できる通信ログの蓄積時間 x は，2 分以上とした．

5.2　通信ログの検索方式の工夫

本システムの検知パターンには，マルウェア感染時の振る舞いを捉える検知パターン（以下，「感染フェーズの検知パターン」という）と，マルウェア感染後の振る舞いを捉える検知パターン（以下,「C & C フェーズの検知パターン」という）の 2 種類が存在する．

5.2.1 感染フェーズの検知パターンの検索方式

4.2 節で述べた感染フェーズの通信ログが，検索対象ログにまたがって検知できない問題を解決するために，図 4 のように，検索対象ログを一定時間（z 分間）重ねて検知パターンで検索する方式を採用した．

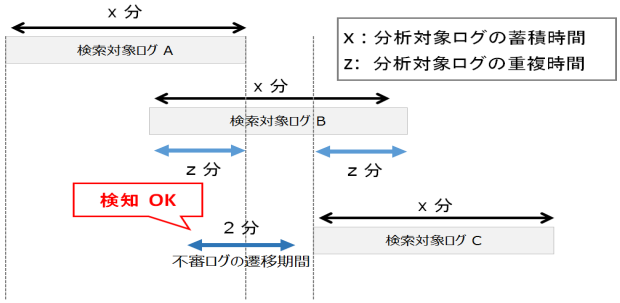

図 4: 通信ログの重複検索

ここで，x：分析対象ログの蓄積時間，z：ログの重複時間，定性的な特徴のある通信ログの遷移時間を 2 分間とすると，本実現方式において，

$$x > z \geq 2$$

が成立する場合，検知漏れは発生しない．近年，マルウェアに感染してから，数分で情報漏えいが発生する事例も確認されている [2] ことを考慮し，分析対象ログの蓄積時間 x は 5 分に，ログの重複時間 z は 2 分に設定した．よって，検知パターンの実行間隔 $x - z$ は 3 分となる．検知パターンの実行間隔が 3 分であるため，感染フェーズの検知パターンの目標実行時間も 3 分以下に設定した．

5.2.2 C & C フェーズの検知パターンの検索方式

C & C フェーズの検知パターンには，不審なアクセス先の URL のパスに見られる特徴的な文字列を定義した検知パターン（以下,「ブラックリスト型検知パターン」という）と，複数行の通信ログの振る舞いを定義した検知パターン（以下,「振る舞い型検知パターン」という）の 2 種類が存在する．C & C フェーズのブラックリスト型検知パターンは通信ログ 1 行のみで判断できるため，分析対象ログの蓄積時間と実行間隔ともに 5 分，ログの重複時間は 0 分とした．C & C フェーズの振る舞い型検知パターンは，マルウェアが C & C サーバと長時間 C & C 通信を行った場合に検知するロジックのため，分析対象ログの蓄積時間を長時間に設定しなければならない．C & C フェーズの振る舞い型検知パターンの分析対象ログの蓄積時間は 6 時間に設定した．

5.3 検索処理の遅延への対応

検知パターンの実行時間を短縮するためには，分析マシンのハードウェア性能を増強する方法がある．しかしハードウェア性能の増強はコストが高く，限界も存在する．そこでハードウェア性能の増強に加えて，検知パターンの処理方式を変更して高速化する方法を採用した．

本システムの検知パターンは，端末の IP アドレスなどの共通点を使って通信ログを複数行のグループへまとめて，その通信ログのかたまりに対して不審な通信の遷移の有無を検索する．この大量の通信ログに対して「共通点をもとに通信ログをグループ化する処理」と「グループ化した通信ログ内から不審な通信の遷移を発見する処理」は時間がかかる．

特に様々な通信が混在した大量の通信ログから共通点をもつ通信を検索してグループ化する処理は，そのままでは非常に多くの処理時間がかかってしまい，通信のピーク時に検索処理遅延が発生する問題点があった．そこで，あらかじめ通信ログを共通点でソートした後に，共通点をもつ通信をグループ化する処理方法へ変更することとした．これにより通信のピーク時でも，設定した目標時間以内に検索処理が終了できると考えた．

6 リアルタイム版サイバー攻撃検知システムの実装

リアルタイム版サイバー攻撃検知システム (以下,「リアルタイム版検知システム」という) の

概要を図 5 に示す.

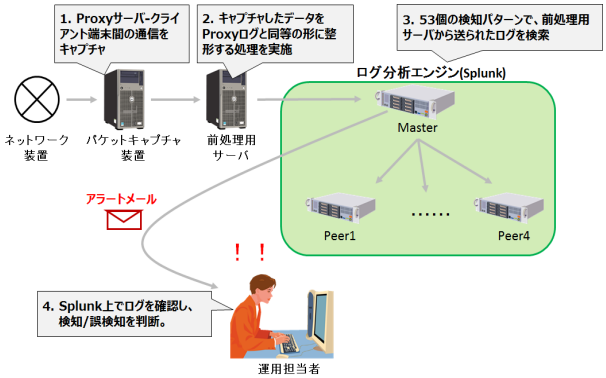

図 5: リアルタイム版サイバー攻撃検知システムの構成

本システムは，データベース処理の基本能力に加え，情報の集約，正規化，相関分析の 3 つの基本処理に対応できるソフトウェア Splunk を用いてデータベース基盤を構築した.

感染フェーズの検知パターンすべてを目標実行時間の 3 分以下で処理を終了させなければならない．そこで検知パターンの同時実行数を増やすため，Splunk をクラスタ構成化し，検索処理用の端末の台数を 4 台へ増強した.

通信ログをリアルタイムで検索して DbD 攻撃によってマルウェアに感染した端末を検知した場合，感染端末を早期に対処するために運用担当者はすぐに検知に気づかなければならない．そこで感染端末を検知した場合，運用担当者へ検知結果のメールを送信して通知する仕組みも実装した.

7 リアルタイム版検知システムの評価

リアルタイム版検知システムを，実環境での検知結果，および検知パターンの実行時間の 2 点で評価する．前者は，リアルタイム化に伴う検知漏れと誤検知を評価した．後者は，設定した目標時間の 3 分以内に感染フェーズの検知パターンすべての検索処理が終了することを検証した.

7.1 検知漏れと誤検知の評価

ある実環境のネットワークへリアルタイム版検知システムを導入し，2015 年 4 月〜8 月の 5ヶ月間，DbD 攻撃を監視した．その検知結果を表 2 に示す.

表 2: ある監視ネットワーク環境の DbD 攻撃の検知件数

Exploit Kit	4 月	5 月	6 月	7 月	8 月	合計
Nuclear EK	7	1	2	4	2	16
Angler EK	4	4	6	4	1	19
不明 (※)			4	14		18
合計	11	5	12	22	3	53

(※)DbD 攻撃が失敗して遷移が止まった攻撃を含む

リアルタイム版検知システムは，マルウェア感染が疑われる不審な通信を 53 件検知した．2015 年 4 月〜8 月は，Nuclear Exploit Kit，および Angler Exploit Kit と思われるマルウェア感染を多く検知した．広告ネットワークを介したものと思われるマルバタイジングと呼ばれる感染手法も検知した.

日次版検知システムを上記のリアルタイム版検知システムと同じネットワークで，同一期間運用したところ，検知数 53 件，および検知した Exploit Kit の内訳も一致した．このことから，リアルタイム化に伴う検知漏れと誤検知はなかった.

7.2 検知パターンの処理時間による評価

検索処理遅延への対応を行う前に検索処理時間が目標時間の 3 分を越えていた感染フェーズの検知パターン 2 個について，5.3 節に記載した検索処理の高速化手法適用後の計測結果を図 3 に示す．分析対象ログは，トラフィックのピーク時刻の 5 分間のログを用いた．高速化手法を採用したことにより，約 7 分近くかかっていた処理時間が，おおよそ 3 分の 1 の時間で実行できることがわかった.

表 3: 検知パターンの処理時間の比較

検知パターン	高速化適用前	高速化適用後
検知パターン A	6 分 28 秒	2 分 16 秒
検知パターン B	7 分 8 秒	2 分 19 秒

ログ行数：384,000 件，データサイズ：0.4GB

同じトラフィックのピーク時刻の 5 分間のログに対して，感染フェーズの検知パターンすべてを検索し，設定した目標時間の 3 分以内に検索処理がすべて終了することを確認した．以上より，ハードウェア性能の増強と処理の高速化により当初の目標を達成することができた．

8 まとめと効果

8.1 まとめ

本稿では，DbD 攻撃によりマルウェアに感染した端末を 10 分以内に検知するために，日次版サイバー攻撃検知システムをリアルタイム化するための手法を検討し，実装した．リアルタイム版サイバー攻撃検知システムの通信ログのリアルタイム取得方式は，パケットキャプチャリング方式を採用し，日次版サイバー攻撃検知システムを運用している既存ネットワーク環境へネットワーク構成の大幅に変更なく導入を完了した．

リアルタイム版検知システムは，通信ログの蓄積時間や重複検索などの検索方式の工夫により，日次版検知システムと同じ DbD 攻撃を検知できることがわかった．これによりリアルタイム化に伴う検知漏れと誤検知はなく，日次版検知システムと同等の検知率があることが確認できた．またハードウェア性能の増強に加えて，検知パターンの処理方式を高速化することにより，トラフィックのビーク時間帯でも感染フェーズのすべての検知パターンの検索処理を目標時間の 3 分以内に完了できるようになった．

8.2 副次効果

日次版検知システムでは，マルウェアに感染してから検知するまでにタイムラグがあったため，感染端末のインシデント対応を行うときには，改ざんされた Web サイトや Exploit コード，マルウェアを配布している Web サイトが存在しない場合が多かった．リアルタイム版検知システムは，DbD 攻撃を 5 分程度で検知して運用担当者へ通知できるようになった．そのため，感染端末のインシデント対応を行うときには，改ざん Web サイトや Exploit コード，マルウェア配布 Web サイトが存在し，迅速に攻撃手法に関する情報などの解析や対応に有益な情報を収集できるようになった．収集できた有益な情報は，攻撃に対するセキュリティ対策や次の研究の足掛かりになった．

8.3 今後の課題

今回は，マルウェアに感染した端末を迅速に検知する仕組みを検討したが，本システムで検知した後の対応は不十分であり，今後の課題であると考えている．本システムで検知した後，改ざんされた Web サイトやマルウェアを配布している Web サイトへのアクセスを迅速に遮断する仕組みを整備することや，迅速にマルウェア配布サイトにアクセスし，そのマルウェアを取得した後，そのマルウェアの挙動を把握する体制を整備することを進めていきたい．

参考文献

[1] JPCERT/CC，JPCERT/CC インシデント報告対応レポート，https://www.jpcert.or.jp/ir/report.html, accessed Aug. 24, 2015.

[2] ベライゾンジャパン，2013 年度データ漏洩/侵害調査報告書，https://www.verizonenterprise.com/jp/DBIR/2013/，accessed Aug. 24, 2015.

[3] 笠間 貴弘，神薗 雅紀，井上 大介，Exploit Kit の特徴を用いた悪性 Web サイト検知手法の提案，マルウェア対策人材育成ワークショップ 2013.

[4] 進藤 康孝，佐藤 彰洋，中村 豊，飯田 勝，マルウェア感染ステップのファイルタイプ遷

移に基づいた Drive-by Download 攻撃検知手法，マルウェア対策研究人材育成ワークショップ 2014.

[5] 北野 美紗，大谷 尚通，宮本 久仁男，Drive-by-Download 攻撃における通信の定性的特徴とその遷移を捉えた検知方式，マルウェア対策研究人材育成ワークショップ 2013.

[6] 大谷 尚通，北野 美紗，重田 真義，企業内ネットワークの通信ログを用いたサイバー攻撃検知システム，マルウェア対策研究人材育成ワークショップ 2013.

[7] 大谷 尚通，益子 博貴，重田 真義，実環境におけるサイバー攻撃検知システムの有効性評価および検知範囲の拡大に向けた検討，マルウェア対策研究人材育成ワークショップ 2014.

[8] 独立行政法人 情報処理推進機構，
”ガンブラー”の手口を知り、
対策を行いましょう，
http://www.ipa.go.jp/files/000008662.pdf, accessed Aug. 24, 2015.