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

2013/10/23 MWS2013

NTTデータ
○北野美紗，大谷尚通，宮本久仁男

Global IT Innovator NTT DATA Group NTT DATA

目次

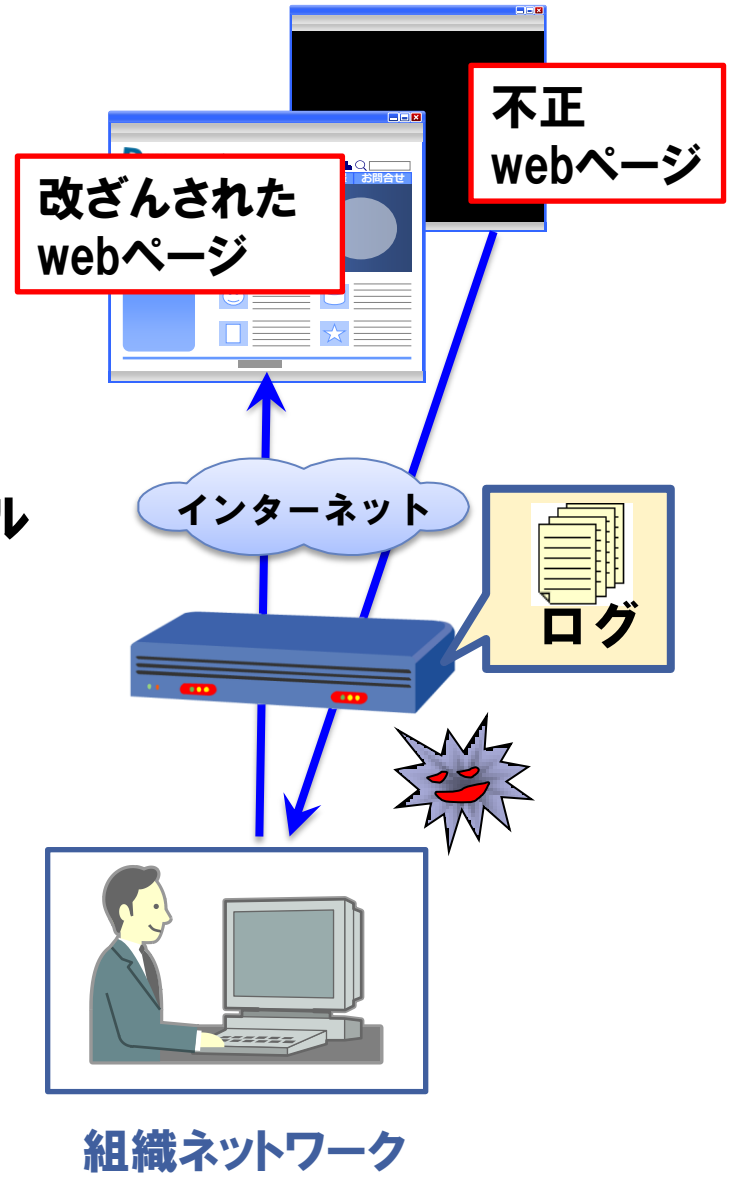
目次

1-1. 背景～DBD攻撃の増加～

Drive by Download攻撃(以下DBD攻撃)により，
改ざんされたwebページから，強制的にマル
ウェアをダウンロードさせる攻撃が増加

■バックグラウンドで感染が進行
　→感染に気づきにくい

■Exploit Kitの多用
- 複数のexploitコードをパッケージ化したツール
- 攻撃の簡易化→感染件数が大幅に増加



クライアント側の検知では限界

通信ログ(proxyサーバなどの通信機
器に蓄積されるログ)を利用し，
Exploit Kitを利用した攻撃を検知する.

1-2 通信ログからDBD攻撃を検知する既存手法

1. URL BlackListを利用する手法
 →攻撃者サーバのURLは短期間で動的に変化・消失

2. URLの文字列上の特徴に着目する手法
 Exploit Kitごとに現れる特徴的な文字列を利用

例： 

Ver 2.0

http://[domain]/ed350d2de32a1fcc/q.php
16文字のパス　1文字のファイル名

- 新しいExploit Kitの登場
- バージョンアップ
- 設定ファイルの操作

- 特徴パターンを変化させる
  例：16文字のパス→32文字のパス
- 特徴を持たせなくすることも可能
  例: http://[domain]/forum/info.php

URL項目を利用する既存手法には限界がある

1-3. 本研究の目標

■目標
URL以外の項目も利用して、感染時のふるまいを捉える汎用的なログ検知ルールの作成

検証

■検知ルール検証の観点
・検知
　- DBD攻撃全体のうち、今回の検知方式で検知できるものはどの程度か、またその原因調査
・誤検知
　- 発生しやすいルール、その原因の調査

目次

2-1. 検知方式の概観

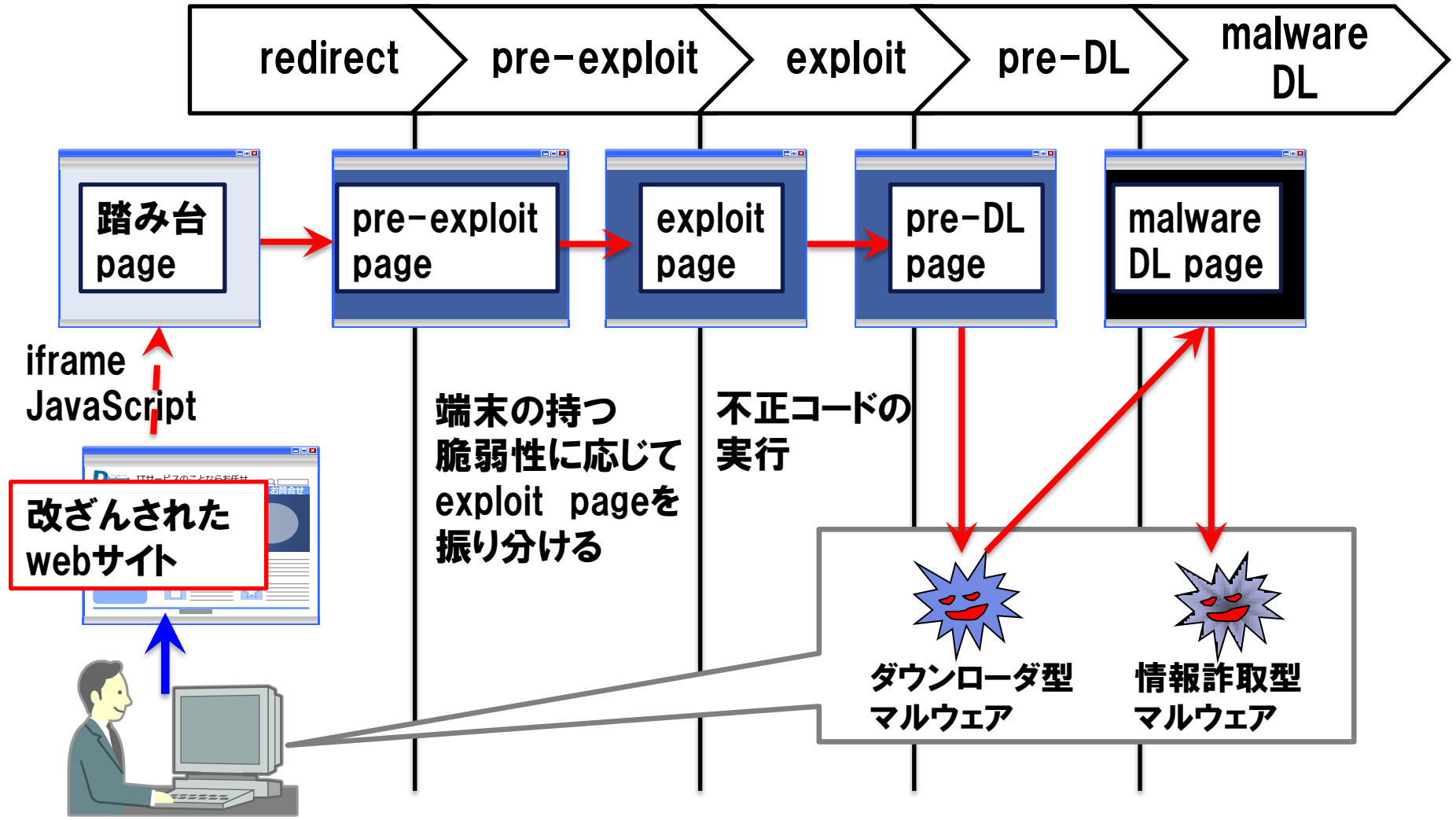
目標：ログ中のURL以外の項目も利用して，
感染時のふるまいを捉える汎用的なログ検知手法

DBD攻撃が共通の感染ステップを踏んで，
段階的に進行することに着目

感染ステップの進行に伴って，「定性的な特徴」が
「遷移」していくことを利用し，検知方式を考案

2-2. DBD攻撃の感染ステップ

通信ログに現れる特徴に着目し，5段階の感染ステップに分類



2-3. 本研究の特徴1～定性的な特徴～

定性的な特徴＝DBD攻撃のメカニズムに依存した特徴

- Exploit Kitの利用者側では，容易に変更できない
- 複数のExploit Kitに共通している

例： 感染時の通信ログ（1件） 赤字：今回利用する定性的な特徴

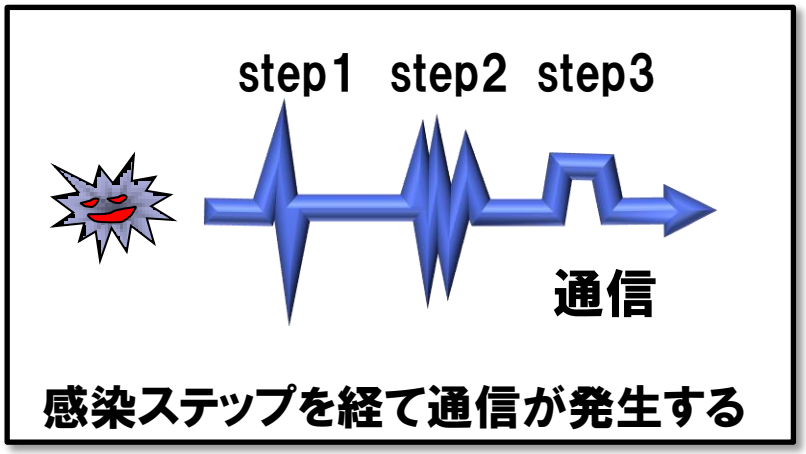
日付日時	18/Sep/2013:12:46:44
メソッド	GET
受信バイト数	325
URL	http:// [domain] /forum/info.php
status	200
UserAgent	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)

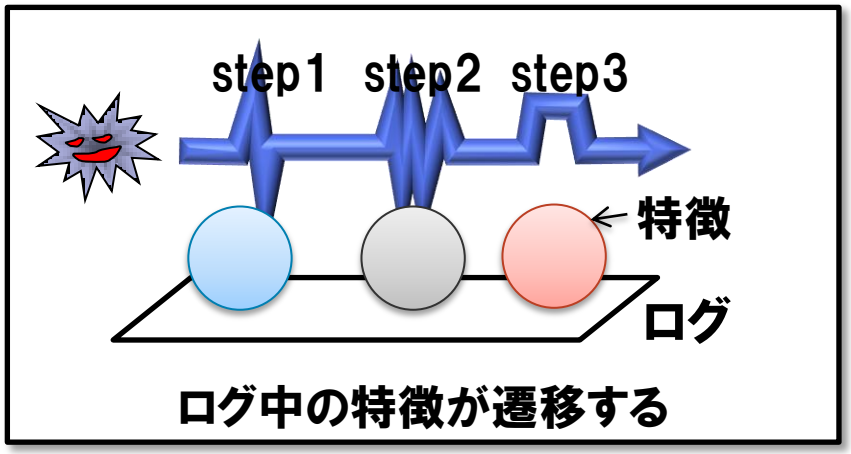
ファイルの suffix

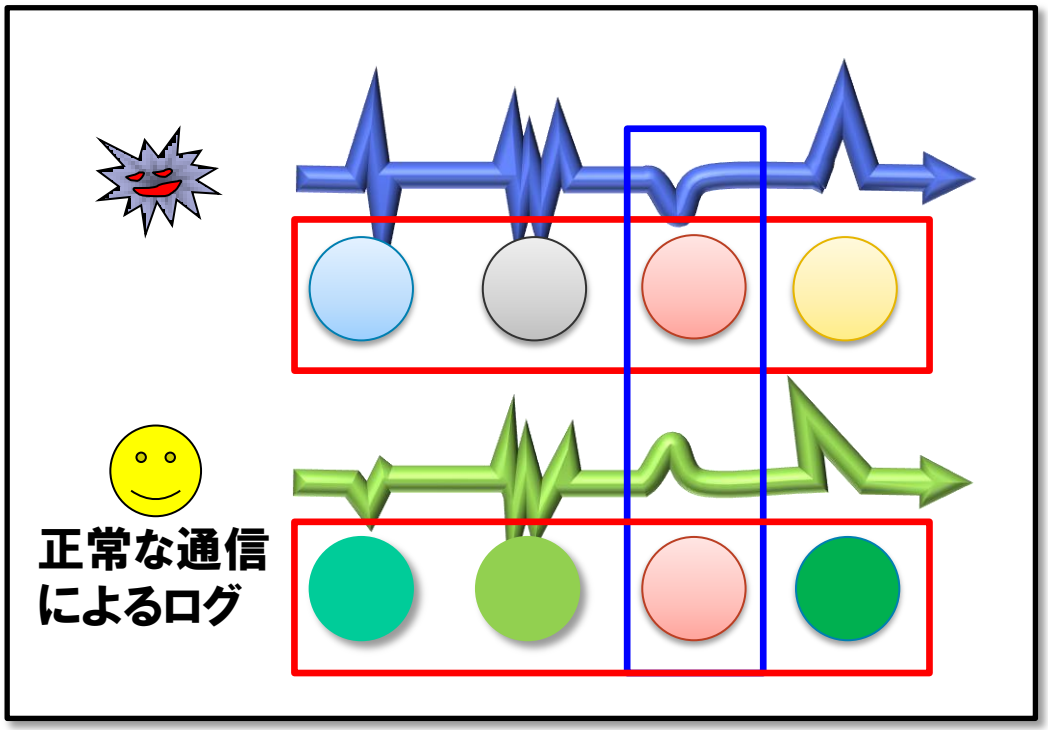
■問題点

通信ログ1件の特徴のみでは，誤検知が発生する

2-3. 本研究の特徴2～遷移の利用～







単独の特徴の利用では誤検知する

遷移を利用し，感染ログと正常なweb通信ログを判別

目次

1. 背景
2. 本研究で提案する検知方式
3. 定性的な特徴の遷移
4. 検証
5. まとめ

3. 定性的な特徴の遷移

今回利用する遷移は2種類

3-1. DBD攻撃に共通する特徴の遷移
3-2. Exploit Kitに固有の特徴の遷移

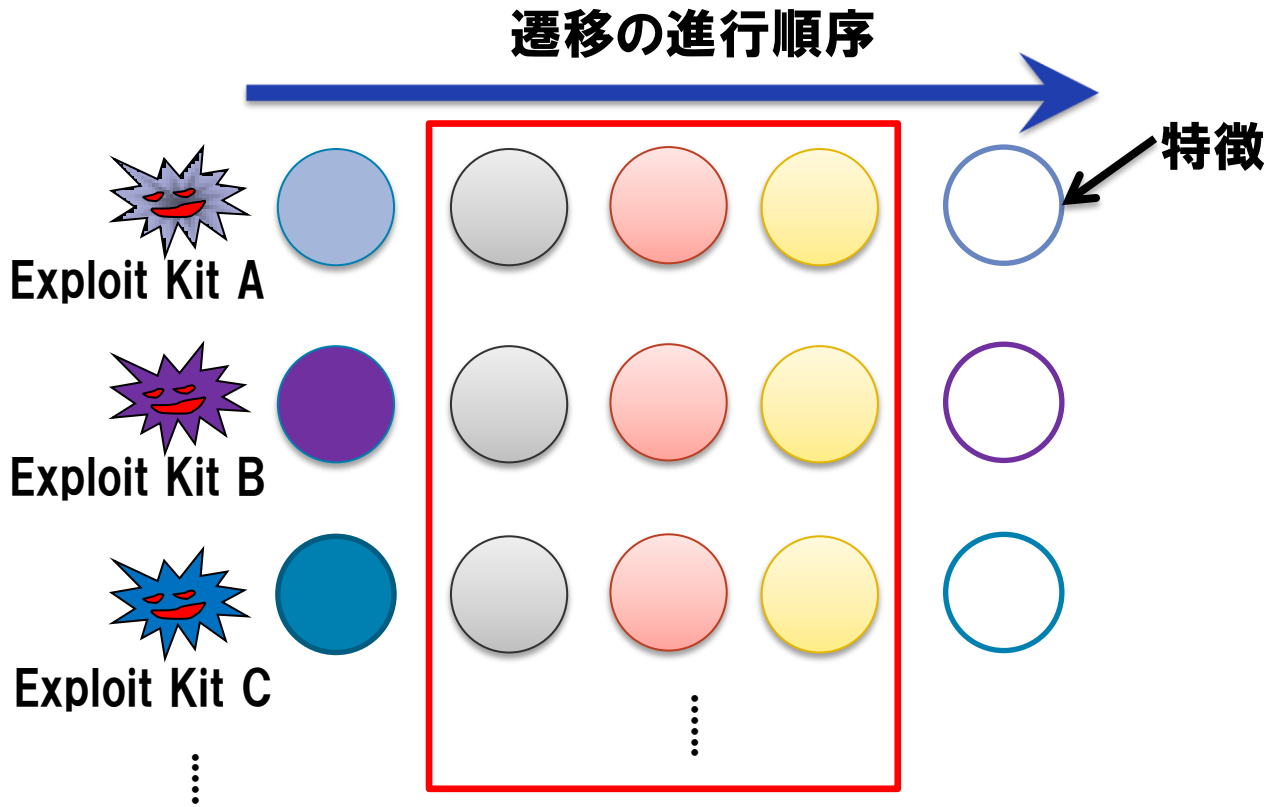
3. 定性的な特徴の遷移

今回利用する遷移は2種類

3-1. DBD攻撃に共通する特徴の遷移
3-2. Exploit Kitに固有の特徴の遷移

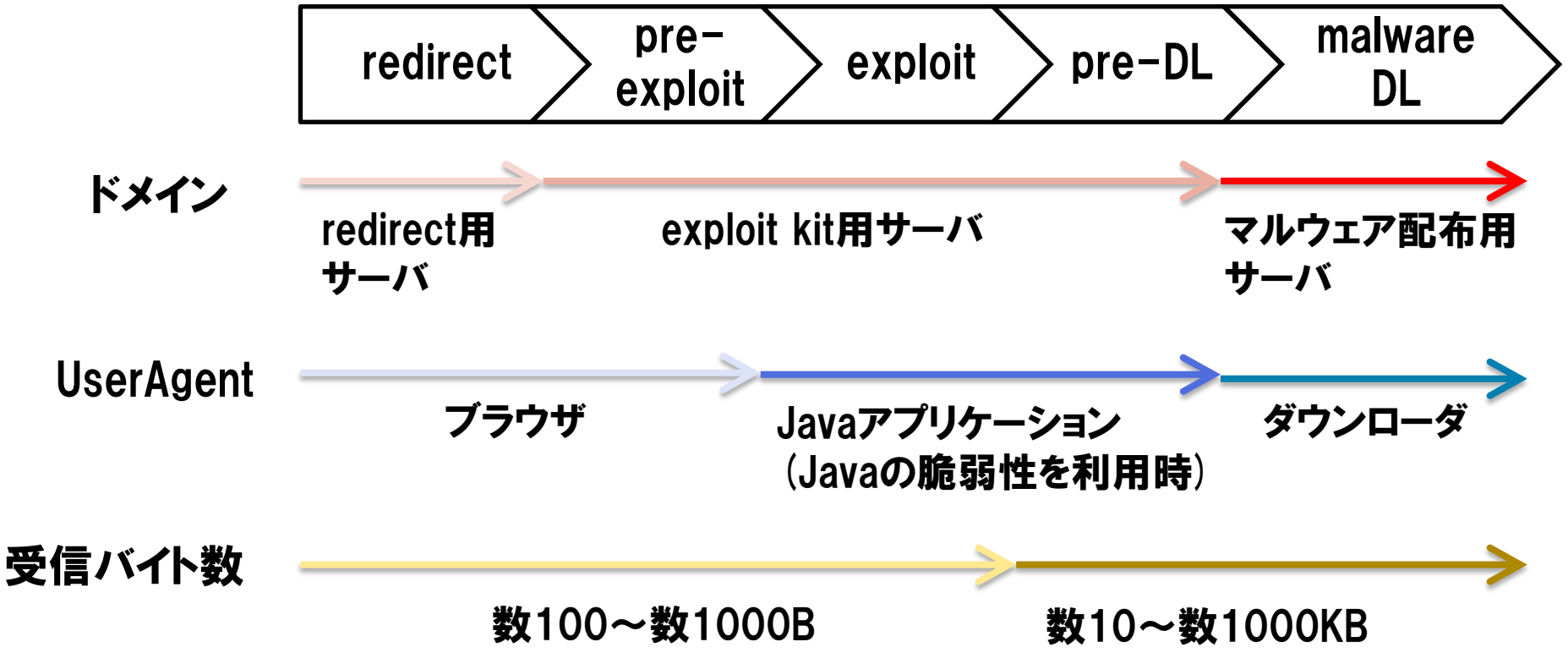
3-1. DBD攻撃に共通する特徴の遷移

DBD攻撃事例，100件以上を分析→共通する特徴の遷移



DBD攻撃の基本メカニズムによって発生し，
Exploit Kitの種類に依存しない

3-1. 今回利用した定性的遷移



Exploit時に利用する脆弱性によって、異なった遷移が発生

→依然、正常webアクセスとの誤検知が多い

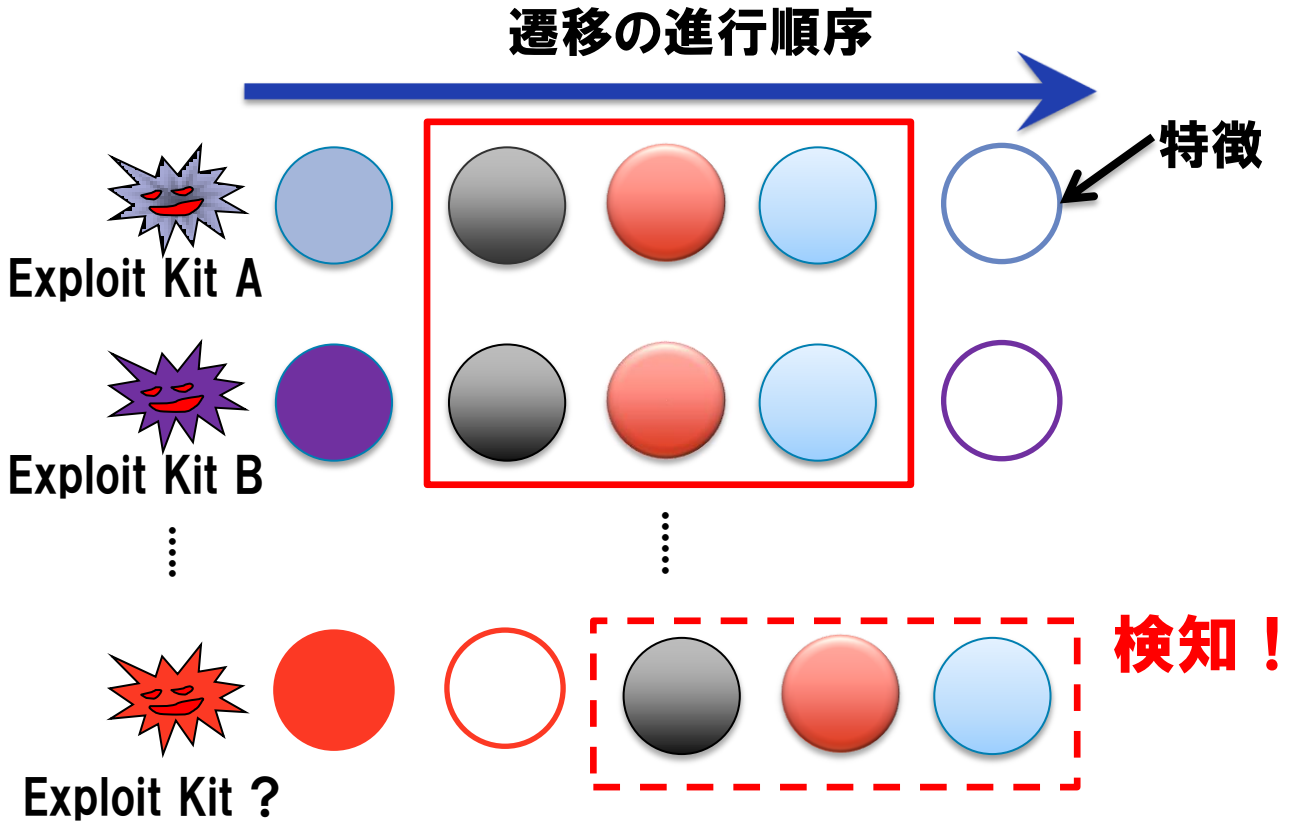
3. 定性的な特徴の遷移

今回利用する遷移は2種類

3-1. DBD攻撃に共通する定性的な特徴の遷移
3-2. Exploit Kitに固有の特徴の遷移

3-2. Exploit Kit固有の定性的な特徴の遷移

複数種類のExploit Kitで，同一の特徴の遷移がみられた



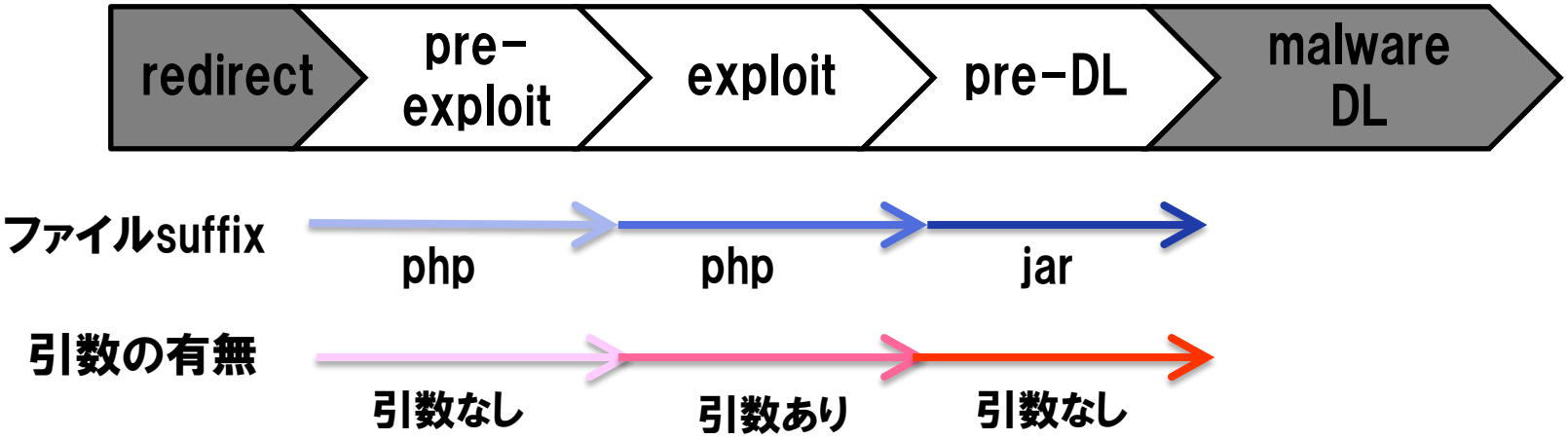
仮説：あるExploit Kitに固有の特徴の遷移を捉えることで，他のKitも検知することができる

3-2. Exploit Kit固有の定性的な特徴の遷移

実際に利用した定性的な特徴

- ファイルのsuffix
- webアプリケーションに渡す引数の有無

■例：BlackHole Exploit Kit　ver 2.0



8種類のExploit Kit（2013年8月時点のバージョン）を調査

3-3. 検知ルールの作成

3要素をパラメータとして検知ルールを作成

（要素1: DBD攻撃に共通する特徴遷移）

※主に、利用する脆弱性に依存

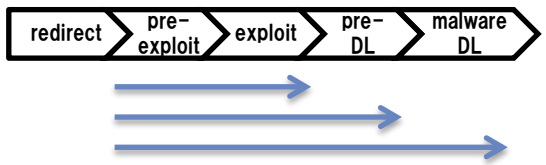
×

（要素2: Exploit Kitごとに固有の特徴遷移）

※主に、Kitの種類に依存

×

（要素3: 遷移の進行度）



=

計32件の検知ルール

ルール例

要素1: Javaの脆弱性を利用する

要素2: BlackHole Exploit Kitに対応

要素3: exploitステップまで進行

目次

1. 背景
2. 本研究で提案する検知方式
3. 定性的な特徴の遷移
4. 検証
5. まとめ

4. 今回実施した検証

今までに作成した，検知ルールの有用性を評価するため
検知・誤検知の観点から2種類の検証を実施

4-1. D3M Datasetを利用した検知の評価

4-2. 通信ログを利用した誤検知の評価

4. 今回実施した検証

今までに作成した，検知ルールの有用性を評価するため
検知・誤検知の観点から2種類の検証を実施

4-1. D3M Datasetを利用した検知の評価

4-2. 通信ログを利用した誤検知の評価

4.1 評価用データと前処理

■評価に利用したデータ
- D3M Dataset 2012, 2013
- 不正URLとの攻撃通信データ

■前処理

1. HTTPパケットデータを通信ログ形式に整形

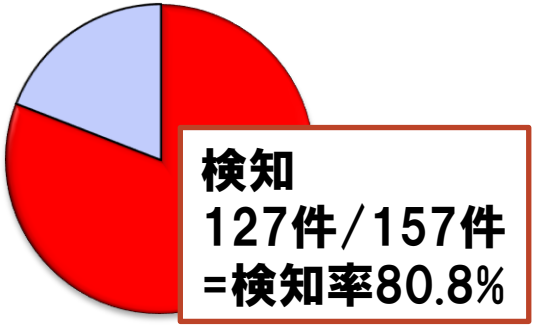
2. 1回あたりの感染によって，
   発生する一連のログを**感染ログ群**として集約

3. 本研究のスコープ外であるものを除外
   - ふるまい不明
   - 感染ステップを経ない単純な手法による感染ログ群
     例：exe,pdfなどを直接ダウンロード

感染ログ群157件を生成
→検知できたもの，できなかったものについて考察

4-1. 検知した感染ログ群の考察

３種類のExploit Kitに対応したルールで検知



■BlackHole用検知ルール
Ver 2.0以降の感染ログ群はすべて検知

■SweetOrange用検知ルール
BlackHoleの旧バージョン(ver1.X)による感染ログ群を検知

例： http://[domain]/showthread.php?t=d7ad916d1c0396ff

■Sakura用検知ルール
実際にはPhoenix Exploit Kitによる感染ログ群

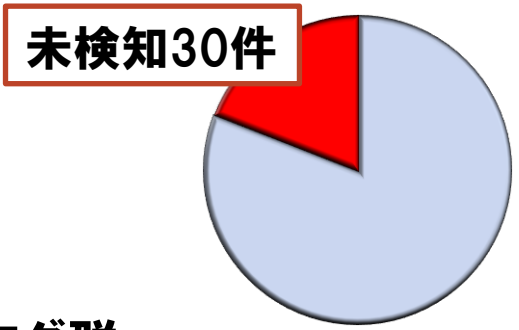
例: http://[domain]/navigator/jueoaritjuir.php
http://[domain]/navigator/dvjwbzawcojqjtx7.jar

想定していないExploit Kitによる感染ログ群も検知できた

4-1. 未検知 感染ログ群の考察

検知できなかった感染ログ群　30件

Bleeding Life，Elenore Exploit Kitなど
当初意図していなかったExploit Kitによるログ群



例：検知できなかったElenore Exploit Kitによる感染ログ群

http:// [domain] /docs/.q/
http:// [domain] /docs/.q/432.js
http:// [domain] /docs/.q/load.php?spl=mdac
http:// [domain] /docs/.q/?spl=2&br=MSIE&vers=6.0&s=
http:// [domain] /docs/.q/pdf.php

Exploit Kitに固有の特徴の遷移
（ファイルsuffix、webアプリケーションに渡す引数の有無）
が想定していた範囲ではカバーできていなかったため
検知できなかった

4. 今回実施した検証

今までに作成した，検知ルールの有用性を評価するため
検知・誤検知の観点から2種類の検証を実施

1. D3M Datasetを利用した検知の評価

2. 通信ログを利用した誤検知の評価

4-2. 誤検知評価

利用したログ
　約89,000ドメインへのアクセスログ （proxyログ約1200万行）

■誤検知が多かったルール　9ルール
- Java以外 (Adobe Readerなど) の脆弱性を利用し，SweetOrange, Sakura, Styx による感染を捉える検知ルール
- 誤検知数：平均10ドメイン

→正常なwebアクセスと類似しているため，誤検知が発生

■誤検知が少ないルール　23ルール
- Javaの脆弱性を突く感染を検知するルールでは，Kitによらず誤検知は少ない
- 誤検知数：平均3ドメイン

→exploit発生時のUserAgentの遷移が特徴的であるため

目次

1. 背景
2. 本研究で提案する検知方式
3. 定性的な特徴の遷移
4. 検証
5. まとめ

5. まとめ

DBD攻撃において定性的な特徴が遷移していくことを利用、
通信ログからふるまいを捉える検知ルールを作成
→検知・誤検知の観点から検証・考察

■課題

- Exploit Kitに固有の特徴が想定と異なっていると,
  検知できない攻撃がある.
  →より多くのExploit Kitの特徴を捉えるよう
  キャッチアップしていく
  →通信ログ以外のデータに現れる共通的な特徴も
  利用して汎用性の高いルールを作成

- Java以外の脆弱性を利用したexploitは,
  特徴が捉えづらく誤検知が多い
  →より多くの事例を収集・分析する

Global IT Innovator
NTT DATA Group

NTT DATA