Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

相補的な Web 感染型 マルウェア検知方式の提案

上松晴信[†]，名坂康平[††]，酒井崇裕[††]，西垣正勝[†††]

近年，Web サイト改ざんによってユーザを自動的にマルウェア配布サイトに誘導する Gumblar と呼ばれる攻撃手法が急増している．Gumblar は Drive-by-download 攻撃の一種であり，閲覧者をマルウェアに感染させることを目的とする．閲覧者が自身の Web サイトを開設していた場合には，マルウェアの感染によって閲覧者の Web サイト管理用パスワードが盗まれ，Gumblar を構成する改ざん Web サイトが増殖していく．この攻撃によるマルウェアの感染を防ぐためには，Web サイトのリダイレクトの有無を判別することが有効であると考えられる．しかし，Web サイトにおける正規のリダイレクトと不正なリダイレクトを確実に切り分けることは一般的に難しい．そこで本稿では，複数の Web サイトからのマルウェア配布サイトへのリダイレクトの集中を検査するグローバル探査と個々の Web サイトのリダイレクトの変更頻度の上昇を検査するローカル探査の併用による相補的な Web サイト改ざん検知方式を提案する．不正者は，Web サイトのリダイレクトを低頻度で改ざんする場合にはグローバル探査によって検知され，Web サイトのリダイレクトを高頻度で改ざんする場合にはローカル探査によって検知される．このため，Gumblar に関与する不正 Web サイトを効果的に検出することが期待される．

A proposal of complementary scheme for Web-based attack malware detection

HARUNOBU AGEMATSU[†]　KOHEI NASAKA[††]
TAKAHIRO SAKAI[††]　MASAKATSU NISHIGAKI[†††]

[†]静岡大学情報学部，〒432-8011 浜松市中区城北 3-5-1，
Faculty of　Informatics, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011 Japan
[††]静岡大学大学院情報学研究科，〒432-8011 浜松市中区城北 3-5-1，
Graduate school of Informatics, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011 Japan
[†††]静岡大学創造科学技術大学院，〒432-8011 浜松市中区城北 3-5-1，
Graduate School of Science and Technology, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011 Japan

Today, there are rapidly increasing numbers of attacks on networks by a type of malware called a ‘Gumblar’, which uses Website manipulation to automatically lead users to go to a malware download site. Gumblar is a type of “Drive-by-download attack” whose aim is to infect a reader with a malware. If a reader has their own Website, the password of the Website can be stolen by malware infection and the Website will be manipulated by Gumblar. To prevent infection by this type of attack, it is effective to determine the presence of a hidden redirect on Websites. But it is difficult to distinguish this type of malware redirect from regular Website update features. In this paper, we suggest a complementary Web-based attack malware detection scheme that combines global investigation (inspecting a concentration of redirect from multiple Websites to malware download sites) and local investigation (inspecting frequencies of redirect updates on several Websites). Crackers can be detected by global investigation when they manipulate Websites with low frequency, and by local investigation when they manipulate Websites with high frequency. So, this scheme is expected to detect the Gumblar-manipulated Websites effectively.

1. はじめに

正規の Web サイトを閲覧しただけで自動的に別の Web サイトに誘導され，誘導先でマルウェアがダウンロードされる，Gumblar とよばれる攻撃手法による被害が増加している．Gumblar は Drive-By-Download 攻撃[1][2]の一種であり，不正者は正規の Web サイトを改ざんし，正規サイトに訪れた閲覧者をマルウェア配布サイトへ誘導する．マルウェア配布サイトへの誘導はリダイレクトによって閲覧者に気づかれないように行われる．リダイレクトは，幾つかのサイトを経由する場合もある．誘導先のマルウェア配布サイトでは閲覧者の PC のアプリケーションの脆弱性が突かれ，閲覧者の PC にマルウェアがインストールされる．

Gumblar によって改ざんされた正規サイトに訪れた閲覧者が自身の Web サイトを開設していた場合には，マルウェアの感染によって閲覧者の Web サイト管理用パスワードが盗まれる．不正者はこのパスワードを使って，閲覧者の Web サイトをマルウェア配布サイトへの誘導サイトへと改ざんする．こうして Gumblar を構成する「改ざんされた正規の Web サイト」が増殖していく．

この一連の攻撃方法は，従来の CodeRed[3]や Nimda[4]に代表されるような能動的攻撃とは違い，閲覧者自身の Web アクセスを起点として不正者の用意したマルウェア配布サイトへ誘導するという受動的な攻撃[5]であるため，ファイアウォールによる防御が困難である．更に，閲覧者にとっては正規の Web サイトを閲覧しただけでマルウェ

Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

アに感染する形になるため，怪しいサイトにはアクセスしない等の閲覧者側の自衛策も機能しない．

この攻撃に対しては，正規の Web サイトのリダイレクトの有無を検知することが重要であると考えられる．しかし，Web サイトにおける正規のリダイレクトと不正なリダイレクトを確実に切り分けることは一般的に難しい．そこで本稿では，不正者の目的と心理を逆手にとることによって改ざんされた正規 Web サイトを検知する方法を提案する．

不正者は，なるべく多くの正規サイトを改ざんし，マルウェアに感染する閲覧者を増やしたい．ここで，不正者が正規サイトに直接マルウェアを埋め込むという戦略をとらず，正規サイトをマルウェア配布サイトへの飛び石として利用する理由は，マルウェアの管理を考えてのことだと推測される．不正者は，マルウェアの感染力を保つために感染に利用する脆弱性を定期的に更新したり，目的の変更に応じて流布するマルウェアを更新すると考えられる．正規サイトにマルウェアを埋め込む方法を採っていた場合は，不正者は今までに改ざんした Web サイトのすべてに定期的に侵入してエクスプロイトコードやマルウェアを更新しなければならず，その分，不正侵入や改ざんが探知される危険性が高まることになる．正規サイトからマルウェア配布サイトへリダイレクトする方法であれば，不正者はマルウェア配布サイトのエクスプロイトコードやマルウェアを変更するだけでよい．

この結果，複数の改ざんサイトのリダイレクト先が 1 つのマルウェア配布サイトへ集中するという特徴が現れることになる．そこで，クローラにインターネットを巡回させ，Web サイトからのリダイレクトが集中しているサイトをマルウェア配布サイトとして検出する．本検出法による検知から逃れるためには，不正者はすべての改ざんサイトにあえて定期的に侵入し，マルウェア配布サイトへのリダイレクトが設置されているサイトが各時刻においてはいずれか一つになるように，各改ざんサイトのリダイレクトを適切なタイミングで切り替える必要がある．すなわちこの場合は，それぞれの改ざんサイトにおけるリダイレクトがある程度の頻度で変更されることになる．そこで，Web サイトのリダイレクトが高頻度で変更される場合も，そのサイトを改ざんされた正規サイトとして検出する．

すなわち本方式は，複数の Web サイトからのマルウェア配布サイトへのリダイレクトの集中を検査するグローバル探査と個々の Web サイトのリダイレクトの変更頻度の上昇を検査するローカル探査の併用による相補的な改ざん Web サイト検知方式となっている．不正者は，正規 Web サイトのリダイレクトを低頻度で改ざんする場合にはグローバル探査によって検知され，正規サイトのリダイレクトを高頻度で改ざんする場合にはローカル探査によって検知される．このため，Gumblar に関与する不正 Web サイトを効果的に検出することが期待される．

2. Gumblar

Gumblar の一連の攻撃手法は Drive-by-Download 攻撃に分類される．Drive-by-Download 攻撃とは，Web ブラウザを通じて利用者に気付かれないようにマルウェアをダウンロードさせるものである．利用者が単に Web サイトを閲覧しただけで Web ブラウザ等のアプリケーションの脆弱性が悪用され，自動的にマルウェアがインストールされる．具体的には，(1)一般の正規 Web サイトが何らかの方法で改ざんされ，(2)閲覧者が改ざんされた Web サイトにアクセスすると，(3)そのアクセスがマルウェア配布サイトへリダイレクトされ（中継サイトを経由する場合もある），(4)閲覧者の PC にマルウェアがダウンロードされる（図 1）．

(3)のリダイレクト方法は，HTTP レスポンスのステータスコード 3xx（301: Moved Permanently，302: Found 等），HTML 文内の Meta タグや iframe タグ，Java スクリプトなどが利用される．最近の Gumblar では多重にリダイレクトされたり，リダイレクトを行うタグやスクリプト（スクリプトコード）が難読化されているなど，その手口が巧妙になっている．(4)のマルウェア感染においては，Web ブラウザの脆弱性だけでなく，Adobe Reader の脆弱性を悪用する PDF ファイルや，Flash Player の脆弱性を悪用する SWF（Small Web Format）ファイルなど，様々な関連アプリケーションの脆弱性が利用される．

Gumblar の被害が深刻となった原因の一つに，感染させたマルウェアを利用して FTP アカウントのパスワードを盗むという特徴がある．不正者が盗んだパスワードを利用して正規の FTP サイトを改ざんすることにより，Gumblar を構成する「改ざんされた正規の Web サイト」が増殖していく．Gumblar は閲覧者自身の Web アクセスを起点とした攻撃手法であるため，外部からの不正通信をブロックするファイアウォールでは防げず，イントラネット内の PC に感染被害を与えることが可能である．また，リダイレクトを用いて閲覧者に気づかれないようにマルウェア配布サイトに誘導されるため，閲覧者にとっては正規の Web サイトを閲覧しただけでマルウェアに感染する形になることから，怪しいサイトにはアクセスしない等の閲覧者側の自衛策も機能しない．

特に，2009 年末から，日本国内の著名サイトが改ざんの被害にあい，不特定多数のユーザに対して大きな感染被害を及ぼした[6][7][8].

Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

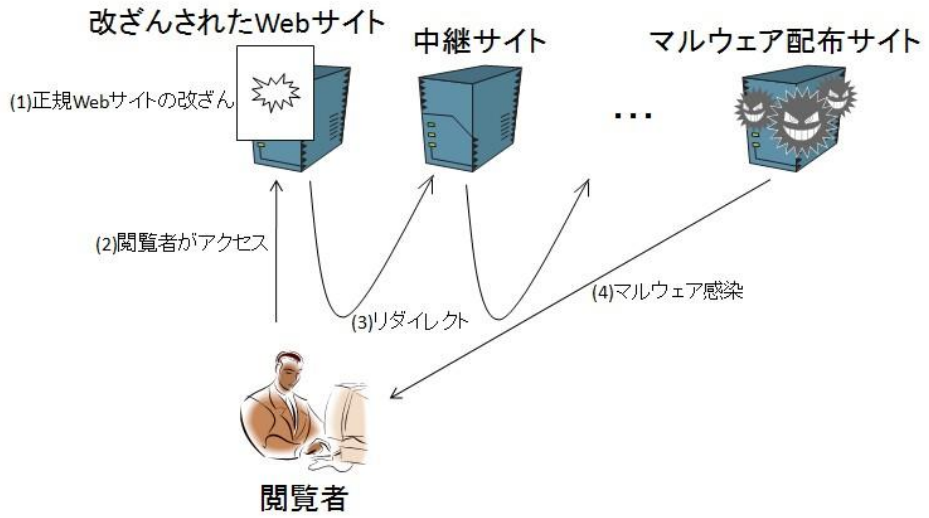

図 1 Drive-by-Download 攻撃
Figure 1 Drive-by-Download attack

3. 関連研究

Gumblar を防ぐためにはマルウェア感染の元となる「改ざんされた正規の Web サイト」を検知することが重要である．改ざんサイトの検知を目的とした既存研究の 1 つとして，ヒューリスティック検知[9]が提案されている．この方式では，改ざんサイトに記載されるリダイレクトコードをブラックリスト化し，それらのコードの有無によってサイトが改ざんされているか否かを判別する．この方式は既知のリダイレクトコードに対しては有効であるが，リダイレクトコードの難読化や，新種のリダイレクトコード（ゼロデイ攻撃）によってマルウェア配布サイトに誘導されている場合には対応できない．

不正なリダイレクトを検知する研究[10]もされている．この研究では，Web サイトにアクセスした際に閲覧者の PC に返される HTTP レスポンスの様子から，マルウェア配布サイトへのリダイレクトを検出する．具体的には，機械学習の決定木学習手法を用いて，マルウェア配布サイト（危険なファイルのダウンロードに至る URL）の判別ロジックを導出している．

また，Gumblar だけでなく Web 感染型マルウェア全般を対象とした対策として，アクティブハニーポット[11]が研究されている．アクティブハニーポットはインターネット上の膨大な Web サイトに自動的にアクセスし，その際に Web サイトからダウンロードされるプログラムを収集する．ダウンロードされたプログラムがマルウェアであれば，プログラムをダウンロードさせた Web サイトをマルウェア配布サイト，そのサイトへ誘導した Web サイトを改ざん Web サイトとして検知する．アクティブハニーポットでは，ヒューリスティック検知[4]の不得手であったスクリプトコードの難読化や新種のコードによるマルウェア配布サイトへの誘導についても検知できる．しかし，この方式の性能は，ダウンロードされたプログラムがマルウェアかどうかを判別する部分のマルウェア検知精度に依存することになる．よって，例えばダウンロードされたプログラムが未知のマルウェアであった場合などには，マルウェア配布サイトや改ざん Web サイトを検知できないといった問題点がある．

4. 提案方式

Gumblar においては，複数の改ざん Web サイトからマルウェア配布サイトへのリダイレクトの集中という特徴が現れると考えられる．そこで我々は，改ざん Web サイトあるいはマルウェア配布サイトの特徴を個々に捉えるのではなく，改ざん Web サイトとマルウェア配布サイトのインターネット上のトポロジに焦点を当てたグローバルな Gumblar 探査方式を提案する．ただし，このグローバルな探査方式だけでは，不正者がこれを回避することが可能である．そこで，これに対処するため，改ざん Web サイト単体の特徴を用いたローカルな Gumblar 探査方式を併用する．改ざん Web サイトとマルウェア配布サイトのトポロジを利用する点，グローバルな探査方式とローカルな探査方式の併用による相補的な検知方式を実現している点が，提案方式の特長である．

4.1 グローバル探査

通常，不正者にとってパスワードを奪取した正規 Web サイトに何度もアクセスすることはリスクが伴う．Web サイトへの侵入を繰り返す内に，正規の管理者によって不正アクセスが発見される可能性が増加し，最悪の場合は通信履歴から Web サイトへの侵入の際に利用している PC の所在が探知されて逮捕にまで及ぶかもしれないためである．よって不正者は，パスワードを奪取した正規 Web サイトにアクセスして，一旦，マルウェア配布サイトへのリダイレクトを設置したら，後はそのまま改ざん Web サイトを放置することが多いと考えられる．この不正者の心理に基づいた仮定の上で，改ざん Web サイトからマルウェア配布サイトへの不正なリダイレクトと Web サイトの移転等の際に用いられる正規のリダイレクトとを判別する方法を提案する．

不正者によって改ざんされた正規の Web サイトは，通常リダイレクトを用いて悪性サイトへの誘導を行う．リダイレクトとは，Web サイトの閲覧において，指定した Web サイ

Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

トから自動的に他の Web サイトに転送させる機能である．リダイレクトは主に，Web サイトの管理者が Web ページの移転を行った際に利用されるため，通常のリダイレクトにおいては, リダイレクト元とリダイレクト先で 1 対 1 の関係が成立する（図 2）. 一方, Gumblar では, 不正者は感染を拡大させるために多数の正規 Web サイトを改ざんすることを試みる．これに対し，マルウェア配布サイトの数は，サイトの準備の手間，および，マルウェアの管理（利用するエクスプロイトコードの更新やダウンロードさせるマルウェアの変更）の容易さなどの観点から，一般に少ないと考えられる．このため，多数の改ざん Web サイトから少数のマルウェア配布サイトへのリダイレクトが集中することになる．すなわち，マルウェア配布サイトと改ざん Web サイトのリダイレクトの間には 1 対多の関係が成り立つ（図 3）.

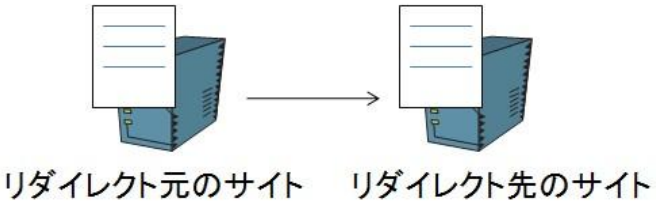

図 2 通常のリダイレクト
Figure 2 Normal redirect

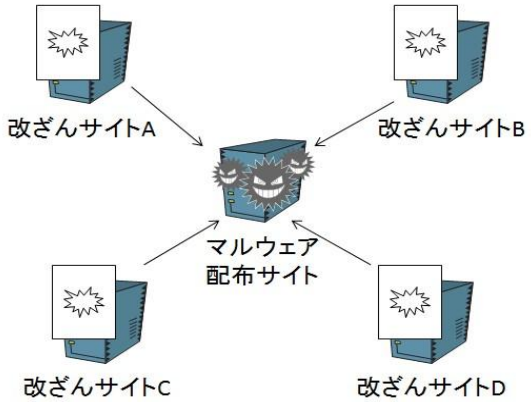

図 3 Gumblar におけるリダイレクト
Figure 3 Wrong redirect by Gumblar

そこで，この通常の Web サイトにおける正規のリダイレクトと改ざん Web サイトにおける不正なリダイレクトのトポロジの違いを利用し，改ざんされた正規の Web サイトを検出する．具体的な検査手順を以下に示す.

(1). Web サイト巡回プログラム（クローラ）を用いてインターネット上の Web サイトを巡回し，リダイレクトを検出する.
(2). リダイレクト元 URL（または IP アドレス）とリダイレクト先 URL（または IP アドレス）を抽出し，集計する.
(3). 複数のサイトからのリダイレクト先となっている Web サイトをマルウェア配布サイトであるとみなし，また，そのサイトへのリダイレクト元となっている Web サイトをすべて改ざん Web サイトであると判断する.

本方式では，クローラを用いて動的に Web サイト間のリダイレクトを検出し，リダイレクト元とリダイレクト先の関係のみを用いて不正サイトの判別を行う．そのため，リダイレクトの方法や感染するマルウェアの種類に依らず，不正サイトを検出することができる．このため, リダイレクトコードの難読化や新種のコードによるマルウェア配布サイトへの誘導であっても, 改ざん Web サイトおよびマルウェア配布サイトを検知することが可能である.

本稿では，この方式をグローバル探査と呼ぶこととする.

4.2 ローカル探査

グローバル探査による検知が既知になると, 不正者はグローバル探査を回避しようとすると考えられる．グローバル探査 を回避するには，マルウェア配布サイトと改ざん Web サイトの間のリダイレクトの関係が常に 1 対 1 になるようにすればよい．よって不正者は，あえて，すべての改ざん Web サイトに定期的に侵入し，マルウェア配布サイトへのリダイレクトが設置されているサイトが各時刻においてはいずれか一つになるように，各改ざんサイトのリダイレクトを適切なタイミングで切り替えるという方法を採ると考えられる（図 4）.

Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

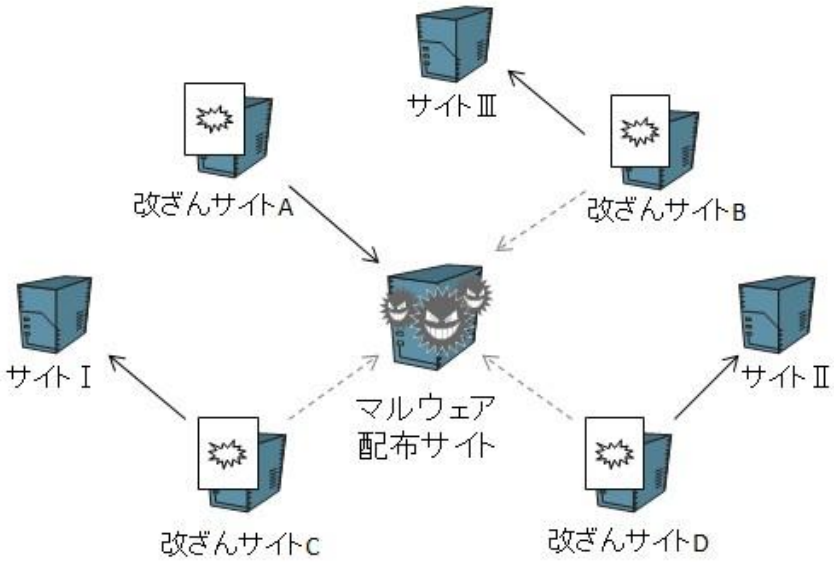

図 4 リダイレクトの切り替え
Figure 4 Change redirect

一方，通常の Web サイトにおける正規のリダイレクト（Web サイトの管理者が Web ページの移転を行った際に利用されるリダイレクト）は，リダイレクト先 URL の打ち間違いがあった場合や，何らかの理由で Web ページの移転が重なった場合にリダイレクトの変更が生じることになるが，変更が何度も繰り返されるようなことは稀であることが一般的である．

そこで，この通常の Web サイトにおける正規のリダイレクトと改ざん Web サイトにおける不正なリダイレクトの変更頻度の違いを利用し，改ざんされた正規の Web サイトを検出する．具体的な検査手順を以下に示す．

(1). Web サイトを正規に更新した時点で Web サイトのソースコードを保存するとともに，Web サイトの変更を常に監視する．
(2). Web サイトの変更があった場合には，変更後の Web サイトのソースコードを前回の正規更新時に保存したソースコードと比較する．
(3). リダイレクト先 URL の変更があった場合は，変更回数カウンタの値をインクリメントする．
(4). ある決められた時間間隔の間にある決められた回数以上のリダイレクト先 URL の変更が観測された場合は，そのサイトを改ざん Web サイトとして検知する．また，そのサイトのリダイレクト先の Web サイトをマルウェア配布サイト（またはマルウェア配布サイトへの中継サイト）であると判断する．

本稿では，この方法をローカル探査と呼ぶこととする．

4.3 グローバル探査とローカル探査の併用による相補的な検知

複数の Web サイトからのマルウェア配布サイトへのリダイレクトの集中を検査するグローバル探査と個々の Web サイトのリダイレクトの変更頻度の上昇を検査するローカル探査を併用することによって，相補的な Gumblar 検知を実現することが可能となる．不正者が不正侵入や改ざんの露見を恐れ，パスワードを奪取した正規の Web サイトに対してマルウェア配布サイトへの静的なリダイレクトを設置した場合には，マルウェア配布サイトへのリダイレクトが集中することになり，インターネット上のリダイレクトに関するトポロジチェックであるグローバル探査によって改ざん Web サイトやマルウェア配布サイトが検知されてしまう．逆に，不正者がグローバル探査を逃れるためにあえて不正侵入や改ざんの露見のリスクを犯し，各時刻においてリダイレクト元とリダイレクト先の関係が 1 対 1 となるように改ざん Web ページのリダイレクトを適切に切り替えた場合には，個々の Web サイトのリダイレクトの変更頻度が上昇することになり，ローカル探査によって改ざん Web サイトが検知されてしまう．このようにグローバル探査とローカル探査が相補的に作用することで，Gumblar に関与する不正 Web サイト（マルウェア配布サイトや改ざんされた正規 Web サイト）を効果的に検出することが期待される．

5. まとめと今後の課題

本稿では，複数の Web サイトからのマルウェア配布サイトへのリダイレクトの集中を検査するグローバル探査と個々の Web サイトのリダイレクトの変更頻度の上昇を検査するローカル探査の併用による Web サイト改ざん検知方式を提案した．本方式は，グローバル探査とローカル探査が相補的に作用し，Gumblar に関与する不正 Web サイトを効果的に検出できると期待される．

今回の方式はリダイレクトを用いてマルウェア配布サイトへ誘導するタイプの Gumblar を対象とした不正 Web サイト検知方法となっているが，改ざん Web サイトにリダイレクトを設置するのではなく，閲覧者自身のクリックによってマルウェア配布サイトに誘導するタイプの Gumblar も存在すると考えられる．このような不正サイトは，リダイレクトのトポロジチェックに基づく本グローバル探査が機能しない．今後はページランクや SEO などの観点から，リダイレクトを用いないタイプの Gumblar のサイトを発見する方法を検討していきたい．

Vol.2011-DPS-146 No.53
Vol.2011-CSEC-52 No.53
2011/3/11

参考文献

1) Marco Cova, Christopher Kruegel, Giovanni Vigna: Detection and analysis of drive-by-download attacks and malicious JavaScript code, Proceedings of the 19th international conference on World wide Web, New York, 2010
2) Manuel Egele, Peter Wurzinger, Christopher Kruegel, Engin Kirda: Defending Browsers against Drive-by Downloads: Mitigating Heap-spraying Code Injection Attacks, Lecture Notes in Computer Science, Volume 5587/2009, pp.88-106, 2009
3) David Moore, Colleen Shannon, k claffy : Code-Red: a case study on the spread and victims of an internet worm, Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement, New York, 2002
4) James Cowie, Andy T. Ogielski, Bj Premore, Yougu Yuan: Global Routing Instabilities Triggered by Code Red II and Nimda Worm Attacks, Tech. Rep, Renesys Corporation, December 2001, http://www.renesys.com/projects/bgp_instability
5) IPA：”脆弱性を利用した新たなる脅威の分析による調査”, http://www.ipa.go.jp/security/vuln/report/documents/newthreat_report_2010.pdf
6) ASCII：”大手サイトも驚愕させた，ガンブラーはどう動く?”, http://ascii.jp/elem/000/000547/547757/
7) JPCERT：“踏み台にされる Web サイト～いわゆる Gumblar の攻撃手法の分析調査”, http://www.jpcert.or.jp/research/#Webdefacement
8) IPA：”Web 媒介型攻撃 Gumblar の動向調査”, http://www.ipa.go.jp/security/fy21/reports/tech1-tg/b_05.html
9) 田中達哉，田村佑輔，甲斐俊文，佐々木良一：”改ざんサイト自動検知システム DICE の開発と評価”，コンピュータセキュリティシンポジウム 2010 論文集，2010.10
10) 寺田剛陽，古川忠延，東角芳樹，鳥居悟：”検知を目指した不正リダイレクトの分析”，コンピュータセキュリティシンポジウム 2010 論文集，2010.10
11) 株式会社フォティーンフォティ技術研究所：”「人柱型」アクティブ・ハニーポット”, http://www.fourteenforty.jp/products/origma/