社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2007－DPS－130 (16)
2007－CSEC－ 36 (16)
2007／3／1

Web応答と時事情報を組み合わせた観測システムの提案

寺田真敏[†1] 枝村和茂[†2] 高橋正和[†3] 有村浩一[†4]

[†1)](株)日立製作所 システム開発研究所
〒212-8567 神奈川県川崎市幸区鹿島田 890
[†2)] 海上保安庁
〒100-8918 東京都千代田区霞が関 2-1-3
[†3)] マイクロソフト株式会社
〒151-8583 東京都渋谷区代々木 2-2-1 小田急サザンタワー
[†4)] 財団法人日本データ通信協会 テレコム・アイザック推進会議 (Telecom-ISAC Japan)
〒107-0052 東京都港区赤坂 2-17-28

概要：Webサイトに対するアクセス集中には，悪意ある第三者がそのサイトのサービス運用を妨害することを意図したDoS攻撃と，注目を集める事象が発生することにより，多数の正規のユーザが同時刻帯にサービスを利用することで該当Webサイトへのアクセスが集中するものがある．本稿では，Webサイトを狙ったDoS攻撃をWebサーバの応答時間から観測するにあたり，正規のユーザによるアクセス集中と悪意あるDoS攻撃によるアクセス集中とを判別するため，Webサーバの応答時間観測とニュースサイトに掲載された記事観測とを組み合わせた観測手法を提案する．また，提案手法を実装したプロトタイプシステムの試行運用を通して，得られた知見について調査事例と共に報告する．
キーワード：Web応答，時事情報，観測

Proposal of monitoring system of combined Web response and current events.

Masato Terada[†1] Kazushige Edaumura[†2] Masakazu Takahashi[†3] Koichi Arimura[†4]

[†1)] System Development Lab. Hitachi Ltd.
890 Kashimada, Saiwai-ku, Kawasaki, Kanagawa, 212-8567 Japan
[†2)]. Japan Coast Guard
2-1-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8918 Japan
[†3)]. Microsoft Corporation
2-2-1 Yoyogi, Shibuya-ku, Tokyo 151-8583 Japan
[†4)]. Telecom-ISAC Japan
2-17-28 Akasaka, Minato-ku, Tokyo 107-0052 Japan.

Abstract:. Malicious users and legitimate users intend the access concentration to the Web server. In a denial-of-service (DoS) attack, a malicious users attempt to prevent legitimate users from accessing Web site. And, legitimate users attempt to access the Web site at same time by the matter, which attracts attention occurring. In this paper, we examine how we can detect the access concentration by legitimate users from remote site. Our proposal is to combine monitoring technique between the social event observation and the Web response observation. Also we implemented a prototype system to show the validity of our approach.
Key words: Web response, Topic, Monitoring

1 はじめに

Webサイトのダウンや応答時間の状況を把握するため，Webサーバからの応答監視やWebサーバがクライアントからの要求に応答するまでの時間を測定するという手法が用いられている．その多くは，一日に何度か測定をおこなって時間帯による表示時間の変化などを知るなど，ユーザの快適な環境を保つために利用されている．このWebサーバの応答時間観測は，Webサイトを狙ったDoS攻撃をリモート観測する手法としても応用することができる．これは，急激な応答時間の劣化が発生した場合，Webサイトに対してなんらかのアクセスが集中したことが考えられるためである．

このようなアクセス集中には，大きく2つの要因が想定される．ひとつめは，悪意ある第三者がそのサイトのサービス運用を妨害することを意図したDoS攻撃である．DoS攻撃は，一見多くの送信元からの一見正常である通信をつくり，攻撃先に大量に送りつけることで成立する攻撃であり，TCP接続の

呼を意味するSYNパケットを攻撃先に大量に送付することで，ある負荷を与えるTCP SYN flood攻撃[1]などがある．このようなDoS攻撃は，ネットワークデモと言われる形で実現される場合もある．これは，掲示板に攻撃対象とするサイト名と時刻を掲示し，大量のユーザが同一時刻帯に同一Webサイトにアクセスすることを意図的に促すという流れを通して攻撃先のサーバや回線の負荷を高める方法である．ふたつめは，社会的な現象に基づくもので，注目を集める事象が発生することにより，該当Webサイトへのアクセスが集中するというものである．これは，正規の多数のユーザが，同時刻帯にアクセスを行なうという事象であり，インターネットにおける正常な活動である．

本研究の目的は，Webサイトを狙ったDoS攻撃をWebサーバの応答時間から観測するにあたり，正規のユーザによるアクセス集中と悪意あるDoS攻撃によるアクセス集中とを判別する観測手法を検討することにある．本稿では，この判別する観測手法のひとつとして，Webサーバの応答時間観測（以降，Web応答観測）とニュースサイトに掲載された記事観測（以降，時事情報観測）とを組み合わせた手法を提案する．また，提案手法を実装したプロトタイプシステムの試行運用を通して，得られた知見について調査事例と共に報告する．

2 関連研究

本章では，Web応答観測と時事情報観測それぞれの関連研究について述べる．

2.1 Web応答観測

トラフィックを直接観測することでDoS攻撃を検出する研究については，通常トラヒックと異常トラフィックの性質の差に着目しDoS攻撃を検出する手法が提案されている[2]．また，トラフィックを間接的に観測する手法については，文献3)ではTCPのSYN flood攻撃において，攻撃用のパケットが攻撃対象となったサーバから跳ね返され，インターネット上にばらまかれる現象の発生に着目したDoS攻撃用の観測システムを報告している．文献4)では測定用パケットの遅延を利用してDoS攻撃を検出する試みが行なわれている．Webサーバからの応答に着目した研究については，文献5)でコンテンツの特徴や変更前後の変化を解析し，改竄を検出するWebサーバリモート監視の際に，応答時間を測定することで，DoS攻撃の発生有無を確認できるとしている．なお，Webサーバの応答観測については，一定時間毎に応答時間を監視する技術が製品化やサービス化されている．

2.2 時事情報観測

時事情報の観測のうち，ネット風評監視については，文献6)では，掲示板やWebページに書き込まれた情報を収集した後，解析を行なう技術ならびにサービスが報告されている．文献7)では，blogを掲示板と同様の情報源として定期的に監視し，そこから情報を抽出するシステムについて言及している．また，代表的な検索ポータルサイト[8][9]では，各サイトから収集した最新ニュースを一覧化して提供するサービスも行なわれている．

2.3 解決したい課題

本研究で解決したい課題は，下記の通りである．

- Webサイトに対するDoS攻撃発生の可能性をリモートから効率的に検知する．
  リモートからの検知は，組織間でDoS攻撃の発生情報を共有する手法のひとつとして有効であると考えられる．
- 上記の検知において，正規のユーザによるアクセス集中と悪意あるDoS攻撃によるアクセス集中とを判別する．
  社会現象を考慮した観測は，事象解析に役立つと考えられる．

しかし，既存技術ならびにサービスだけでは，これらの課題を解決することができない．そこで，次に示す方法を提案する．

- Web応答観測：DoS攻撃発生の可能性をリモートから検知する．
  一定時間毎にWebサーバの応答時間を観測する．Webサーバの応答時間の劣化が発生した場合，DoS攻撃や正規ユーザのアクセス集中が発生した可能性があると判断する．
- 時事情報観測：正規ユーザによるアクセス集中の可能性を判別する．
  一定時間毎に主要ニュースサイトに掲載されている記事を収集する．Webサーバの応答時間の劣化が発生した場合，その該当サイトに関連する時事情報が掲載されている場合には，正規ユーザによるアクセス集中の可能性を検討する．

Web応答観測と時事情報観測とを組み合わせることにより，Webサーバの応答時間が劣化した場合，その劣化が社会的な現象に誘発されたものか否かの判定に役立てることができる．

3 Webサイト観測システム

本節では，Web応答観測と時事情報観測を組み合わせた手法の有効性を検証するために実装したプロトタイプシステムについて述べる．

3.1 システム構成

Webサイト観測システムは，次の3つのコンポーネントから構成する(図 1)．

(1) 応答時間観測機能

Webサーバの応答時間を測定し，遅延が発生している場合には，アラート情報として，サイト名，応答時間などを管理者にメール通知する．

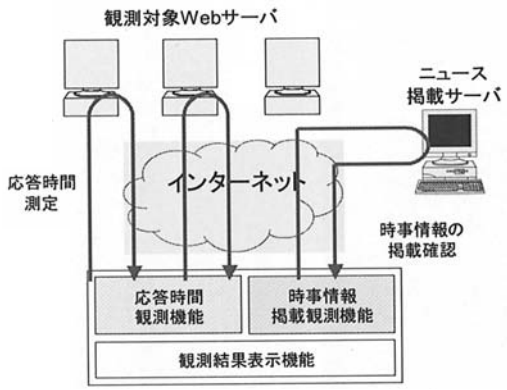


図 1　プロトタイプシステムの構成

(a) 応答時間測定

観測対象とする Web サーバに定期的(試行運用では 5 分に一回)に HTTP 要求を送信し，HTTP 応答を受信するまでの時間を測定する．

(b) 応答時間の遅延検出

下記のいずれかの事象が 2 回続いた場合，遅延が発生したと判断する．

- 応答時間が 10 秒を超えた場合
- “(応答時間－平均時間)÷標準偏差”の値が，4.0 を超えた場合

  なお，平均時間，標準偏差は，最新の一週間のデータを元に算出する．

(2) 時事情報掲載観測機能

応答時間観測機能が応答時間の遅延を検出した場合に，そのサイトに関連する記事の有無を確認する．記事が存在する場合には，補足情報として，サイト名，記事名などを管理者にメール通知する．

(a) 時事情報掲載確認

掲載確認対象とするニュースサーバに定期的(試行運用では 30 分に一回)にアクセスして記事を収集する．収集の際に実施する作業は次の通り．

- HTML を解析し，記事名を記事掲載 URL と共に格納する．
- HTML を解析し，URL リンクを抽出する．抽出した URL リンクと観測対象とする Web サーバの URL とを比較し，合致する場合には，記事と Web サイトとの関連付け情報として格納する．
- トップページやトピックスページに，収集した記事が掲載されているか否かを確認し，その結果を格納する．

(b) 正規ユーザによるアクセス集中の可能性判定

応答時間観測機能で応答時間の遅延が検出され，さらに，ニュースサイトのトップページやトピックスページに，該当 Web サイトに関する記事が掲載されている場合，注目を集める事象が発生したことにより，該当 Web サイトへのアクセスが集中した可能性があると判断する．

(3) 観測結果表示機能

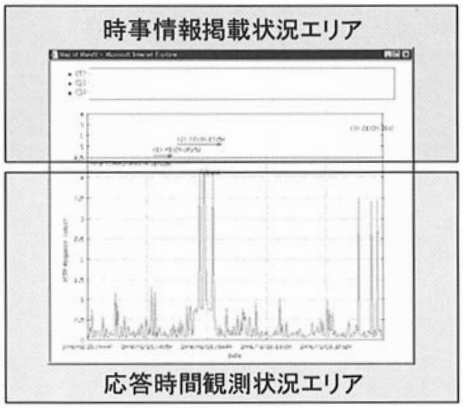


図 2　観測結果表示画面

観測結果表示機能は，Web 応答観測と時事情報掲載との関連を時間軸上で示すために，グラフ化する機能である(図 2)．図 2上段の時事情報掲載情報エリアでは，記事名とその記事がトップページやトピックスページに掲載されていた期間を示す．また，下段の応答時間観測状況エリアでは，Web 応答時間の推移と共に，応答時間の遅延を検出した箇所に“□”を発生時刻と共に示す．

4　観測システムの評価

本章では，プロトタイプシステムの試行運用の結果について述べる．

(1) 試行運用環境

試行運用は，2006 年 3 月から 2006 年 10 月までの 8 ヶ月間，Web 応答観測対象サイト：約 80 箇所，時事情報掲載確認サイト：1 箇所を対象とする Web サイト観測システム 1 台を稼動させた．

(2) 評価項目

試行運用を通したプロトタイプシステムの評価は，次の視点で行なった．

- 正規ユーザによるアクセス集中の可能性ありと判定した件数を算出する．
- 上記件数をさらに，Web サイトに関連付けられた記事の妥当性を 3 段階(○：適切，×：不適切，？：判定できず)で評価する．

(3) 結果

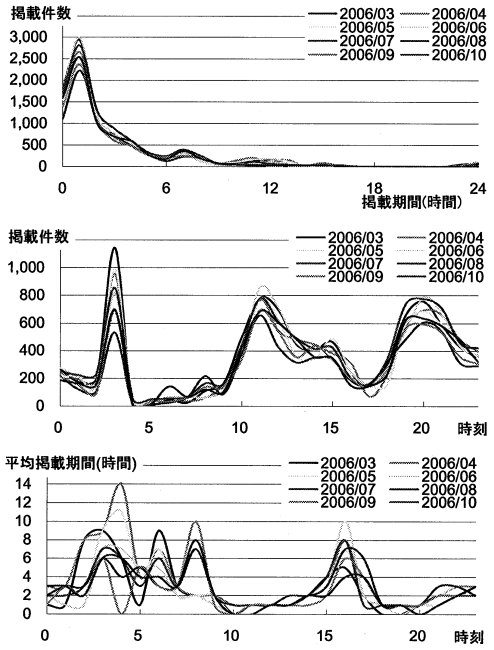
観測システムの発行したアラート数を表 1に示す．Web 応答アラート数(a)は，応答時間観測機能が遅延を検出して発行したアラート件数，掲載記事数(b)は，時事情報掲載観測機能が Web応答観測対象サイトに関連付けた記事数，時事情報アラート数(c)は正規ユーザによるアクセス集中の可能性ありと判定して発行したアラート件数である．比率(d)は全アラート件数に対する正規ユーザによるアクセス集中の可能性ありと判定した比率で，平均すると約32%となった．記事関連付けの妥当性(e)は Web サイトに関連付けられた記事の妥当性の評価であり，正規ユーザによるアクセス集中の可能性ありと判定したうち，記事

内容としても“○：適切”であった比率は平均すると約24%となった．

(4) 考察

試行運用では，トップページやトピックスページに記事が掲載されている期間を記事への注目度が高い期間と仮定し，アクセス集中の可能性判定に利用している．

試行期間中の掲載記事について，トップページやトピックスページに掲載されている期間，掲載開始時刻ならびに，掲載開始時刻と掲載期間を調査した結果を図 3に示す．利用した時事情報掲載確認サイトでは，記事掲載期間は約2時間が多く(図 3上段)，掲載開始時刻は明け方(3時)，昼(11時)，夜(20時)にピークが見られる(図 3中段)．また，昼間の記事掲載は2時間前後と短い(図 3下段)．このような掲載条件は，注目度に関する重み付けパラメタとして利用できると考えられる．



上段：記事掲載期間
中段：記事掲載開始時刻
下段：記事掲載開始時刻毎の平均記事掲載期間

図 3　記事掲載開始時刻と記事掲載期間

5　観測システムを用いた事例調査

本章では，観測システムで捕捉できた事例と捕捉できなかった事例について述べる．

5.1　観測システムで捕捉できた事例

観測システムで捕捉できた事例として，2006年10月16日に発生した海難事故を取り上げる．

(1) 事象

- 2006年10月16日午前1時頃，静岡下田沖でタンカーと貨物船が衝突して重油が流出した事件が発生した．
- 午前9時頃，ニュースサイトに関連記事が掲載された．
- 午前9時から10時にかけ，観測システムが該当Webサーバの応答時間遅延を検知した．

(2) 観測システムの状況(図 4)

- アラート件数：2件 (9:25,10:10)
- 関連記事収集件数：11件
- 正規ユーザによるアクセス集中の判定：図 4中の記事番号2“静岡・下田沖でタンカーと貨物船が衝突，重油が流出”が該当記事であり，可能性ありと判定した．

(3) 考察

- 応答時間遅延と送受信トラフィック

本事例の場合，Web応答時間と送受信トラフィックは同期性がみられる(図 5)．このことから，サイトによっては，Webサーバの応答時間遅延を観測することにより，DoS攻撃や社会的な現象に基づく正規ユーザのアクセス集中を検知できる可能性がある．

- Webアクセス数と応答時間

アラートが報告された9時台のWebアクセス数とWeb応答時間に同期性がみられる(図 5)．また，Webアクセスへの誘導経路(図 6)を調べると，検索ポータル系サイトから誘導され(図 6中段)，さらに，そのサイトの記事掲載から誘導されている(図 7)．このことから，事象によっては，社会的な現象に基づく正規ユーザのアクセス集中の判定に，時事情報掲載観測を利用できると言える．

5.2　観測システムで捕捉できなかった事例

観測システムで捕捉できなかった事例として，Winnyに関する報道が行なわれた2006年3月15日のTelecom-ISAC Japan(https://www.telecom-isac.jp/)のWebサイトを取り上げる．

表 1　観測システムの発行したアラート件数

項目		3月	4月	5月	6月	7月	8月	9月	10月
a)Web応答アラート数		1,454	1,262	1,753	1,081	1,011	2,736	1,016	1,003
b)掲載記事数		8,942	8,316	8,076	8,587	9,113	6,789	7,217	7,753
c)時事情報アラート数		547	412	784	417	564	407	241	114
d)比率(=c/a)		38%	33%	45%	39%	56%	15%	24%	11%
e)記事関連付けの妥当性[*]	○	334(23%)	289(23%)	503(29%)	456(45%)	359(13%)	210(21%)	78(8%)	373(35%)
	×	108	64	211	62	13	17	18	25
	?	105	59	70	46	35	14	18	19

[*] Webサイトに関連付けられた記事の妥当性 ○：適切，×：不適切，？：判定できず

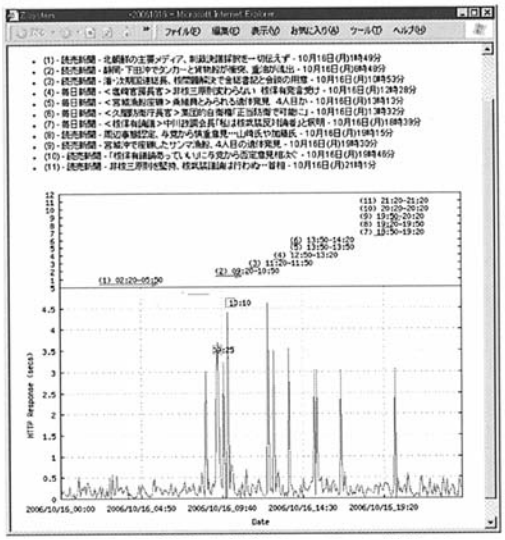
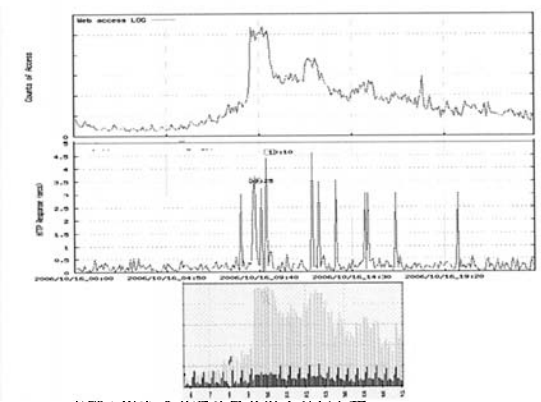


図 4　事例 1：観測システムの結果



上段：Web アクセスログのイベント数
中段：Web 応答時間
下段：送受信トラフィック

図 5　事例 1：アクセス数と送受信トラフィック

(1) 事象

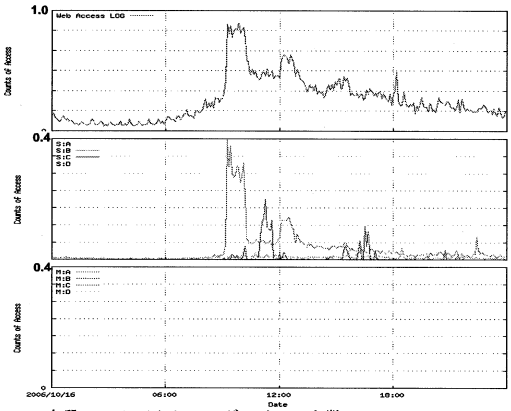
- 3 月 15 日午前中に，“Winny を介した情報漏えいについて”の官房長官発表が発表された[10].
- Telecom-ISAC Japan では，“ISP との連携による ANTINNY ウイルス感染ユーザへの注意喚起の取り組み”に関するニュースをリリースした[11].

(2) 観測システムの状況(図 8)

- アラート件数：4 件 (9:40,11:00,22:40,22:50)
- 関連記事収集件数：0 件
- 正規ユーザによるアクセス集中の判定：可能性なし

(3) 考察

プロトタイプシステムで実装した時事情報掲載確認による関連付け機能は，掲載記事に直接関係性が記載されている Web サイトと関連付けを行なう．



上段：Web アクセスログのイベント数
中段：ポータル系の Referrer 数
下段：一般メディア系 Referrer 数

図 6　事例 1：Web アクセスへの誘導経路

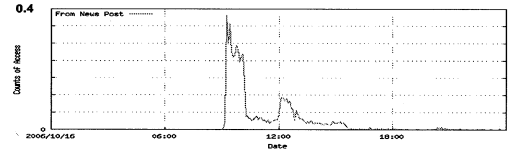


図 7　事例 1：記事掲載からの誘導

本事例の場合，掲載記事に直接関係性が記載されていなかったこと，TV 報道により Web サイトへの誘導が発生したことが，観測システムで捕捉できなかった理由と考えられる．

- Web アクセス数と TV 報道

図 9に示す通り，応答時間と Web アクセス数の同期性はあまりないが，Web アクセス数には何回かのピークが見られる．調査したところ，図 9の上段に示す Web アクセスログのイベント数のピークと中段の NHK ニュース報道時刻がほぼ一致していることがわかった．この結果から，NHK ニュース報道が Telecom-ISAC Japan サイトへのアクセス数増加に寄与していると考えることができ，本事例は，TV 報道から誘導され Web アクセス数が増加した例と言えるであろう．

- 社会的な事象に関する Web アクセスへの誘導経路

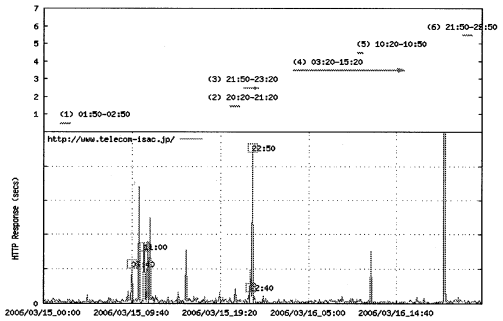
図 10に示す通り，Web アクセス数のピークとポータル系の誘導ピークがほぼ同期していること，特に検索からの誘導が多いことから(図 11)，TV 報道，検索という流れが Telecom-ISAC Japan サイトへのアクセス数に繋がっていると推定できる．

6　おわりに

本稿では，Web サーバの応答時間観測とニュースサイトに掲載された記事観測とを組み合わせた観測手法を提案すると共に，プロトタイプシステムの試行運用を通して得られた知見について報告した．

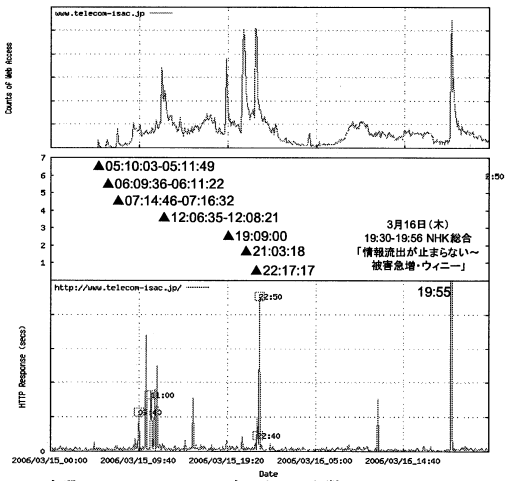
- 試行運用を通して，Web サーバの応答時間観測とニュースサイトに掲載された記事観測とを組み合わせにより，正規ユーザによるアクセス集中を推定する手段のひとつとなることを示した．
- 事例調査の結果，Web サーバの応答時間観測と TV 報道観測を組み合わせにより，捕捉可能な範囲が広がる可能性があることを示した．

今後の課題は，観測システムを用いた DoS 攻撃の発生検出の検証，時事情報掲載確認による関連付け機能の拡張など，実運用を想定した観測システムの検討があげられる．



注：時事情報掲載確認による関連付けでは，該当する記事なし．このため，時事情報掲載情報エリアには，関連記事の掲載開始終了時刻をマニュアルで追加している．

図 8　事例 2：観測システムの結果

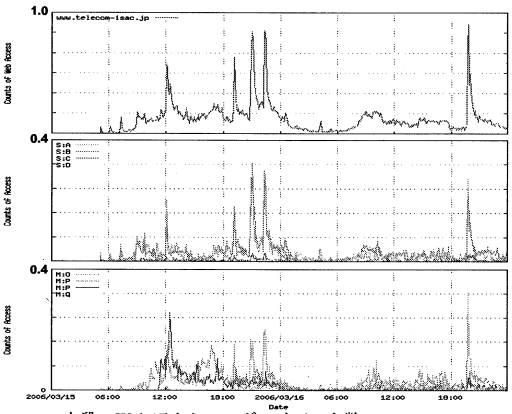


上段：Web アクセスログのイベント数
中段：NHK ニュース報道時刻
下段：Web 応答時間

図 9　事例 2：Web アクセス数と TV 報道

謝辞

本研究の一部は独立行政法人情報通信研究機構からの委託研究である「広域モニタリングシステムに関する基盤技術の研究開発」の支援を受け実施している．本研究を進めるにあたって有益な助言と協力を頂いた関係者各位に深く感謝致します．



上段：Web アクセスログのイベント数
中段：ポータル系の Referrer 数
下段：コンピュータメディア系 Referrer 数

図 10　事例 2：Web アクセスへの誘導経路

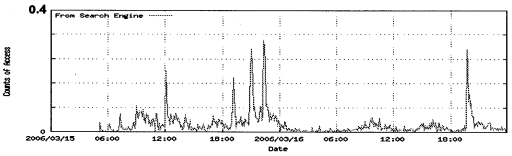


図 11　事例 2：検索から誘導

参考文献

[1] CERT, “CERT Advisory CA-1996-21 TCP SYN Flooding and IP Spoofing Attacks”, (1996)

[2] 大下裕一 他，“観測トラヒックの統計的性質を利用した DDoS Attack の検出方法”，電子情報通信学会技術研究報告 vol.103No.651(2004)

[3] 警察庁，“SYN flood 攻撃被害観測システムについて”(2004)

[4] 萱島 信 他，“One-Packet 推定法を用いた DoS 判定手法の評価”，情報処理学会コンピュータセキュリティシンポジウム 2006 (2006)

[5] 竹森敬祐 他，“Web サーバリモート監視におけるホームページ改竄判定”，情報処理学会研究報告コンピュータセキュリティ Vol.2002 No.68 (2002)

[6] 蔀由紀 他，“インターネットにおける風評リスク対策支援サービス：BBSwatch”，雑誌 FUJITSU 2003-3 月号 (2003)

[7] 南野朋之 他，“blog の自動収集と監視”，人工知能学会論文誌 Vol.19, No.6(2004)

[8] Yahoo! ニュース，http://headlines.yahoo.co.jp/

[9] Google ニュース，http://news.google.co.jp/

[10] 官房長官発表，“Winny を介した情報漏えいについて”，
http://www.kantei.go.jp/jp/tyoukanpress/rireki/2006/03/15_a.html

[11] Telecom-ISAC Japan，“ISP との連携による ANTINNY ウイルス感染ユーザへの注意喚起の取り組み”，https://www.telecom-isac.jp/news/news20060315.html