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

2004－QAI－13　(2)
2004／10／28

トラフィック監視による新出ワームの検出システム

鈴木 和也† 馬場 俊輔† 田中 貴志† 金山 卓矢†

†横河電機株式会社セキュリティプロジェクトセンタ 〒180-8750 東京都武蔵野市中町 2-9-32
E-mail: †{Kazuya.S,Shunsuke.Baba,Takashi.Tanaka,Takuya.Kanayama}@jp.yokogawa.com

あらまし　現在のインターネットにはワームが蔓延し，恒常的なトラフィックを生成している．ワームトラフィックはDoS攻撃や間違いアクセスと同様に未利用アドレスにも到達する．したがって未利用アドレスに対するトラフィックを監視することによって，新種ワームの出現によるインシデントの早期発見が期待できる．今回，未利用のアドレスに対するトラフィックを精査することにより，定常的に蔓延しているワームとその他のスキャン行為などを分類するシステムを構築した．実験を通し提案システムにより新種ワームの出現が初期段階で確認できたことを本報告で述べる．

キーワード　ネットワークトラフィック，未使用アドレス，パケット監視，新種ワーム

New worm detecting system with traffic monitoring

Kazuya SUZUKI† Shunsuke BABA† Takashi TANAKA† and Takuya KANAYAMA†

† Security Project Dept. Business Development Div. Yokogawa Electric Corporation 2-9-32 Nakacho, Musashino-shi, Tokyo, 180-8750 Japan
E-mail: †{Kazuya.S,Shunsuke.Baba,Takashi.Tanaka,Takuya.Kanayama}@jp.yokogawa.com

Abstract Worms are spreading in the Internet these days and generating constant traffic. Worm traffic arrives at unassigned addresses just like DoS attacks and misuses. Thus, monitoring the traffic to unassigned addresses enables us to detect new worm outbreak instances. We have established a system that can distinguish well-known worm traffic and irregular traffic like scanning or new worm outbreak. This report describes that we found a new worm occurrence in its early stage with our proposed system.

Keyword Network traffic, unassigned IP address, packet monitoring, new worm

1. はじめに

近年，ワームやDoSなど様々なセキュリティインシデントが絶えず発生するようになって来ている[1]. CERT/CC発表の資料[2] "Incidents reported"によると、セキュリティインシデント数は急速に増加する傾向にあり、2000年には21,756件であったインシデントが2003年には137,529件に達している．これは、ワームやウィルスなど、自動的に攻撃を行うツールの蔓延によるものと考えられている。セキュリティはネットワーク管理者にとってはかなり重要な問題であり，一番始めに導入するツールとしてはファイアウォールや IDS(Intrusion Detection System)ツールなどの類であろう. IDS は最近注目されて来ており，様々な攻撃をリアルタイムに検知してくれる優れたツールである．

実際に行われている監視活動としては，警視庁やJPCERT/CCの活動がある．警察庁では全国に設置されたIDS及びファイアウォールにおける検知状況の結果をウェブサイト(@police[3])にて公開しており，JPCERT/CCでも定点観測システムを立ち上げセンサを分散配置し，ワームの感染活動や弱点探索のためのスキャンなど，セキュリティ上の脅威となるトラフィックの観測を行いこれも公開[4]している．さらに海外でもいくつか観測活動を行い、結果を公開しているサイトがある．なかでもDShield.org[5]は，ファイアウォールやIDSのログを収集し統計処理を行って公開している．

このように監視活動が行われるようになって来ているが，

このIDSでは基本的には既知の攻撃しか検出することができない．このため未知の攻撃には対応することができず，新種のワームなどが流行した場合には問題となってしまう．これでは最近のワームの発生速度，感染速度を考慮すると対処が間に合わなく恐れがある．実際にトレンドマイクロ社の報告[6]によると，2001年に発生したNIMDAワームはセキュリティパッチが公開されて336日後に発生しているが，2004年4月に発生したSASSERワームはセキュリティパッチが公開されてからわずか17日後に発生している．このような感染状況の中において，未知の攻撃や新種のワームを早期に発見することは極めて重要になって来ている．

2. 目的

IDS等の既存の監視ツールだけでは未知の攻撃や攻撃の予兆，さらには準備行為等を検出することはかなり困難になって来ており，近年のワームの感染速度を考えると現在導入されているシステムでは不十分である．よって，シグネチャを介さずに直接ネットワークトラフィックを監視する必要がある．現在導入されているネットワークトラフィックの監視システムはトラフィックの流量や遅延などの計測が主流であり，攻撃が大規模にならないと発見が遅れてしまう．つまり攻撃の初期段階，もしくはわずかな攻撃の特徴を検出することはほぼ不可能である．さらにIDSのアノマリ検出機能では，主に統計処理により異常検出を行うため誤検出，過剰検出が発生しやすいなどの問題がある．ネットワークトラフィックを監視する際，全てのパケットを精査していると通常の問題のないアクセスも解析しなければならず，大量のリソースを必要としログも膨大になってしまうためパフォーマンスに影響が出てしまう．したがって，ネットワークトラフィックを監視する場合は，まず精査すべきパケットと精査しなくてもよいパケットつまり攻撃パケットを判別する工程が必要になってくる．そのためには，攻撃パケットを具体的に把握する必要がある．

よって今回の目的は，ネットワークの状況を正確に把握するために攻撃トラフィックを分類することである．パケットそのものの監視を行うことで既に知られている攻撃だけでなく，未知の攻撃や新種のワームを早期に容易に発見することを目標にする．

3. システム

3.1. システム概要

本システムの概念図を図1に示す．本システムでは，インターネットの末端にセンサを設置した．これはエンドユーザと同じ環境で監視を行い，実際に攻撃パケットの振舞いを把握するためである．ただし，今回はエンドユーザのいない未利用アドレスにセンサを設置し，エンドユーザのトラフィックと攻撃トラフィックを分離する作業の省略を行った．この未利用アドレスブロックに到達するパケットは，基本的にワーム，間違いアクセス，不正アクセスのどれかであり，正常なアクセスのパケットは到達しない．したがってこのブロックに到達するパケットは，通常のアクセスとは考えにくく攻撃パケットみなすことができ，これらを精査すれば良い．この未利用アドレスブロックには，定常的に蔓延しているワームも到達するためこれと新種ワームや攻撃のための準備行為などを分離する手法が重要になる．

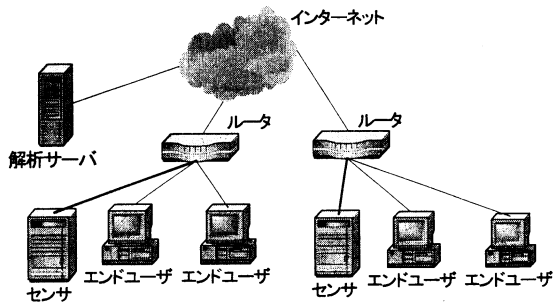

図1 システム概要

3.2. システム構成

本システムの構成を図2に示す．本システムは，末端に設置した複数のセンサとその観測データを解析するサーバから構成される．センサは監視アドレスブロックに到達するパケットを全てキャプチャし，そのキャプチャデータの保存を行う．解析サーバは，データ回収モジュール，分類モジュール，分析モジュールの各モジュールから構成される．モジュール単位に分割することによりデータの流れや処理を明確にし，システムにかかる負荷のバランスを分散することができる．回収モジュールは，センサから定期的にデータを回収し，分類モジュールへ引き渡す役割を担う．分類モジュールは，実際にデータの分類を行い，分類済みログとして保存する．分析モジュールは，分析者が解析を開始するとログを解析し，

解析結果を提供する．

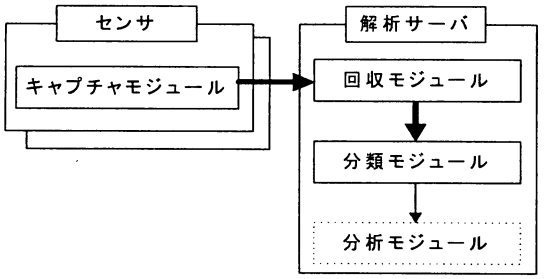

図２システム構成

4. 分類手法

未利用アドレスブロックに到達するトラフィックを精査するためには，まず適切に分類しなければならず，分類法もいくつか存在する．一般的には TCP や UDP などのプロトコル別，ディスティネーションポートやソースポート別などがある．この一般的な分類法で実際にトラフィックを分類していくと

・複数のワームが同じポートを用いる事がある

・ポートスキャンやネットワークスキャンのノイズが大きい

・恒常的に蔓延しているワームとの区別がつかない

などの問題が発生してしまう．これでは新たな攻撃や微細な準備行為を検出することが難しい．したがってこの問題を解決するために，今回は特に各パケットのソースアドレスを基準とする．このパケットの時刻を t とし，分離するための時間間隔を Δt_1，Δt_2 として，t から $t-\Delta t_1$，$t+\Delta t_2$ までの間に到達したパケットを解析範囲とする．

ここで

N：送信先ネットワークのアドレスの種類数

H：送信先ホストのアドレスの種類数

SRC：送信元ポート番号の種類数

DST：送信先ポート番号の種類数

とすると，表１のような判定を行うことにより６種類に分類することが可能となる．

表１判定条件

判定条件	N = H	N < H
SRC > DST	Port_scan	Network_scan
SRC = DST	Normal	Network_scan2
SRC < DST	Port_scan2	Network_scan3

したがって，あるホストから観測しているネットワークに到達するパケットの振舞いに着目し，

・Normal

・Port_scan

・Port_scan2

・Network_scan

・Network_scan2

・Network_scan3

の６つのタイプに分類する．

1. Normal: 送信元ポート種類数＝送信ポート種類数
   同じ送信元ポートから，数回のパケットが到達

例

xxx.xxx.xxx.xxx:3145 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:3145 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:3145 aaa.bbb.ccc.ddd:445 SYN

2. Port_scan：元種類数＞先種類数
   観測ホストの複数のポートに対して，複数のポートからのパケットが到達

例

xxx.xxx.xxx.xxx:62304 aaa.bbb.ccc.ddd:135 SYN

xxx.xxx.xxx.xxx:62769 aaa.bbb.ccc.ddd:135 SYN

xxx.xxx.xxx.xxx:63037 aaa.bbb.ccc.ddd:135 SYN

xxx.xxx.xxx.xxx:60225 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:60785 aaa.bbb.ccc.ddd:445 SYN

3. Port_scan2：元種類数＜先種類数
   観測ホストの複数のポートに対して，複数のポートからのパケットが到達

例

xxx.xxx.xxx.xxx:63644 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:63644 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:63644 aaa.bbb.ccc.ddd:445 SYN

xxx.xxx.xxx.xxx:63644 aaa.bbb.ccc.ddd:135 SYN

4. Network_scan：元種類数＞先種類数
   観測ネットワークの複数の IP アドレスの１つまたは複数のポートに対して，複数のポートからのパケットが到達

例

xxx.xxx.xxx.xxx:3594 aaa.bbb.ccc.100:445 SYN

xxx.xxx.xxx.xxx:3596 aaa.bbb.ccc.101:445 SYN

xxx.xxx.xxx.xxx:3597　aaa.bbb.ccc.102:445 SYN

xxx.xxx.xxx.xxx:3598　aaa.bbb.ccc.103:445 SYN

5. Network_scan2：元種類数＝先種類数

観測ネットワークの複数の IP アドレスの１つまたは複数のポートに対して，一つまたは複数のポートからのパケットが到達

例

xxx.xxx.xxx.xxx:4230　aaa.bbb.ccc.100:1023 SYN

xxx.xxx.xxx.xxx:1640　aaa.bbb.ccc.101:445　SYN

xxx.xxx.xxx.xxx:2117　aaa.bbb.ccc.102:9898 SYN

6. Network_scan3：元種類数＜先種類数

観測ネットワークの複数の IP アドレスの複数のポートに対して，一つまたは複数のポートからのパケットが到達

例

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.100:3127 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.100:3127 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.100:1080 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.100:1080 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.103:3127 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.103:3127 SYN

xxx.xxx.xxx.xxx:22022　aaa.bbb.ccc.103:1080 SYN

5. 観測結果

今回構築したシステムでの観測結果をそれぞれ図３から図１０に示す．各図の観測期間は，SASSER ワームが流行した 2004 年 4 月末から 6 月初旬までのものである．なお，x 軸は日付，y 軸はアクセス数を示している．

5.1. TCP/445 に関する観測結果

図３は 2004 年 4 月 29 日から 2004 年 5 月 6 日までの期間における TCP の 445 番ポートへの総アクセス数のグラフである．このグラフのデータは分類する前のデータであり，アクセス推移の振幅が大きい．これは各種スキャンを受けるとそのままアクセス数にカウントされるためである．このトラフィックを本システムで分離したものが図４である．このグラフの推移の振幅は，図３に比べて小さくなっている．これは，各種スキャンを分離したためであり，分離後のプロットデータで一番大きな値を示しているのは Normal タイプとなっている．

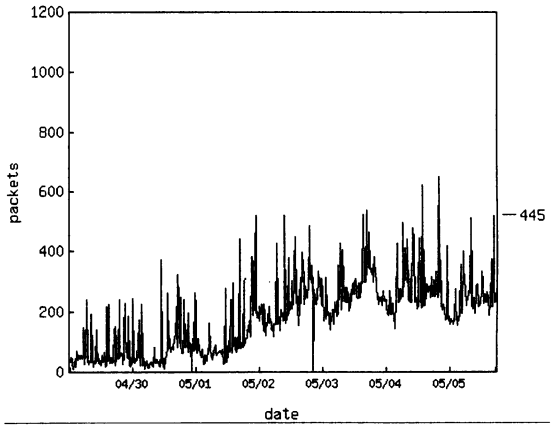

図３　TCP/445（分離前）

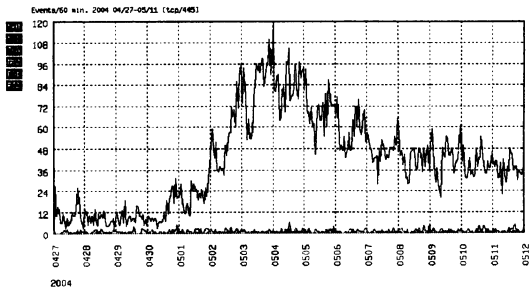

図４　TCP/445（分離後）

5.2. ICMP を考慮にいれた TCP/445 の観測結果

図５は，ICMP EchoRequest に関する観測結果である．このグラフを見ると 5 月 3 日頃に急激に増加していることがわかる．この急激な増加は攻撃の予兆とみなすことができると思われる．この ICMP EchoRequest と TCP/445 との連続したアクセスをセットとみなして解析を行うと，図６のような結果が得られる．このグラフを見ると図５とほぼ同時期に増加していることから ICMP を要求した後に TCP/445 へアクセスするタイプのものが発生したと思われる．さらにこのタイプを長期間に渡り観測した結果が図７である．このグラフを見るとこのタイプのアクセスが 6 月 4 日から 5 日にかけてピークを迎え，約 10 日間前後で沈静化していくのがわかる．

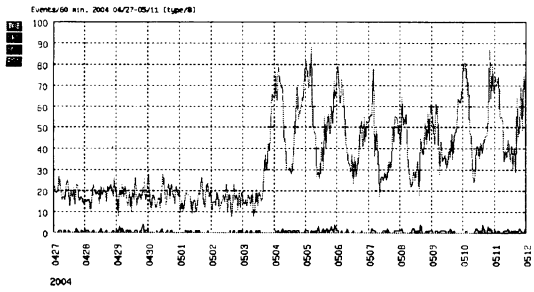

図５　ICMP EchoRequest

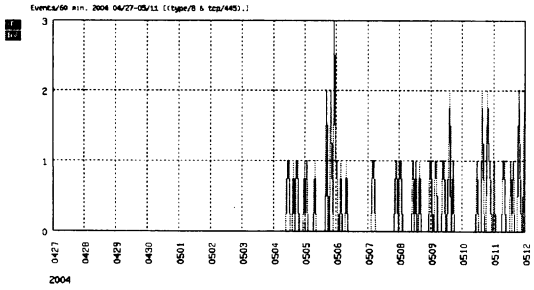

図６　ICMP EchoRequest 後 TCP/445

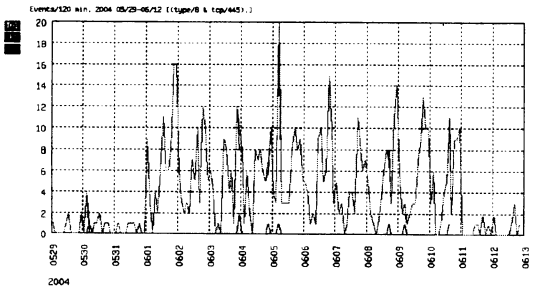

図７ ICMP EchoRequest 後 TCP/445(期間:5/29～6/12)

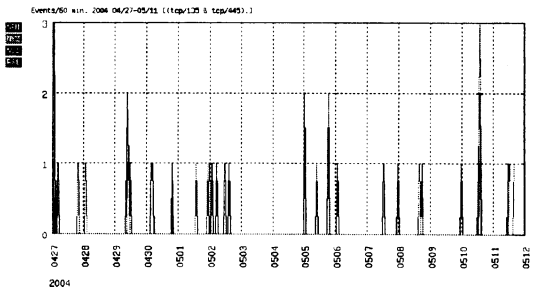

図８　TCP/135 と TCP/445

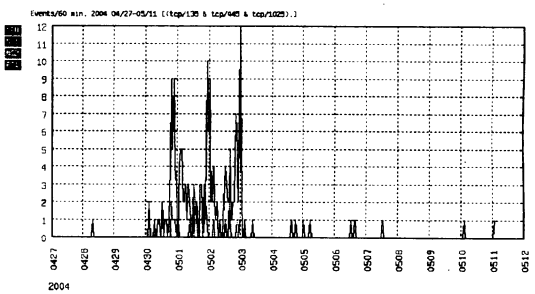

図９　TCP/445，TCP/135，TCP/1025

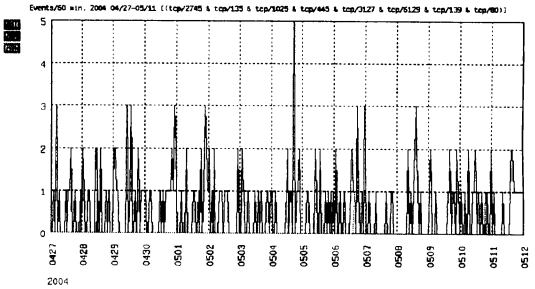

図１０　TCP の 2745，135，1025，445，3127，6129，139，80

5.3. TCP/445 とその他のポートに関する観測結果

図８は，TCP/135 と TCP/445 のセットでアクセスするものに関するグラフである．このタイプは，アクセス数としては小さいが，常に存在していることがわかる．さらに図９は，TCP の 445，135，1025 のセットでアクセスするものに関するグラフである．このタイプは，4 月 30 日頃に流行したことがわかる．さらに図１０は，TCP の 2745，135，1025，445，3127，6129，139，80 のセットでアクセスするものに関するグラフである．このタイプも常に存在していることがわかる．

6. 考察

今回の観測では，2004 年 5 月に流行した SASSER ワームに関してのデータの採取を行った．観測結果の図３では，TCP の 445 番ポートに対する総アクセス数を示したが，このグラフではスキャンなどのノイズにより変動がかなり大きく SASSER ワームと他のアクセスの判別が難しいことがわかる．そこで今回構築したシステムで分類を行った結果を図４に示してある．この図４では，ノイズが除去され SASSER ワーム

の流行が顕著に現れているのがわかる．さらに図５に示したように ICMP EchoRequest の推移を見てみると数日遅れて極端にアクセス数が増加している．この ICMP EchoRequest の増加を考慮して図６のように ICMP EchoRequest を送信した後に TCP/445 へのアクセスをグラフにしてみると確かに増加していることがわかり，これは亜種の発生を示している．さらにこのタイプのワームに関して期間を長くして観測，分離してみると６月に急激に増加している．つまりこのタイプの亜種が流行したことを示している.さらに別の亜種として図８に示したように TCP の 135 番ポートと 445 番ポートへセットでアクセスするものや，図９に示したように 445 番ポートと 135 番ポートさらに 1025 番ポートへセットでアクセスするもの発生が現れている．また図１０に示したように 2745, 135, 1025, 445, 3127, 6129, 139, 80 のセットでアクセスするものの存在も発見することが出来た．

7. あとがき

今回のシステムでは，攻撃トラフィックを６種類のタイプに分類する手法を提案し実際に観測したトラフィックに適用し，実験を行った．この分類法により恒常的に蔓延しているワームや各種スキャンを分離することができ，攻撃パケットの振舞いを具体的に把握できることを示した．

今回のシステムでは，この分類法でトラフィックを分類したが，さらにポート番号の組み合わせを分類法に加えることでより詳細に把握することが可能となる．この組み合わせの種類も自動的にリストアップする手法を検討中である．この機能は，トラフィック分析支援に利用することが可能であると思われる．

文　献

[1] 財団法人インターネット協会，“インターネット白書 2004，pp.104–05，株式会社インプレス ネットビジネスカンパニー，2004

[2] http://www.cert.org/stats/cert_stats.html#incidents

[3] http://www.cyberpolice.go.jp/

[4] http://www.jpcert.or.jp/

[5] http://www.dshield.org/

[6] http://www.trendmicro.com/jp/security/report/report/archive/2004/mvr0406.htm

[7] 武田圭司，磯崎宏，“ネットワーク侵入検知”，pp.109 - 125 ソフトバンクパブリッシング株式会社，2000

[8] William Stallings，“Cryptography and Network Security”，pp.559 – 604，株式会社ピアソン・エデュケーション，2001

[9] Carol Taylor，Jim Alves-Foss，“Session 5: less is more: NATE: Network Analysis of Anomalous Traffic Events, a low-cost approach”，Proceedings of the 2001 workshop on New security paradigms，Sep.2001

[10] R. Sekar, A. Gupta, J. Frullo, T. Shanbhag，“A. Tiwari, H. Yang, S. Zhou，Intrusion detection: Specification-based anomaly detection: a new approach for detecting network intrusions”，Proceedings of the 9th ACM conference on Computer and communications security，Nov.2002

[11] Carol Taylor，Jim Alves-Foss，“Intrusion detection and response: An empirical analysis of NATE: Network Analysis ofAnomalous Traffic Events”，Proceedings of the 2002 workshop on New security paradigms，Sep.2002

[12] Paul Barford，David Plonka，“Passive measurements: Characteristics of network traffic flow anomalies”，Proceedings of the First ACM SIGCOMM Workshop on Internet Measurement，Nov.2001

[13] Cliff Changchun Zou, Lixin Gao, Weibo Gong, Don Towsley，“Information warfare: Monitoring and early warning for internet worms”，Proceedings of the 10th ACM conference on Computer and communication，Oct.2003

[14] Stefano Zanero, Sergio M. Savaresi，“Computer security (SEC): Unsupervised learning techniques for an intrusion detection system”, Proceedings of the 2004 ACM symposium on Applied computing，Mar.2004

[15] Robin Sommer，Vern Paxson，“Intrusion detection: Enhancing byte-level network intrusion detection signatures with context，Proceedings of the 10th ACM conference on Computer and communication security”，Oct.2003

[16] Matthew V. Mahoney, PhilIP K. Chan，“Industry track papers: Learning nonstationary models of normal network traffic for detecting novel attacks”，Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining，Jul.2002

[17] Anoop Singhal, Gary Weiss, Johannes P.“Ros, A model based reasoning approach to network monitoring”，Proceedings of the workshop on on Databases: active and real-time，Nov.1996