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

2006－DPS－126（46）
2006－CSEC－32（46）
2006／3／17

複数プローブによる異常トラフィック検知システム

中村 信之，中井 敏久

概要

インターネット上にはボットと呼ばれるワームに感染した端末が多く存在しており，それらボットの構成するボットネットの潜在的な危険性が問題になっている．本稿ではそれらボットネットの危険性が顕在化する早期段階において発生する異常なトラフィックを検知するために，トラフィックの異常状態を表す”異常度”を定義してその”異常度”の変化を元に異常トラフィックを検知する手法を述べる．また複数のプローブにおいて同様の検知手法を用いその結果をあわせて解析することで，大規模な異常トラフィックの早期検知と異常と判定した原因の推定ができることを示す．

Anomaly Network Traffic Detection System using Multi-Probes

Nobuyuki Nakamura, Toshihisa Nakai

Abstract

Many nodes, which are infected by worms, are spreading all over the internet. These nodes are called bots, and these bots may construct botnets. Botnets' behavior is now a potential risk to the everyday operation of the internet. This paper proposes the way to detect the anomaly of the network traffic at the early stage of the strange behavior of the botnets by defining "anomaly degree". Anomaly degree is calculated by the statistics of the network traffic observed by a monitoring node, called probe. By combining, comparing and analyzing the results of multiple probes, we can detect a large scale network traffic anomaly events in the early stage and moreover we can estimate the source of anomaly.

1. はじめに

インターネット上には，ワームやウイルスに感染したままのコンピュータが多く存在している．さらに，ワームは感染したコンピュータの利用者に対して悪意をもっていないため利用者への影響が少なく，利用者はワームに感染していることに気づきにくい．また，ワーム対策を何もしていないコンピュータでは，感染したままの状態でインターネットに接続され続けるために，ワームの感染は広がる一方である．その上，ワームのオープンソース化によりすでにウイルス対策ソフトでは亜種の全てを検知するのが間に合わない状態であり，今後もワームはウイルス対策ソフトに検出されないように改造される可能性が高い．これらワームに感染したコンピュータはボットと呼ばれている．

これらのボットの構成するネットワークはボットネットと呼ばれ，ボットネットの機能を利用して大量のスパムメールを送信,あるいはDDoS攻撃を行うなどの悪用をされる．そのため，ボットネットがそれらの挙動を示した際，早急に検知して防ぐことが求められている．

沖電気工業株式会社 研究開発本部
ユビキタスシステムラボラトリ
Ubiquitous System Laboratory, Corporate Research and Development Center, Oki Electric Industry Co., Ltd.

2. ワーム関連技術動向

2.1. ワーム検知技術の動向

従来，ワームはウイルスと同様に扱われることが多く，ネットワークのエッジに接続されたコンピュータ上やメールサーバーなどで検出されるものであった．これらワームやウイルスの検出方法として，たとえばウイルス対策ソフトがあげられる．同ソフトに組み込まれた既知ウイルスのパターンを用いたパターンマッチング法，プログラムのビヘイビアを

監視するビヘイビア法やヒューリスティック法，パーソナルファイアウォールがある．しかしながら，すでにパターンマッチング法は新しいワームや亜種ワームに対して効果的でなく，ビヘイビア法やヒューリスティック法は誤検知に関する問題がある．一方，パーソナルファイアウォールはネットワークインタフェースの入出力パケットを監視し，管理者が認めたプログラムの通信のみを通過させるため，ワームに対して効果が期待できる．また，ワームに感染してしまった場合にもワームのインターネットに対する通信を防ぐことによってワームの拡散を抑制できる．しかし，オペレータの設定ミスや，OSから検出されずに活動できる機能を持つワームなどに対しては効果が期待できないといった問題がある．

一方，最近コンピュータやメールサーバー上ではなくネットワーク側で異常を検知する手法が多く実施されている．たとえば，snort[1]のようなネットワークIDSやISDAS[2]のような定点観測があげられる．ネットワークIDSは，通過するパケットのペイロード部分を監視して，ウイルス対策ソフトと同様にパターンマッチングによって異常なトラフィックを検出するものである．これは先に述べたウイルス対策ソフトと同様に新しいワームや亜種ワームに対して効果が期待できないという問題がある．

定点観測では一般的に能動的に通信をしないコンピュータによってある種のスキャンパケットやその跳ね返りパケットを取得し，その統計情報を用いてネットワーク状態の変化を観測している．しかしながら，ワームはランダムにアドレスを生成するため，たまたま観測点にScanパケットが流れてくることはあってもそれらScanパケットのみを用いてワーム発生を早急に知るのはかなり難しい．

2.2. ワーム生成技術の動向

現在では，ワームのオープンソース化とプロトコルの高度化により従来技術を用いてワームを検知することが難しくなってきている．

ワームのオープンソース化は，Gaobot（又はAgobot）と呼ばれるワームのソースコードが公開されたことから一般に知られるようになった．オープンソース化の結果として亜種ワームの作成が容易にできるようになったため，大量の亜種ワームが出現している．Telecom ISAC Japanの調査では1日に80種程度のワームの亜種ワームが出現していることが確認され[3]，それらの亜種ワームのうちいくつかは従来技術ではワームとして検出できなかったとされている．その理由として，従来のパターンマッチング方法では，ウイルス対策ソフトもネットワークIDSも未知ワームを検知するできないことがあげられる．

また，近年インターネットプロトコルの高度化が進んでおり，NATやファイアウォールの制限を回避して外部と通信するプロトコルが多数開発されている．これらのプロトコルを用いたアプリケーションはネットワーク管理者がその利用を把握するのが非常に困難であるところに問題があり，該当する通信のみを遮断するといった対応ができない．例としてSkype[4]のような高機能なP2PアプリケーションやSoftEther[5]のようなVPNアプリケーションのプロトコルがあげられる．このような別の正常なアプリケーションを装って通信するプロトコルの出現によって，ワームを検知することの難しさは今後ますます増大すると考えられる．

3. 提案する異常検知方式

本稿では，プローブと呼ばれるネットワークモニタリング装置を用いてネットワークトラフィックを観測し，異常トラフィックを検知する観測手法を提案する．ここで異常トラフィックとは，過去一定期間のトラフィックデータ（以下，教師データ）と判定したいトラフィックデータ（以下，評価データ）を比較し，教師データと評価データの間に大きな乖離があるものと定義する．

はじめに，3.1でどの程度の異常を示しているかを”異常度”という数値によって表現する．次に3.2では，3.1で計算した”異常度”の変化をもとにしてトラフィックに発生した異常を検知し，その原因となるトラフィック内の要素を特定する手法を示す．3.3では，ネットワーク内に複数配置したプローブにおいて同様の”異常度”と異常の原因となる要素を比較することにより広域なワーム検知ができる可能性を示す．

3.1. “異常度”の計算

ネットワークトポロジーやトラフィックの流量など，プローブの設置場所によって得られるデータが異なるため，各々のプローブを同じ指標で評価するための仕組みが必要である．ここで，”異常度”を定義する．”異常度”とは，各プローブにおける教師データに対して評価データがどの程度乖離しているかを示す値である．”異常度”の計算において利用する要素は，Etherフレームのペイロード部分の要素としてIP, ARP, RARP, OTHER(その他)．IPパケットのペイロード部分の要素としてTCP, UDP, ICMP．また，TCPのフラグとして URG, ACK, PSH, SYN,

RST, FINの合計13種類とする．これらの要素は図1で示すようにそれぞれ３つのグループに分けることができ，それぞれ全体を100%とする割合として表すことができる．この割合はトラフィックに含まれる要素のバランスであり，このバランスが崩れることを利用して異常トラフィックを検知することを考える．

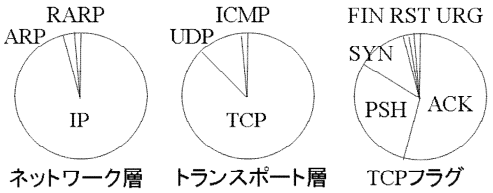

図1 トラフィックの要素

図2にトラフィックを評価するための教師データと評価データの関係を示す．評価データの取得時間から，それぞれ24時間前のデータを7つ集めたものを教師データとする．このように教師データを設定したのは，ボットの発生させるトラフィックを含めてネットワーク上を流れるトラフィックが人間の活動によって左右されるものであり[6]，24時間を7日繰り返す人間の活動を反映することが適当と考えたためである．

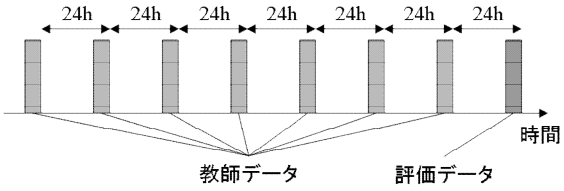

図2 教師データと評価データの関係

次に，教師データを用いて評価データを評価する前に正規化によって教師データの各要素のとりうる値の大きさの差を解消し，要素間の大きさの影響を少なくする．正規化には式（1）を用い，各要素は平均値0，分散1に正規化される．

$$y = \frac{x - \bar{x}}{s_x} \qquad \cdots \quad (1)$$

(y:正規化値, ${s_x}^2$:xの分散, x:教師データ, $\bar{x}$:xの平均)

また，教師データの正規化の際に用いた各要素の分散値と平均値と同じ値を用いて式（2）により評価データを計算する．この計算により，教師データの値を元にした評価データの各要素の値が得られる．

$$z = \frac{a - \bar{x}}{s_x} \qquad \cdots \quad (2)$$

(z:異常値, a:評価データ)

このようにして得られたzをそれぞれの要素名ごとに$z_{要素名}$と表し，その要素の異常値と呼ぶことにする．次に，トラフィックの”異常度”は式（3）に示すように各要素の異常値の絶対値の和と定義する．

$$\begin{aligned} \text{異常度} = &|Z_{ip}| + |Z_{arp}| + |Z_{rarp}| + |Z_{othe}| + |Z_{tcp}| + |Z_{udp}| \\ &+ |Z_{icmp}| + |Z_{syn}| + |Z_{rst}| + |Z_{fin}| + |Z_{ack}| \\ &+ |Z_{psh}| + |Z_{urg}| \end{aligned} \qquad \cdots \quad (3)$$

式（3）で示した”異常度”を用いると，トラフィックが普段と変わらないとき”異常度”は0に近づき，普段のトラフィックと大きく異なるときに”異常度”は大きな値をとることになる．そのため”異常度”は，ワームが既知・未知に関わらず，また特殊なプロトコルが利用されている場合にも，ワームがトラフィックを発生させた際に通常流れているトラフィックに対して有意な影響を示す場合には大きな値を取り異常の度合いを計算することができる値となる．

3.2. 単一プローブを用いた異常検知手法

異常なトラフィックが発生すると，各要素のバランスが崩れて”異常度”が大きくなることを先に述べた．しかし，全く同じトラフィックが発生することはなく常に変動する性質を持つため，”異常度”が高いからといって必ず異常があるとはいえない．こうした異常検知は一般にFalse Positiveと呼ばれ誤検知に当たる．ここでは，単一プローブにおいて”異常度”を用いて誤検知の少ない異常検知をするための手法を述べる．

今回，我々が検知したいのは，異常トラフィックである．また，ワームの発生させる異常トラフィックは継続して拡大していく性質のものであり，”異常度”は異常の程度が大きいほど大きな値を示すことが既にわかっている[7]．よって，ワームが動作して継続的にトラフィックに影響を与えた場合，継続的に大きな異常値が観測されるはずであり，その”異常度”の推移をもとに以下のⅰ，ⅱ，ⅲの検知手法に合致するものを異常トラフィックの発生と判定する．

ここで，tを時間，”異常度”を時系列に$A_{(t-9)}, A_{(t-8)}, \cdots, A_{(t)}$とし，異常と判定するための閾値を$TH$と表すこととする．

i. 一定区間内での”異常度”の傾きが正である

ex. $TH < A_{(t-2)} < A_{(t-1} < A_{(t)}$

ii. 一定区間内で異常の占める割合が大きい

ex. $A_{(t-9)}$～$A_{(t)}$のうち過半数がTHより大きい

iii. 一定区間内で”異常度”が増加傾向である

$$\text{ex.}\quad \begin{aligned} TH &< (A_{(t-8)} + A_{(t-7)} + A_{(t-6)})/3 \\ &< (A_{(t-5)} + A_{(t-4)} + A_{(t-3)})/3 \\ &< (A_{(t-2)} + A_{(t-1)} + A_{(t)})/3 \end{aligned}$$

以上の3つの検知手法により異常トラフィックの判定をする．これらの検知手法にはそれぞれ以下のような特徴がある． i は急激な”異常度”の立ち上がりの検知に向いており，早急な対応をするための検知手法である． ii は異常の継続性を判定しており，急激に広がりつつある異常ではないが，異常な状態が定常的に続いていることを検知する手法である． iii は，急激ではないが，ゆるやかに異常な状態に推移しつつあることを検知する手法である．

また，これらの検知手法を用いて異常を検知した際に，各要素で異常値が高くなっているものが異常の原因と推測できる．例えば図3のように”異常度”が増加し異常と判断した際に各要素の異常値の推移は図4のようになったとする．このとき，異常と判断した原因は図4において大きな異常値を示している要素であるといえ該当する要素はICMPである．よって異常の原因はICMPの急激な増加であり，該当するパケットを足がかりに詳細に解析することで原因の特定が可能であると考えられる．

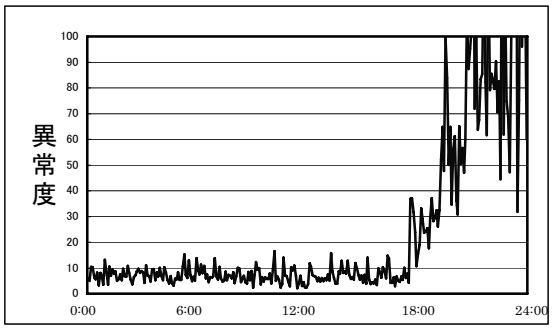

図3 “異常度”の推移

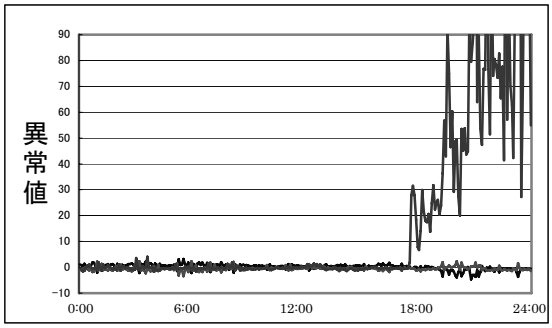

図4 図3に対応する異常値の推移(要素名は省略)

3.3. 複数プローブを用いた異常検知手法

ここまで，単一プローブでの”異常度”の計算及び異常検知手法とその原因要素の特定手法を示してきた．次の目的は，広域にワームが発生して異常トラフィックを発生させた際に，同種の異常が各々のプローブにおいて検知できることである．最も単純なシーケンスを図5に示す．このシーケンスでは，前半で2つ以上の複数のプローブで同時に異常が検知されたかどうかを判断している．次に，複数プローブで異常が検知されている場合それぞれの異常の原因が同種の要素によるものかどうかを判断し，同種の原因によるものであれば同種のワームによる異常である可能性が高いと判断している．この複数のプローブを用いた異常検知手法はプローブをまとめて管理するセンター装置によって行うことを想定している．また，複数のプローブで同じ要素に起因する異常が発生した場合，他プローブに対して同要素の異常値の変化に注意するように喚起することで同種の動きが確認された際に，より早急に検知できると考えられる．ここまでに述べた異常検知手法を用いて，次章で実際のトラフィックを評価する．

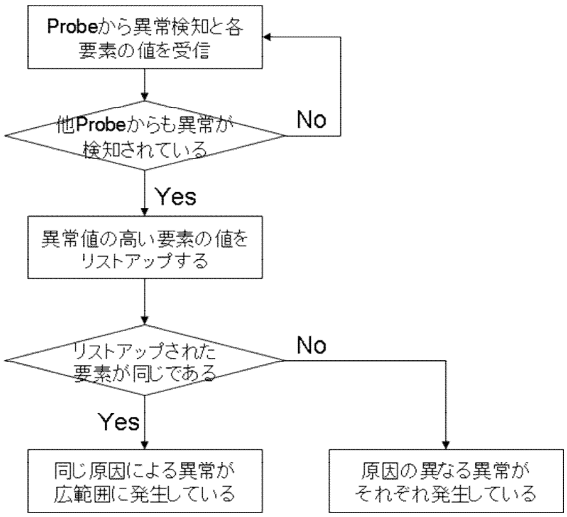

図5 複数プローブによる異常検知シーケンス

4. 評価

今回の評価にあたり，以下に示す箇所に設置された2つのプローブで観測されるトラフィックを用いた．また，これらのプローブでは5分間隔にトラフィックをサンプリングしている．

- イントラネットの無線LAN（以下，WL点）
- イントラネットの有線LAN（以下，WR点）

まず，先に用いた図3はWR点の2005年12月のある日の”異常度”の推移を示しており，図4は同日の各要素の異常値を示している．図3によると，18時を過ぎた

あたりから”異常度”が急激に上昇しているのがわかり，この原因は図4によりICMPと判断できた．これは図4において同時刻に高い異常値を示しているのがICMPであるためである．図示していないがこの時刻に対応する実際のトラフィックはICMPの値が平常時の3倍ほどになっていた．よってこの異常は普段は「安定」しており変動の少ないICMPと比べて異常な変化だと判断されたと考えられる．また，閾値として適当な値[7]を用いると，3.2に示した3つの異常検知手法の初検知時間はそれぞれ，19時20分，18時55分，19時45分であった．その後ICMPの値が平常時に戻るまでの間，3.2で述べたⅱの検知手法では断続的に異常と判断し続けたが，検知手法ⅰでは稀に，検知手法ⅲでは頻繁に異常と判断される結果となった．以上のことから，この時間帯に起きた異常はICMPの増加によるものであり，ゆるやかに異常状態へと遷移し，そのまま異常状態が継続したと考えられる．また，同時刻のWL点では上記検知手法にも目視にも異常が検知されなかったため，WR点のみの単独の異常であったと判断できる．

次に，図3とは別の2005年12月のある日のデータを評価する．図6にWR点での”異常度”の推移を，図7に図6に対応する各要素の異常値の推移を示す．

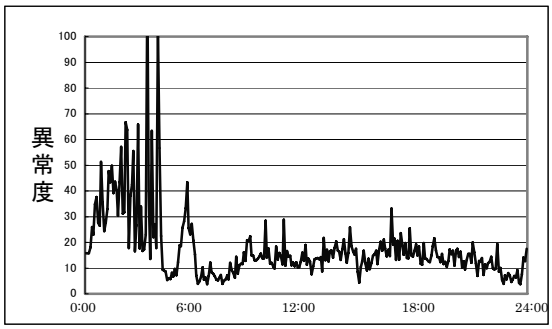

図6 WR点での”異常度”の推移

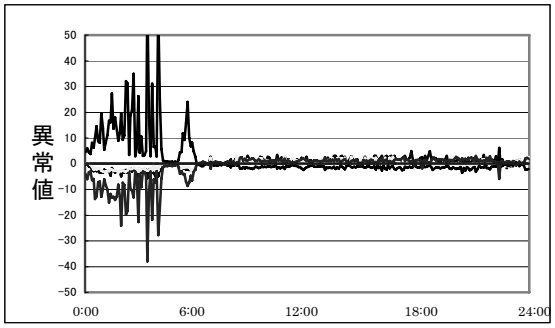

図7 WR点での異常値の推移(要素名は省略)

まず，図6であるが，0時から4時頃まで全体的に”異常度”が高い状態が続いている．この原因としては，図7より主にTCP-ACKとTCP-PSHのバランスの崩れにあると考えられる．図の該当する時間にプラス方向に大きく振れているのがTCP-ACKであり，マイナス方向に大きく振れているのがTCP-PSHである．このことからTCPを用いたなんらかの特殊なトラフィックが発生した可能性が高い．また，ここでは図示していないが，実トラフィックにおいて深夜時間帯にはおあまり大きなトラフィックがないのが普通であるにも関わらず，この日は日中と同程度のトラフィックが観測されていた．異常値としてTCPの値が目立っていないのは，教師データと比べて異常と呼べるほど大きい値をとっていないためであった．

また3.2に示した3つの異常検知手法では，検知手法ⅰで検知せず，検知手法ⅱで0時55分から4時過ぎまで，検知手法ⅲでは1時半から3時にかけて数回検知するという結果になった．このことから，この時間帯にTCPを用いた異常トラフィックが発生したのは確かである．しかし，ICMPやTCP-RSTに大きな異常が見られないためTCP ACK Scanではないと思われる．よって，この異常はワームのように感染拡大しないものであり，決まった数のホストがTCPを用いた特殊なトラフィックを一定時間発生させた後に収束したと考えられる．より詳細な解析はトラフィックのダンプデータを用いて行う必要がある．

さらに，図6や図7と同日のWL点での”異常度”の推移を図8に，図8に対応する各要素の異常値の推移を図9に示す．

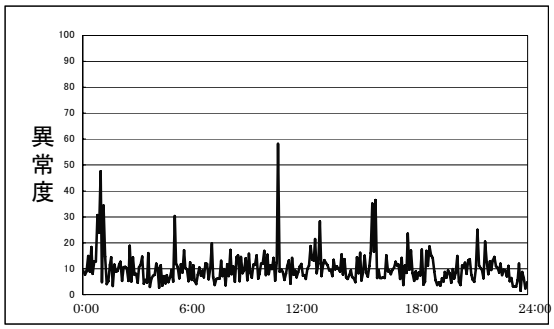

図8 WL点での”異常度”の推移

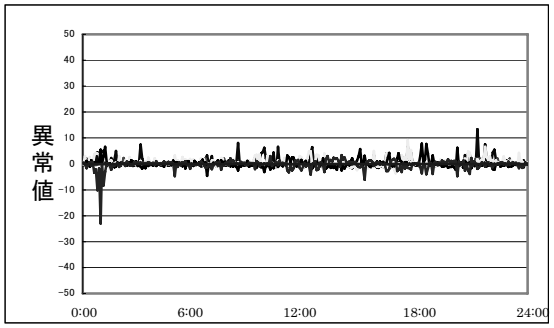

図9 WL点での異常値の推移(要素名は省略)

この図を評価する前に，まずWL点の特徴を述べる．WR点が常時なんらかのトラフィックが流れているのに比べて，WL点では勤務時間帯を外れるとほとんどトラフィックが流れなくなるという特徴がある．この理由は，無線LANを用いたトラフィックの大半がモバイルPCによるものであり，利用者の活動と密接に関連しているためである．このことから，WL点では基本的にユーザーの操作によるトラフィックしか流れないために統計処理をするのに十分なデータが用意できない状況となり，”異常度”のみを用いるとFalse Positive及びFalse Negativeの両方の誤検知が発生しやすいという特徴を持つことがわかる．図8においても，いくつかの”異常度”の高い箇所が目視で確認できるが， 3.2の検知手法により誤検知を回避している．

さて，図8のデータの評価であるが，3.2に示した3つの検知手法を同時刻，同じ閾値で実行したところ，どの検知手法でも異常を検知できなかった．理由として先に図3と図4を用いた評価結果のようにWR点でのみ観測される異常である可能性もあるが，図8を見ると同時間帯のうちの一部箇所で”異常度”が大きくなっている箇所がある．また，図9により同時刻にTCP-PSHがマイナス方向に大きく振れていることがわかる．よって，WR点と同様の異常がWL点で観測されていた可能性も考えられる．この異常をWR点で異常を検知した際にWL点のプローブに対してTCP-ACKのプラス方向への振れとTCP-PSHのマイナス方向への振れを検知した際に，異常と検知するように働きかけることで，WL点においても同様の異常が検知できると考えられる．

今回の評価期間内にワームやその他の特殊な異常が発生することがなかったため，実際にワームの発生させるトラフィックを検知できるかどうかを評価することはできなかった．したがって3.3の異常検知方式の評価が不十分ではあるが，複数のプローブに対して同種の異常を検知できる可能性は示すことができた．また，複数のプローブを連携させることでより精度が高く早急な異常検知ができる可能性を示した．

5. まとめ

本稿では，”異常度”を用いてネットワーク内に設置したプローブ装置で異常を検知する手法，及び複数のプローブ間でのデータ比較による広域異常に対応する手法を示した．また，複数のプローブ装置を用いてトラフィック異常を検知し，原因となるトラフィック要素を特定することができることを示した．

実トラフィックにおいては過去と全く同じトラフィックが来ることは無く，定義した”異常度”が0になることは評価データ内では無いことが確認できた．また，図3，図6，図8よりわかるように，トラフィックボリュームやトポロジー，利用されているアプリケーションが異なっても平常時の”異常度”はほぼ同程度の小さい値となる．このことから，複数プローブを同様の手法で比較・評価できることがわかった．

しかしながら，無線LANに見られたようにトラフィックが少なく統計処理をするのに十分なデータが用意できないときには各要素の割合が大きく変化することがあり，誤検知がおこりやすいプローブも存在する．この点に関して本稿では3.2の検知手法により改善されたことが確認できたが，動的にサンプリング間隔を変える，あるいはパケットカウントベースのサンプリング手法を用いることで統計処理に十分なデータを用意することも検討課題である．

また，本方式はより深い解析を始めるための前処理と位置づけているため，用いる要素をかなり絞っている．そのため１つ目の評価に見られたようなICMPの異常に関する原因推定において，ICMPのtype/codeも異常検知に含めることにすれば検知段階でより多くの情報が得られる．この他にも利用可能な要素を追加することと，その妥当性を検討することも今後の課題である．

6. 謝辞

本研究は，独立行政法人情報通信研究機構（NICT）の委託研究「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」の一環として行われた．ここに深謝する．

参考文献

[1] snort (http://www.snort.org/)

[2] ISDAS (http://www.jpcert.or.jp/isdas/)

[3] Telecom ISAC Japan，” 第１回 JPCERT/CC 共催セミナー”，2005 年 7 月．

[4] Skype (http://www.skype.com/)

[5] SoftEther (http://www.softether.com/)

[6] David Dagon, Cliff Zou, Wenke Lee, “Modeling Botnet Propagation Using Time Zones”, Network and Distributed System Security Symposium, February 2006.

[7] 中村信之，中井敏久，”トラフィック内部状態変化を利用したネットワーク異常検知”，(社)電子情報通信学会信学技報 NS2005-5，2005 年 4 月．