# トラフィック分析による異常検出システムの試作

2000MT086 杉野 裕一　　2000MT093 田代 亮
2000MT097 角岡 新一　　2000MT098 内田 幸治
指導教員　後藤 邦夫

## 1　はじめに

近年，ネットワークを流れるパケットを監視して不正アクセスと思われるパケットを検出する侵入検出システム (Intrusion Detection System 以下 IDS) の重要性が増している．

商用の IDS の多くは，監視するパケットを既知の攻撃の特徴を記したデータ (シグネチャ) と照らし合わせることによって不正なパケットの検出を行っている．この方法は頻繁にシグネチャの更新が必要であること，未知の攻撃には対応できないこと，処理が重たく高速のトラフィックが発生するネットワークへの対応が難しいといった問題点がある．

一方，統計的手法を採用した IDS は，“通常の通信”によるトラフィックと現在のトラフィックとの相違から警告を発する．商用 IDS でこの手法を用いているものもあるが，分析対象はシグネチャベースの手法で検出された一部のプロトコル違反のパケットに限られている．また，snort で一部実装されているが，検出目的は portscan に限定されている．

そこで本研究では，ネットワークトラフィックを監視し，パケットヘッダだけを収集し，特に異常なパケットだけでなく，正常な通信も含めた全ての通信に対し，応用プロトコル別に分類計数し，パケットの数を統計的に分析することで異常を検出するシステムを試作する．

過去 2 年間の研究 [1][2] ではパケットキャプチャと分析を同一ホスト上で行っていたが，本研究ではこの二つを異なるホストで行うことにより効率面においてシステムを改善した．

本研究は，大学院生の高木陽司さんとの共同研究であり，研究の全体的な管理は高木さんが担当した．異常検出方法については杉野，田代，角岡が担当し，プログラムは内田が担当した．

## 2　異常検出方法

一般的に異常検出とは異常の原因を直接検出するのではなく，システムやユーザのふるまいを監視し，通常とは異なるふるまいを検出することである．異常を検出することで，管理者の調査を促し，結果的に不正や障害を発見することを狙いとしている．

### 2.1　MSD モデル

本研究では数値で表されるデータのうち，ある一定の範囲の値を持つものに対して，事前に上限値および下限値を設定しておき，観測データがその範囲を越えた場合に異常とするものを MSD(Mean and Standard Deviation) モデルとする．MSD モデルでは，観測結果からしきい値を自動的に設定する．本研究ではしきい値を平均と標準偏差から求める．標準偏差を用いることで，過去の観測値の散らばり具合を考慮したしきい値を設定することができる．したがって，異なる 2 つの時系列データから同じ平均が得られても，データの変動の幅の大小によって異なるしきい値が設定され，観測対象に応じた評価を行うことができる．この検出手法では，短期の異常の検出が期待できる．

本研究では過去の観測結果のデータ $(x_1, x_2, x_3, \cdots, x_n)$ から平均 $(\overline{x})$ と標準偏差 $(S)$ を用いてしきい値を算出する．それぞれのデータが互いに独立で，正規分布に従うとして、データ $x_{n+1}$ のしきい値を

$$\overline{x} \pm (d \times S) \tag{1}$$

と設定する．ここで d は標準正規分布の信頼度区間幅で，例えば 95%の信頼区間を設定する場合は $d = 1.96$ である．本研究では，$x_{n+1}$ の値が式 (1) の範囲を越えた場合を異常とする．このモデルでは，短期異常の検出が期待できる．

$n$ と $d$ をパラメータとして与えることで，しきい値の間隔を調節することができる．

### 2.2　Anomaly Score モデル

本研究では [5] の論文で述べられている anomaly score という変数を導入する．本研究では [5] で述べられている anomaly score を導入し，その値の急激な変化をに注目することで異常を検出する．以下では，項目を $dstPort, dstIP$ とした場合を例として、anomaly score を用いた異常検出方法を説明する．

$x$ を，ある dstPort,dstIP に来たパケット数とする．一般的なトラフィックで，そのポートをターゲットにしている確率 $P(x)$ は，

$$P(x) = \frac{\text{ある } dstPort, dstIP \text{ に来たパケット数}}{\text{総パケット数}} \tag{2}$$

となる．これを全ての $dstIP$・$dstPort$ の組合せごとに計算する．anomaly score $A(x)$ はこの確率の対数をとって，

$$A(x) = -\log_2(P(x)) \tag{3}$$

とする．

確率が小さいほど anomaly score$A(x)$ は大きくなるので，高い anomaly score が算出された時は，普段あ

まり観測されないパケットが観測されたということになる．これを利用して [5] では，anomaly score を用いることで主にポートスキャンの検出をしている．

本研究では，anomaly score の急激な変化を異常とする．あるパケット数間隔で観測データを区切ったとき，過去のデータから得られた anomaly score を $A(x_{old})$，最新のデータから得られた anomaly score を $A(x_{new})$ とする．$A(x_{old})$ と $A(x_{new})$ を比べることでパケット数の変化を見る．

具体的には，以下の方法で異常を定義する．
$A(x_{new}) - A(x_{old}) > threshold(decrease)$ ：この種の通信の比率が過去に比べて減ったと報告
$A(x_{old}) - A(x_{new}) > threshold(increase)$ ：この種の通信の比率が過去に比べて増えたと報告

2 つのしきい値 $threshold(decrease), threshold(increase)$ はパラメータとして与える．anomary score モデルでは一定のしきい値を用いる．

### 2.3 LOF(local outlier factor) モデル

本研究では [6] の論文で述べられている LOF を異常検出手法の一つとして導入する．Local Outlier Facter(LOF) とは，各データのはずれ値の程度を表す値である．はずれ値の程度は，そのデータの回りの局所的な密度を考慮して計算される．データの回りの密度が，そのデータの近くにある他のデータの回りの密度と比べて相対的に大きい場合 (はずれ値の程度が高い場合) は，LOF の値が大きくなる．そうでない場合は，LOF の値は 1 に近付く．

データ $P$ とデータ $O$ の距離を $d(P,O)$ とした時 (図 2 参照)，LOF は以下のように求められる．

1. 各データ $P$ に対して，$k\text{-}distance(P)$(データ $P$ と，データ $P$ から $k$ 番目に近いデータとの距離) と $k\text{-}distance\ neighborhood$(距離が $k\text{-}distance(P)$ 以下のデータの集合) を計算する．
2. 各データ $P$ に対して，$reachabilitydensity(each - dist_k(P,O))$ を計算する．ただし，$O$ は $P$ を含まない全てのデータである．
3. 各データ $P$ に対して，$local\ reachability\ density(lrd_{MinPts}(P))$ を計算する．$lrd_{MinPts}(P)$ は $P$ から $k\text{-}distance$ の範囲にあるデータ $O$ の $reach-dist_{MinPts}(P,O)$ の平均の逆数である．$MinPts$ とは近隣データの数を定義するパラメータである．
   ここで，$N_{MinPts}(P)$ はデータ $P$ からの距離が $k\text{-}distance(O)$ 以下となるデータの数である．
4. データ $P$ の LOF を計算する．LOF はデータ $P$ の $lrd_{MinPts}P)$ と $P$ から $k\text{-}distance$ の範囲にあるデータ $O$ の $lrd_{MinPts}(O)$ の比率の平均として定義される．

本研究では，$n$ 個の時系列データ $(x_1, x_2, x_3, \cdots, x_n)$ と評価対象のデータ $x_{n+1}$ の集合から，$x_{n+1}$ の LOF を求める．これによってデータ $x_{n+1}$ が過去のデータと比べてどれくらいはずれているのかが分かる．この程度が大きい場合を異常とする．

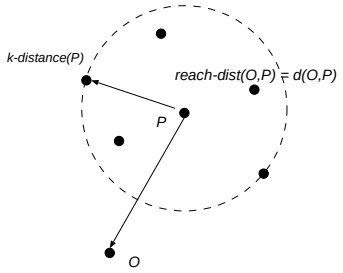


図 1: k-distance

### 2.4 Holt-Winters Forecast モデル

Holt-Winters Forecast は，周期性を持つ時系列データの過去の観測値を基に，未来の値を予測する方法である．この手法では観測された時系列データが 3 つの構成要素 (baseline, trend, seasonal effect) に分解できることを仮定している．$t$ を現在の時刻，$m$ を 1 周期のデータ数として，予測値 $\hat{y}_{t+1}$ は次のように求められる．

$$\hat{y}_{t+1} = a_t + b_t + c_{t+1-m} \tag{4}$$

ここで，$a_t, b_t, c_t$ は $\alpha, \beta, \gamma$ をパラメータとして，以下のように定義される．

- Baseline
$$a_t = \alpha(y_t - c_{t-m}) + (1-\alpha)(a_{t-1} + b_{t-1}) \quad (0 < \alpha < 1) \tag{5}$$
- Linear Trend
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \quad (0 < \beta < 1) \tag{6}$$
- Seasonal Trend
$$c_t = \gamma(y_t - a_t) + (1-\gamma)c_{t-m} \quad (0 < \gamma < 1) \tag{7}$$

この方法を異常検出に応用することで，周期的な変動を考慮したしきい値を算出することができる．[7] と同様，本研究では予測値を基に実測値のしきい値を設定し，その範囲を越えた場合を異常としている．

## 3 処理の流れ

検出処理は以下の流れで行う (図 2).

## 4 実験

2 節で述べた検出手法のうち，MSD モデル，LOF モデル，Holt-Winters Forecast モデル，Anomaly Score モデルによる検出結果を比較し，さらに商用 IDS，snort の検出結果とも比較することにより試作 IDS の性能評価を行う．

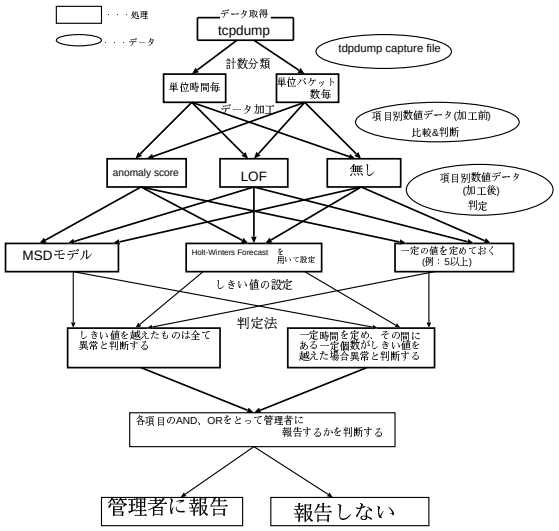


図 2: 処理の流れ

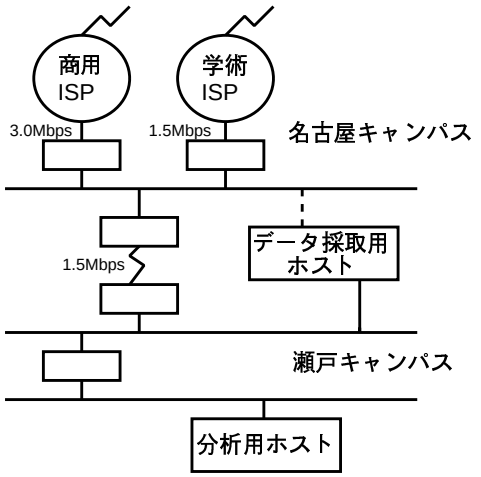


図 3: システム構成図

### 4.1 データ

2 週間分のデータを tcpdump を利用して採取した．以下にその詳しいデータを示す．

南山大学 2003 年 12 月 5 日 0 時 0 分 ～ 2003 年 12 月 18 日 23 時 59 分

以下に，このデータを用いた実験について，その方法と結果を示す．

### 4.2 システム構成

本研究で試作するシステムはパケットを採取するホストと，そのホストから受け取ったデータを分析するホストの 2 種類から構成される (図 3 参照)．データ採取用ホストでは tcpdump を動かしておき，分析用ホストからデータ転送の要求が来ると，HTTPS を用いて今までに採取したデータを圧縮して転送する．分析用ホストは，ある一定時間毎に (現在は 10 分毎) データ転送要求を出し，データを保存し, 解析を行う．

### 4.3 実験方法

最初に，得られたパケットヘッダの一つ一つを 10 進数の数値データに変換し, 以下の項目を抽出する．

$Time, TCPflag, Protocol, ICMPtype, SourceIP, ICMPtype$
$DestinationIP, ICMPsubtype, SourcePort, payload$
$DestinationPort$

また，パケットヘッダ中に該当する項目が存在しない場合は-1 を代入している．次に，パケットヘッダを項目別に分類し，その数を数える．分類計数のプログラムは，オブジェクト指向言語の JAVA で作成した．オブジェクト指向言語を用いた理由は，パケットの分類計数に用いる木構造の実装が簡単だからである．

分類項目は，次のようにした．

$direction, Protocol, ICMPtype, ICMPsubtype$
$TCPflag, DestinationPort, SourcePort$

“direction” は通信の向きを表し，パケットヘッダを

- 学内から学外
- 学外から学内
- 学内から学内

の 3 種類に分類したものである．

採取したデータを上記の方法で分類計数し，採取したパケットの時系列データを作成する．作成された時系列データを基に MSD モデル，LOF モデル，Holt-WintersForecast モデル，anomaly score モデルで異常検出を行った．

### 4.4 結果

一定時間毎に異常を検出する MSD モデル，LOF モデル，Holt-Winters Forecast モデルでは，項目別にパケット数の急激な変化をほぼ捉えることができた．一定パケット数毎に異常を検出する Anomaly Score モデルでは，パケット数の多い項目がパケット数の少ない項目に影響し，項目別にパケット数の変化を捉えられた部分が少なかった．また全てのモデルにおいて，しきい値を大きくすることで検出数が減少することが確認できた．この結果からしきい値の幅を変化させるパラメータは，アラームレベルに応用できると考えられる．

### 4.5 商用 IDS，snort との比較

本研究では Cisco IDS と snort を用いて，試作 IDS で異常と見なした部分とアラームとを比較し評価を行う．

**利用した異常検出システムのパラメータ**

評価に使用したパラメータは以下である．

MSD モデル: $n = 12$ $D = 4$

LOF モデル: $n = 24$ $MP = 12$ $threshold = 8$

Holt-Winters Forecast モデル: $\alpha = 0.195$ $\beta$ =0.00208 $\delta = 3.5$

Anomaly Score モデル: 200000 パケット毎 $\alpha = 0.5$ $threshold = 5$

**利用した CiscoIDS の項目**

本研究で対象とした項目と比較した Cisco IDS のアラームを以下の図 4 で表す.

| | |
|---|---|
| A : OUTgoing TCP SYN+ACK | 1 : IP fragments overlap |
| B : OUTgoing TCP RST | 2 : Sendmail Data Header Overflow |
| C : OUTgoing TCP RST+ACK | 3 : IP Fragment Attack |
| D : INcoming TCP SYN | 4 : FTP real path Buffer Overflow |
| E : OUTgoing UDP | 5 : Nachi Worm ICMP Request |
| F : INcoming UDP | 6 : TCP High Port Sweep |
| G : inside UDP | 7 : TCP Port Sweep |
| H : ALL UDP | 8 : TCP SYN Host Sweep |
| I : ICMP type0 | 9 : Half-Open SYN |
| J : ICMP type3 sub3 | 10 : UDP Port Sweep |
| K : ICMP type4 | 11 : Nmap UDP Port Sweep |
| L : ICMP type5 | 12 : Net Flood ICMP Echo Reply |
| M : ICMP type8 | 13 : NET Sweep Echo |
| N : ICMP type11 | 14 : Net Flood ICMP Echo Request |
| O : traffic outgoing | 15 : Snort SYN alert |
| P : traffic incoming | 16 : Snort SYN/ACK alert |
| Q : traffic inside | 17 : Snort ACK alert |
| R : traffic total | |

図 4: 実験に用いた検出項目とアラーム

**比較方法**

CiscoIDS のアラームの中から致命的な被害になりそうな攻撃についてのアラームを 4 個，2003 年に流行した Nachi ワームに関するアラーム，TCP，UDP，ICMP それぞれ通信に関係するであろうアラームを用いて評価を行う.

**比較結果**

- MSD
  項目 15(Snort SYN alart), 16(Snort SYN/ACK) において，非常に高い確率で異常を検出できた.
- LOF
  項目 5(Nachi Worm ICMP Request) において 91%重なっていた．また項目 15(Snort SYN alart), 16(Snort SYN/ACK) において，非常に高い確率で異常を検出できた.
- Holt-Winters Forecast
  項目 15(Snort SYN alart), 16(Snort SYN/ACK) において，非常に高い確率で異常を検出できた.
- Anomaly Score
  すべての項目について，うまく検出できていなかった.

## 5 おわりに

実験の結果から，MSD モデル・LOF モデル・Holt-Winters Forecast モデル・Anomaly Score モデルにおいてパケットの急激な変化を捉えていることが確認できた．すべてのモデルで異常と思われる部分を検出することができた箇所もあったが，得られる結果に大きな違いがある部分もあった．この違いは，モデルによって検出に優れているアラームが異なることが考えられる.

snort での比較は検出率が全体的に高いという結果がえられたが，これも偶然に snort と試作 IDS でのアラームが一致した可能性が考えられる．実際はうまく検出できていないのだが試作 IDS では異常と判断されている誤報 (False Positive) が snort の圧倒的なアラーム数に埋もれてしまっている可能性も考えられる.

また，snort のアラームと試作 IDS の各モデルでの外向き，内向き，内部，合計のトラフィック量とも比較を行った．SYN と SYN/ACK のアラームではあまり違いは見られなかったが，理由はわからないが ACK アラームに関しては合計のトラフィック量をどのモデルと比較した場合にも他の項目に比べて高い検出率を示していた．さらにパラメータを変化させてアラームを厳選することで，これらの関係がより明確になると思われる.

今後の課題としては，より検出率が高くなるようアラーム数の調整やパラメータを変化させた実験を繰り返すことが必要である.

また検出率を算出する方法も検討しシステムを評価を行うことが必要である.

## 参考文献


[1] 福山勉，上田絵美，“ネットワークトラフィック分析による侵入と異常検出システムの試作” 南山大学経営学部情報管理学科卒業論文, (2002)

[2] 荻野大介，大田渡，山中勇樹，“トラフィック分析に基づく異常検出システムの試作” 南山大学経営学部情報管理学科卒業論文，(2003)

[3] 武田 圭史, 磯崎 宏, **ネットワーク侵入検知**, ソフトバンクパブリッシング株式会社, 2000.

[4] www.silicondefense.com

[5] Stual Staniford, Janes A.Hoagland, and Joseph M. McAlerney: “Practical Automated Detection of Stelthy Portscans” , www.snort.org.

[6] Markus M. Breunig, Hanse-Peter Kriegel, Raymond T. Ng, Jorg Sander: “LOF:Identyfying Density-Based Local Outliers”, ACM SIGCOMM 2000 Int. Conf. On Management of Data, Dalles, TX, 2000

[7] Jake D. Brutlag, :“Aberrant Behavior Detection in Time Series for Network Monitoring”, www.usenix.org/ publications/library/proceedings/lisa2000/index.html., 2000.

[8] Paul Barford, Jeffery Kline, David Plonka, Amos Ron.: “A Signal Analysis of Network Traffic Anomalies”, Proceedings of ACM SIGCOMM Internet Measurement Workshop , 2002.

[9] B.B ハバード (山田道夫 西野躁 訳), **ウェーブレット入門**, 朝倉書店, 2003.