日本オペレーションズ・リサーチ学会
2004年秋季研究発表会

ベイジアンネットワークを用いた不正アクセスの統計的検知

岡村寛之 (01013754)†, 福田達徳 ‡, 土肥正 (01307065)†
† 広島大学大学院工学研究科情報工学専攻
‡ 広島大学工学部第二類（電気系）

1. はじめに

不正アクセスからコンピュータを守るために，侵入検知システム (IDS: Intrusion Detection System) の開発が行われている．IDS の主な不正アクセス検出方法は，ネットワークのトラフィックかサーバへのアクセスに対する正常状態のパターンを学習・認識し，正常状態から外れたパターンを不正アクセスとして検知する方式（異常検知方式）と，不正アクセスのパターンを登録し，そのパターンに合致するものを不正アクセスとする方式（不正アクセス検知方式）が用いられる．異常検知方式は未知のパターンに対応できるが，設定が難しく誤検出が多い．一方，不正アクセス検知方式は誤検出が比較的少ないが，未知のパターンに対応できないことが問題になっている．

文献 [1] では，異常検知方式に着目し，不正アクセス検知法を検証している．具体的には IP 監視ツールを用いてポートにおける通信履歴（ポートプロファイル）を記録し，ポートプロファイルと現在の処理パターンを比較することで不正アクセスの検知を行っている．

本稿では，文献 [1] の不正アクセス検知法に対して，次の二つの観点から改良を行う．一つ目はポートプロファイルを作成するためのデータに関する改良であり，送受信したパケット数やバイト数に関する相関と宛先ポート番号を考慮することで，精度の高い不正アクセス検知を目指す．二つ目の改良では複数因子への対応を考える．一般にサーバアクセスの確率的挙動は時間帯や曜日に依存して変化することが報告されている [2]．つまり，ポートプロファイル以外に時間帯や曜日といった因子を考慮する必要がある．また，不正アクセスされることによって引き起こされるプロファイルの統計的な異常は，不正アクセスの種類や方法によって単一の因子に現れる場合，或いは複数の因子に現れる場合など多岐にわたると考えられる．そのような観点から，不正アクセスの検知を複数の因子から分析する枠組みは必要不可欠であると考えられる．特に本稿では複数の因子からなる分析を統計的に行える枠組みとしてベイジアンネットワーク (BN: Bayesian Network) を用いた不正アクセス検知法を提案する．

2. ポートプロファイルを用いた統計的分析

ポートプロファイルとは，ネットワークのポート（他のパソコンと接続するための出入口）における通信履歴を記録した度数分布表である．ポートプロファイルに基づいた不正アクセス検知は，現時点のアクセス状態のプロファイル (SP: Short-term Profile) を正常稼動時のプロファイル (LP: Long-term Profile) と比較することで行われる．ポートプロファイルは，通信履歴から対象とするデータを取り出して，対応するクラスへ出現頻度を記録することで作成される．ポートプロファイルの更新は以下の式を用いて行われる．

$$\begin{aligned} L_i &= r_l L_{i-1} + X_i, \\ S_i &= r_s S_{i-1} + X_i. \end{aligned} \tag{1}$$

X_i は新規のアクセスデータであり，r_l と r_s は LP と SP の減衰率である ($0<r_s<r_l<1$)．即ち，r_l と r_s によって古いデータの価値が割引かれる．文献 [1] では，サーバから入ってくるパケット数・バイト数とサーバから出て行くパケット数・バイト数の記録を個別に用いていた．本稿では，それらに相関を持たせた記録を用いた不正アクセスの検知（方法 A），さらに宛先ポート番号の増減を用いた不正アクセスの検知（方法 B）を考える．

上記の手順で作成した LP と SP に基づいて不正アクセス検知を行う．統計的分析による不正アクセス検知は，通常稼動プロファイル (LP) と現時点におけるプロファイル (SP) を比較することで行われる．LP と SP はともにポートの通信記録に対する度数分布表であるため，度数分布表から生成される経験分布に対する χ^2 値によって LP と SP の乖離度を推測することができる．特に本稿では χ^2 値に対する p 値 (χ^2 値が棄却される最小の有意水準) を算出することで不正アクセスされている確率を表現する．方法 A，B によって得られる p 値を q_A, q_B とする．次章では，q_A，q_B を利用して BN による不正アクセス検知を提案する．

3. ベイジアンネットワークによる確率推論

BN とは，様々な事象に対する因果関係を有向グラフを用いて表現した確率モデルであり，観測可能な事象から観測不可能な事象の発生確率を推定するための枠組みである．

次の三つの事象を考える．方法 A で異常が発生する（事象 A），方法 B で異常が発生する（事象 B），不正アクセスを受けている（事象 M）．観測可能な事象 A, B から事象 M の生起確率を推定することを考える．

三つの事象に対して，図 1 で表現される BN を考える．これは，不正アクセスを受けた結果として，パケット数・バイト数あるいは宛先ポート番号に異常が発生することを表現している．図中の表は各事象の条件付き確率を表す．例えば，確率 a は事象 M が発生したもとで事象 A が発生する確率を表す．いま，事象 A，B の発生する確率が q_A と q_B によって与えられるため，ベイズの定理から以下の式によって不正アクセスを受けている確率を算出する．

$$\begin{aligned} p_M = {} & \frac{abp'_M q_A q_B}{abp'_M + a'b'p'_M} + \frac{\bar{a}bp'_M \bar{q_A} q_B}{\bar{a}bp'_M + \bar{a'}b'p'_M} \\ & + \frac{a\bar{b}p'_M q_A \bar{q_B}}{a\bar{b}p'_M + a'\bar{b'}p'_M} + \frac{\bar{a}\bar{b}p'_M \bar{q_A}\bar{q_B}}{\bar{a}\bar{b}p'_M + \bar{a'}\bar{b'}p'_M}. \end{aligned} \tag{2}$$

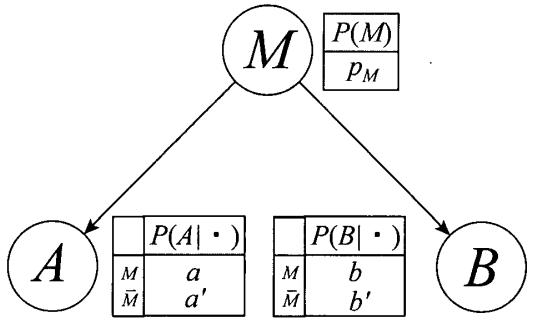

図 1: BN による不正アクセス検知確率モデル.

一般に $\bar{p} = 1 - p$ である．ここで，p'_M は事象 M に対する事前確率であり，$p'_M = p_M$ によって更新される．

4. 評価テスト

ここでは実際に観測したデータを用いて，BN による不正アクセス検知の有効性を検証する．ここでは DoS アタックをシミュレーションしたデータを用いて不正アクセスの検出を試みる．観測は 1ヶ月間行われ，総アクセスデータ数は 77598 個である．59083 個目の観測記録から 63494 個目の間が DoS アタックによるアクセスデータである．方法 A，B におけるポートプロファイルは三つのクラスを持ち，クラスの幅はパケット数およびバイト数に関して 10 個および 6000 バイトとした．さらに減衰率を方法 A では $r_l = 0.9999$ と $r_s = 0.99987$，方法 B では $r_l = 0.999$ と $r_s = 0.9985$ とした．方法 A，B で得られた q_A，q_B と式 (2) から不正アクセスされている確率 p_M を算出した．図 2 は横軸にアクセスデータ，縦軸に式 (2) から算出される不正アクセスされている確率を示したグラフである．図 2 では，特に DoS アタックを受けている 59083 個目から 63494 個目の部分で不正アクセスされている確率が高くなっており，DoS アタックを検出することができていることがわかる．

また比較として，不正アクセスされている確率が次の式で与えられる場合を考える．

$$p_M = 1 - \bar{q_A}\bar{q_B}. \tag{3}$$

これは，方法 A あるいは方法 B のどちらか一方に異常が発生する (OR 条件) という確率を表している．図 3 は式 (3) によって導出される不正アクセスされている確率を示したものである．図 2 と比較すると，図 3 は実際に不正アクセスを行っていない部分で確率が高くなる傾向が見られ，誤検出が多いことがわかる．また，2 では DoS アタックのアクセスデータ部分で，不正アクセスされている確率が 1 になってるが，図 3 の場合は 0.8 程度である．このことから BN によって複数の因子を考慮することが不正アクセス検知に有効であることがわかる．

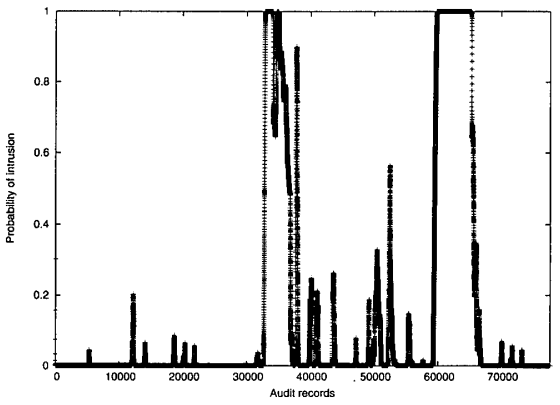

図 2: 不正アクセスされている確率のふるまい (BN).

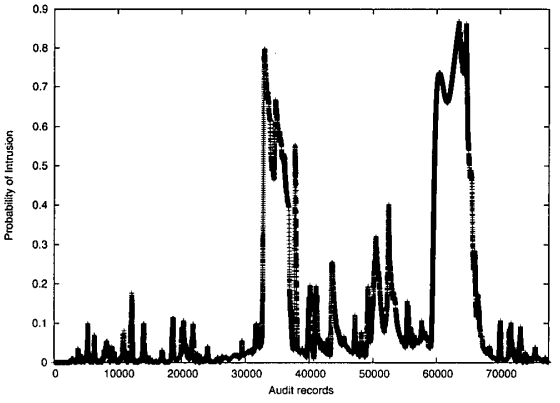

図 3: 不正アクセスされている確率のふるまい (OR 条件).

5. まとめ

本稿では，BN による複数因子を考慮した不正アクセス検知に関する手法を提案した．数値例では，パケット数やバイト数に関する記録を用いた分析と宛先ポート番号に関する記録を用いて，BN による不正アクセス検知を，複数因子の単なる和事象とした場合の結果と比較し，BN によって複数因子を考慮する確率推論の枠組みが最も有効であることがわかった．

参考文献

[1] M. Iguchi and S. Goto, Detecting Malicious Activities through Port Profiling, IEICE Transactions on Information and Systems, Vol. E82-D, No.4, pp.784-792, 1999.

[2] 藤本 衡, 清水 貴史, ネットワーク不正アクセスの確率的挙動について, 待ち行列シンポジウム「確率モデルとその応用」報文集, pp.249-254, 2003.