IDSのログ視覚化システムの構築

水谷 正慶† 白畑 真† 土本 康生†† 村井 純†

概要

近年，ネットワーク上の攻撃を検知する手段として Intrusion Detection System（以下 IDS）が多く用いられるようになってきた．IDS には誤検出やログが膨大になるという特性があるため，ネットワークの管理者等が速やかに状況を把握できるような形式へのログの加工が重要となる．しかし，これまでのログ視覚化の手法では攻撃数に着目しており，危険性と攻撃数が比例しない攻撃を管理者が見逃してしまう可能性が高い．また攻撃毎に状況を比較するにも不向きな手法であった．それらを踏まえて本論文では色による警告発生の傾向を表現するシステムを提案する．本システムでは IDS による警告が無い時間帯はなにも表示せず，ある時間帯は警告数によって色を決定し，それを時系列にそって一覧する．さらにそれらを並べることで，行われた攻撃の種類，時間，頻度などを容易に比較できる．本研究ではこのように攻撃を効果的に視覚化するシステムを構築し，実ネットワーク上で評価を行った．

Development of Visualization System of IDS's Log

Masayoshi Mizutani† Shin Shirahata† Yasuo Tsuchimoto†† Jun Murai†

Abstract

In recent years, the Intrusion Detection System (IDS) has been became more popular the technique for detecting attacks on computer networks. The IDS often detects incorrect incidents and generates huge amount of logs. Therefore, converting the IDS's logs to the format which administrators can understand the network situation easily is very important. However conventional techniques for visualization focus on the number of attacks. Also, they are not adequate to compare each attacks. Based on the current situation, I am going to suggest a system performing the trend of alerts by color in this paper. This system displays nothing during the term it detects nothing, and when it detects attacks, it displays color which is based on the numbers of attacks. Also those colors are lined up by the time order. It is easy to compare types, time and frequency of attacks by listing them. In this research, I have developed the attack visualization system and evaluated it on an actual network.

1 はじめに

今日，インターネットの利用範囲の拡大・利用者の増加に伴ってセキュリティレベルの向上が求められている．このため，ネットワークを監視し不審もしくは攻撃であると判断されるトラフィック (以下インシデント) を特定しログに残す Network Based Intrusion Detection System(以下 IDS) が広く利用される．

本論文ではログの閲覧，及び既存の視覚化手法の問題点を指摘し，新しい視覚化手法を設計，実装をした．このログ分析ツール「Mimir(ミーミル)」について述べる．

† 慶應義塾大学環境情報学部
Keio University, Faculty of Environmental Information
†† 慶應義塾大学政策メディア研究科
Keio University,
Graduate School of Media and Governance

2 背景

2.1 IDS のログ閲覧の必要性

IDS のインシデント検出の必要性は被害発見前と被害発見後に分けることができる．発見前には定期的もしくは即時的なインシデント監視，過去との比較による状況調査を行う必要がある．また，被害発見後には被害状況把握やインシデント発生日時特定，侵入経路特定，侵入されたホストからの攻撃の有無といった情報を得ることが必要である．

2.2 IDSのログ閲覧における弊害

IDSはインシデントを検出しログに残すシステムであるため，どのような脅威が発生したかが，ログを見れば一目でわかることが，IDSの要件である．しかし，現実にはそのログをそのまま利用するのは難しい．その理由として，ログが膨大な量になること，誤検出が発生することの二点が挙げられる．ネットワーク全体のトラフィックを監視するIDSは，単数のホストのトラフィックから検出されるログに比べ，必然的にログの量が多くなってしまう．テキスト形式で保存されたログは，管理者が閲覧する際にそのままでは閲覧に多大な時間が必要となり，また全体的にどのような傾向があるかを把握することが困難になる．

また，IDSが多様なトラフィックを発生するネットワークを監視する場合，正常なトラフィックであるにも係わらず，異常なトラフィックと判断してしまうことが多い．そのためログのほとんどを誤検出が占めるという場合も少なくないので，ログ全体の傾向を把握し，それが誤検出か否かを判断しなくてはならない．

これらの理由により，IDSのログを閲覧する際は視覚化，もしくは集約化が必要となる．

3 既存の視覚化表現における問題点

IDSのログ視覚化についての研究は以前から行われており[1][2]，既存のデータ視覚化表現であるグラフによって視覚化が行われている．しかし，もともとグラフは量的な割合や変化を見ることを主たる目的としている表現手法であるため，重要なインシデントの情報を得るためには不十分な視覚化になっている．以下にその理由を挙げる．

3.1 円グラフ，帯グラフ

円グラフによって視覚化する場合，インシデントの種類毎の発生件数をグラフの要素にする事が考えられる．しかし，円グラフはそれぞれの発生件数と全体の発生件数を比較する視覚化であり，根本的にインシデントの種類によってインシデントの合計発生件数がもつ意味は異なってくるためあまり意味をなさない．さらには発生件数の多いインシデントが目立つばかりか，発生件数の少ないインシデントは存在すら見えなくなってしまうため，IDSの視覚化としては適さない．

3.2 折れ線グラフ，棒グラフ

折れ線グラフでは時系列に沿って発生件数の推移を表示する視覚化ができ，[1] これはインシデントの発生状態を的確に表しているように見える．しかし，折れ線グラフでは数の多い項目が強調され，数の少ない項目を目立たなくする．そのため，全体の状況を把握するために全てのインシデントを同時に表示しようとして全体で同じ最大値を用いると，発生件数の多いインシデントばかりが強調される事となってしまう．すると発生件数が少ないインシデントは発生件数の多いインシデントに埋もれてしまい，管理者が認識できないという問題がある．逆にインシデント毎に最大値を変えると非常に見づらくなってしまい，比較が難しくなる．

3.3 分布図

分布図は存在の有無を示す事を目的としたグラフなので，例えばx座標を時間経過，y座標をインシデントの種類とすることでインシデントの発生を的確に表現できる．しかし，発生件数の変化について表すことができないため，いつ・どのくらいの数のインシデントが発生したかを判断ができない．また，分布図は二つの連続性を持ったデータをもとにグラフの表示位置を決定するため，それぞれが全く異なる性質を持つインシデントについての表現にはあまり適さない．

4 ログ視覚化における要求

既存のログ視覚化手法での問題点を解決する新しい視覚化手法を提案するために，IDS のログ視覚化における要件を以下にまとめる．

4.1 時系列にそった表現

ログ情報を集約し特定の期間中に発生したインシデントの数だけを表そうとすると，そのインシデントの発生にどのような傾向があるのかを掴めない．インシデントが集中的に発生したのか分散的に発生したのか，分散的であればそれは定期的か不定期的であるか，あるいはどの時間帯にどのような発生件数の差があるか，発生開始と発生終了がいつか，といった時系列に沿って見ることで得られる情報は多い．

4.2 分類別の比較

IDS では不正侵入に限らず，その攻撃準備とされるホストの調査や DoS 攻撃まで様々な種類のインシデントを検出できる．複数のインシデントは互いに関連性を持つ場合があり，各インシデントの発生頻度や時間を比較し，関連性を発見できる．また，ある一つのソース IP アドレスからのインシデントを一つにまとめて視覚化するだけではなく，インシデントの種類毎に分類して比較することによって，そのソース IP アドレスの行動パターンを分析できる．

4.3 量的表現と質的表現の両立

IDS が検出するインシデントは量と危険性は相関関係を持たない場合が多い．例えばポートスキャン，フィンガープリンティングの行為は DoS 攻撃となる量でなければ，何度行われても実質的な被害にはならない．しかし，バッファオーバーフローによる攻撃やバックドアからの侵入は，ただ一度の通信により，重大な被害を及ぼすことになる．このような，数そのものは少ないが危険性の高いインシデントは，数量の多い他のインシデントや誤検出に隠れてしまいがちである．そのためインシデントが「多い，少ない」という量的なものではなく，「有無」という質的な視点が必要となる．ただし，インシデントの量的推移が重要な情報となることもある．そのため，「量」と「質」の両方を得られる視覚化でなくてはいけない．

5 視覚化方法の設計

4 章で述べた要求を踏まえ，色を用いてインシデント数の推移・変化である「量」とインシデントの発生そのものの「質」を同時に表現し，かつ項目毎の比較を実現する方法を提案する．

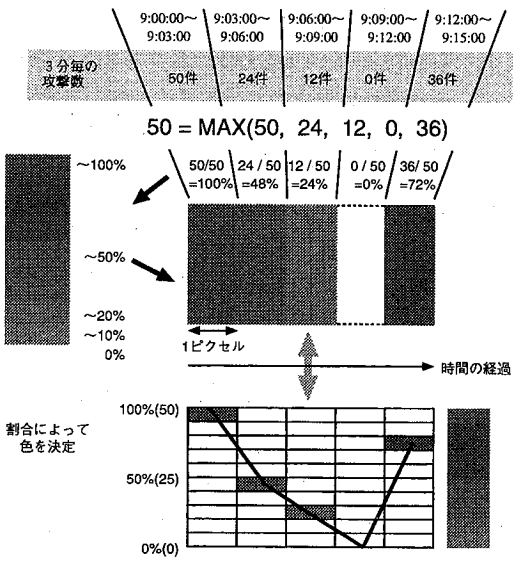
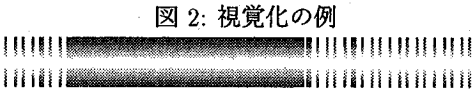


図 1: 色による視覚化の方法

まず，数行あるいは数十行からなる表にして，一行で一つの項目についての情報を表す．項目の分類方法は，主にインシデントの種類・ソース IP アドレス・ディスティネーション IP アドレスが主に挙げられる．これによって，項目毎の情報を並べて表示できるので，それぞれを容易に比較できる．

図 2: 視覚化の例



次に，図1に各行における視覚化方法について述べる．まず，一つの行において横軸を時間軸とし，左から右へ進むことで時間が経過が表される．そして，時間軸を表現可能な最も小さい単位に分割する．(コンピュータのディスプレイ上に表示するならば1ピクセルとなる) これを便宜的に「コマ」と呼称する．

その後，表示する時間の範囲を指定する．ここでは，それぞれのコマに時間軸にそった時間の範囲を割り当てる．(例 9:00〜9:03，9:03〜9:06，…)

割り当てたそれぞれの時間中に検知されたインシデントの数をIDSのログから算出する．そして表示する時間範囲の中で最もインシデント数の多かったコマの数値をとりだし，それを通常の折れ線グラフの最大値にあたるMAXとする．それぞれのコマにおけるインシデント数を，このMAXを分母としMAXに対する割合を計算する．コマには，その値が低い程「標準色」に近い色となり，高い程「強調色」に近い色を着色する．さらに，インシデントが無い時間は無色にする．2に色を用いて視覚化されたインシデントの例を示す．これによりインシデント数の差で検知したインシデントが隠れることなく，明示的にインシデントがあった時間帯となかった時間帯を認識できるようになり，同時に量的な変化を目で捉えることが可能となる．また，表示する時間の範囲を変化させることによって，細かなインシデントの発生間隔を捉えたり，長期的な傾向を見ることができる．さらに各行を比較することでインシデント発生状況の把握が容易となり，項目の分類方法を変えることによって様々な側面からの把握も可能とする．

6 実装

6.1 実装方法

5章の機能を備えた解析ツールとして「Mimir」(ミーミル) を実装した．図3に実装方法を示す．実装は，Debian GNU/Linux 2.2上でおこない，Webサーバに Apache 1.3.27を用い，RDBMSに MySQL 3.23.54を利用した．IDSはオープンソースのSnort1.9.0[3]を利用した．本システムはC++によるCGIとなっている．

スイッチがポートミラーによってコピーしたトラフィックをSnortが監視し，検出したログはSnortによってRDBMSに出力される．そして，Webサーバが受け取ったリクエスト情報をもとにRDBMSから取得したデータを加工し，Webブラウザを通して管理者に視覚化した情報を表示する．

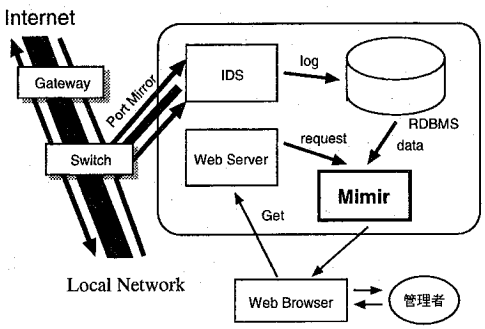


図 3: 実装方法

6.2 Mimir

システムの動作画面を図4に示す．本システムでは指定した期間に発生したインシデントについてのログを分析して表示する．図4は分析期間を24時間に設定したものを示している．左部には分類した項目の一覧を表示する．これは分類する内容によって異なり，図4ではインシデントの種類の一覧となっている．また，視覚化の部分で行う量的推移は分類された項目毎の最大値を用いるため，中央部で示される項目毎のインシデントの数を見て他の項目との比較を行う．最後に右部では5章で説明した視覚化を行う．これにより項目毎の比較も実現している．他にインシデント，IPアドレスを種類毎に分類して表示する機能を備えている．

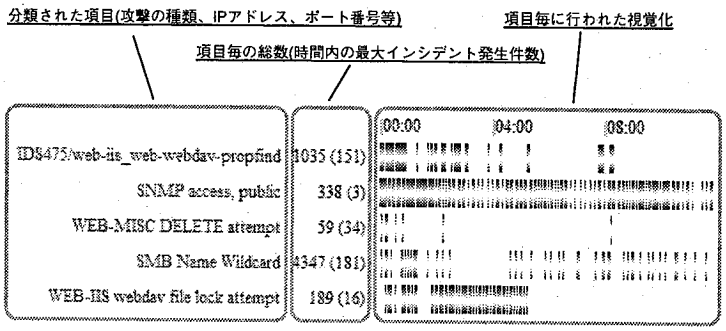


図 4: 「Mimir」画面

7 評価

7.1 質的なインシデント視覚化の評価

図 5 で，同一のデータを本システムと折れ線グラフで表現したものを比較する．これはある一つのホストに対するインシデント情報であり，特定のポート番号に特定のデータグラム，例えば”/bin/sh” の文字列が送られてくるといった，あるインシデントについて表示したものである．折れ線グラフ上では左部に大量のインシデントが発生していることを示す突出が見られる．一方，本システムでは，当該インシデントは強調色として表示されている．

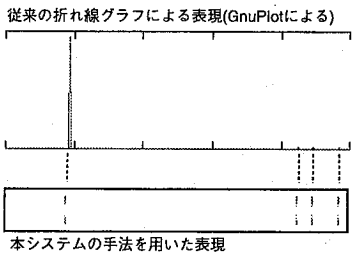


図 5: 一項目についての視覚化効果の比較例

しかし，右部では本システムがインシデント発生を示しているにも係わらず，折れ線グラフ上では僅かにしか表示されておらず，相当な注意が無いと見落としてしまう．もし項目毎に比較を行うなら，折れ線グラフの上下間隔はさらに狭まり，インシデントの存在は視認できなくなってしまう．しかし，このインシデントで重要なのは左部の突出した部分よりも，右部の少ない発生である．一つのホストが全く同じバッファオーバーフロー攻撃を短時間に受けても攻撃者にとってのメリットは少ない．よってこれは偶然このホストが特定のポート番号を使用してデータ転送を行い，その際に特定のデータグラムを多く含むデータであったことが伺える．しかし，右部の場合はごく少ない発生件数だがバッファオーバーフローによる攻撃はただ一度の通信で成功するため，一つのホストに対する発生件数はごく少ないものと考えられる．そのため，こちらの方が意図的な攻撃である可能性が高いと考えられる．

これが実際にバッファオーバーフローによる攻撃であったか否かは，トラフィックのデータグラムやヘッダー情報を見ることで明らかにできる．しかし折れ線グラフや棒グラフではその存在を認識できない場合があり，本システムではそれを解決している．

7.2 項目毎の比較と誤検出判断

図 6 ではインシデント，IP アドレスを種類毎に分類した表示について表している．上部ではあるインシデントの種類についての視覚化を行っているが，これだけを見ると不規則に多数のインシデントが発生しているように見える．しかし，下部に示すようにフロー毎に分解し，それを並べて表示することで，実は規則的な大量のインシデントの発生を確認できる．規則的

に行われる攻撃が無いとは言い切れないが，この例では発生間隔も非常に短い訳ではなく，さらにフローも内部から内部へのものなので，何らかの設定ミスかあるいは正常なトラフィックが攻撃と誤認されているのではないか，といった推測をたてる事ができる．このように視点を変え項目毎の比較を行うことで，誤検出であるトラフィックの判断・推測を行うことが可能となり，ひいては全体の状態把握を容易にしている．

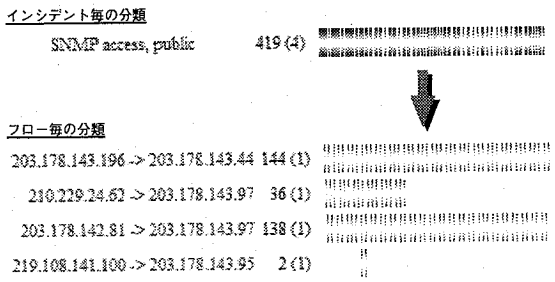


図 6: 項目の分類による比較の例

8 問題点，今後の課題

8.1 危険性示唆の必要性

本システムはIDSのログを一定の法則に従って視覚化する機能のみを持つ．そのため，危険性の高さは管理者本人の知識と視覚化された表示をもとに判断しなくてはならない．このような運用方法では管理者に豊富な知識が要求されるが，実際に全ての管理者が豊富な知識を備えているとは言えない．そのため，危険性の高いインシデントやその状況を判断し管理者に示唆できるような視覚化が必要となる．それにより，さらに多くのネットワークにおいて安全性の高い運用が行われる事が期待できる．

8.2 IDSの本質的な問題

本システムはあくまで出力されたログから管理者が状況を把握するのを助けるためのものである．そのため，IDSが抱えている問題の本質である誤検出の多さを解決する事にはなっていない．誤検出にはそのネットワーク特有のトラフィックが攻撃と似ている，という状況も多々存在する．今後は管理者が判断した内容をそのまま状況判断のみに使うのではなく，それを用い動的にIDSの設定を変更・反映し動作させるようなシステムが求められる．

さらに，IDSによる情報だけではなく，管理ネットワーク内のマシン毎のOS，利用アプリケーション，アクセスログの情報を把握することによって，脅威となるインシデントか否かを判断する事が可能となる．

9 むすび

本論文では，IDSのログ視覚化をするにあたっての要求を定義し，それを解決するシステムの実装を行った．このシステムによって，インシデントの発生件数の差に左右されずインシデントを視認，比較できるようになった．そのため，インシデント発生状況の傾向分析の手間を大幅に減らし，誤検出の特定や危険性の高いインシデントの容易な発見を実現した．

参考文献

[1] *Analysis Console for Intrusion Databases*, Roman Danyliw, http://www.acidcool.com

[2] *Snort Report*, Circuits Maximus, 2000. http://www.circuitsmaximus.com

[3] *Snort*, Martin Roesch, "SNORT-LIGHTWEIGHT INTRUSION DETECTION FOR NETWORKS", USENIX LISA '99 Conference, 1999. http://www.snort.org

[4] 高田 哲司，小池 英樹，"ログ情報の視覚化による不正侵入検知手法の提案"，コンピュータセキュリティシンポジウム'98，情報処理学会，pp.153-158，Oct. 1998.