ネットワーク不正侵入監視のための視覚化の一手法

伊藤 貴之　高倉 弘喜　沢田 篤史　小山田 耕二

京都大学学術情報メディアセンター
京都大学高等教育研究開発推進センター

計算機ネットワークのセキュリティ維持管理の目的で、近年 IDS（侵入検知システム）技術の研究開発が進んでいる。しかし現状の IDS 製品には、ログ出力情報の膨大さ、複雑化するネットワーク攻撃の分析技術の不足、などの課題が残っている。本報告はこの課題を改善する一手段として、情報視覚化の技術を適用してネットワーク侵入検知データの概要理解と詳細探索を支援する技術を提案する。本報告では計算機ネットワークに分布する計算機群を IP アドレスで階層化して、「平安京ビュー」という情報視覚化技術を用いて画面表示する。この画面表示上に侵入検知データの統計量をマッピングすることで、現実性のあるネットワーク侵入検知の振る舞いをいくつか観察できた。

A Visualization Technique for Monitoring of Network Intrusion Detection Data

Takayuki ITOH　Hiroki TAKAKURA
Atsushi SAWADA　Koji KOYAMADA
Academic Center for Computing and Media Studies, Kyoto University
Center for the Promotion of Excellence in Higher Education, Kyoto University
itot@computer.org, {takakura, sawada, koyamada}@media.kyoto-u.ac.jp

IDS (Intrusion Detection System) is an active research topic for the purpose of cost reduction of security maintenance of computer network. However, existing IDS technologies still have some issues, including enormous log output data, and lack of analysis technologies of complicated behavior of recent intrusions. This report proposes a technique to support understanding and exploration of such behavior of intrusions, by applying an information visualization technique. The technique constructs hierarchical data according to IP addresses of computers in a target network, and visualizes the data by Heiankyo-view, which is a new technique for hierarchical data visualization. By mapping statistics of intrusions onto the visualization display, we could observe some behavior of real intrusions.

1. はじめに

インターネットの不正アクセスやウィルス等による被害の拡大に伴い、近年では IDS (Intrusion Detection System) の研究が活発であり、その商用化も進んでいる。これら IDS 製品の多くは、監視対象のネットワークに一定の安全対策が施されていることを前提にして、不正現象（インシデント）を漏れなく検出している。しかしながら、特に大規模で開放性の高いネットワークにおいて IDS 製品を実際に運用してみると、以下のような問題が生じることがわかった[Saw03]。

- インシデント 1 件ごとに 1 件の警報を送信するメール通報システムでは、警報メールの量が膨大になるだけでなく、その相関性や統計的傾向を理解することが困難である。
- 最近の不正アクセスは複雑化する傾向にあり、複数のシグニチャ（インシデントの種類

やパターン）によって一連の不正アクセスを構成するケースが多く、その全貌を把握するには機械的処理だけでは不十分である。
- インシデントの大量さゆえに、インシデントを事後分析するためのデータベースの運用も容易ではない。
- GUI による従来の探索的な閲覧システムには、重要なインシデントを検索することが難しい、多忙な管理者では手に負えない、多数の管理者間の知識共有に向かない、遠隔からの業務に向かない、などの問題がある。

これらの問題点を解決し、安全性の高いネットワーク運用を実現するための手段はいくつか考えられる。本報告では、CG 技術によって情報の特徴を提示する「情報視覚化」という技術を適用した手法を提案する。本手法は、大規模階層型データを対象とした「平安京ビュー」[Ito03]という情報視覚化手法を用いて、数千、数万もの計算機を有する大規模ネットワークに広がるインシデントに対して、その統計的傾向を一画面に全貌表示する。この手法により、大量なインシデントの傾向を容易に把握できるだけでなく、従来の GUI 技術の問題をも改善できると考えられる。

また本報告では、実在する大規模ネットワークに実際に生じた大量のインシデントの視覚化例を提示し、どのような現象を視覚的に理解できたか、について解説する。

2. 関連研究

不正アクセスの検出、警告、分析に関する技術は、すでにいくつか商品化されているが、いずれの製品にしても 1 章で列挙した問題点を解決してはいない。この問題に対して、いくつかの改善手法が提案されている。例えば IDS が検出するインシデントを蓄積する統合管理型のデータベースにより、インシデントの事後分析を支援する研究がある[Ohy02]。また、テキストマイニングや周期性解析などの分析的手法を用いて、精度の高いインシデント分析を目指す研究がある[Miy02]。

情報視覚化によってインシデントの統計的傾向を表現する手法もいくつか発表されている。高田らは「見えログ」という手法を提案している[Tak02]。この手法では、画面を水平方向に分割し、左側にインシデント全体の時間別件数を一列に棒グラフ表示し、その右側にユーザーが指定したある時間におけるインシデントの内訳を表示する。この手法はインシデントの時間的な全体的傾向と局所的傾向を同時に表現できる点に特徴がある。また IDS とはデータ構造が異なるが、Axelsson は 1 台のウェブサーバーのアクセスログファイルを対象として、横軸に属性の種類、縦軸に属性値、を配置した Parallel Coordinate という座標系に個々のアクセスを配置する手法を提案している[Axe03]。この手法は不正アクセスの属性値の相関性を表現することで、不正アクセスの詳細な性格の理解を支援している。

以上の従来手法と異なって本報告は、数千、数万規模の多数の計算機が接続された大規模ネットワークを対象として、その個々の計算機にて送受信されるインシデントの計算機ネットワーク空間上での分布を視覚化することを目的としている。本報告の提案手法は、大量な不正アクセスの中に潜む人為的な要因、例えば
- 一台の計算機から大量の計算機を攻撃する。
- 大量の計算機から一台の計算機を攻撃する。
- ある特定の組織に属する計算機に組織外からウィルスを配置すると、その計算機が攻撃者に転じて、一斉に他者を攻撃する。
- ある特定の組織に属する多数の計算機を、組織外から一斉に攻撃する。

というように、計算機の空間的分布に関係ある人為的要因の視覚化に向いていると考えられる。

3. 平安京ビュー

本報告で用いる視覚化手法「平安京ビュー」[Ito03]は、大規模な階層型データの全貌を一画面に表現することを目的とした視覚化手法である。平安京ビューでは、階層型データを構成する葉ノードをアイコンで表現し、枝ノードを入れ子状の長方形の枠で表現している。

図 1 に、平安京ビューによる階層型データの視覚化例を示す。平安京ビューは、数千、数万もの葉ノードを有する階層型データに対し、画面上で互いに重なり合うことなく、できるだけ小さい画面空間に余すことなく、またすべての葉ノードを対等な大きさで表現することができる、という点において他の階層型データ視覚化手法に対する優位性をもつ。

本報告では、ネットワークに接続されている計算機の IP アドレスに注目して計算機群をグループ化する。まず IP アドレスの上位 1 バイト目が同一である計算機をグループ化し、続いて上位 2 バイト目が同一である計算機をグループ化し、さらに上位 2 バイ

ト目が同一である計算機をグループ化することで、4段階の階層を有する階層型データ（図2参照）を構築する。これを平安京ビューで視覚化することにより、ネットワーク上の計算機の分布図を表現する。

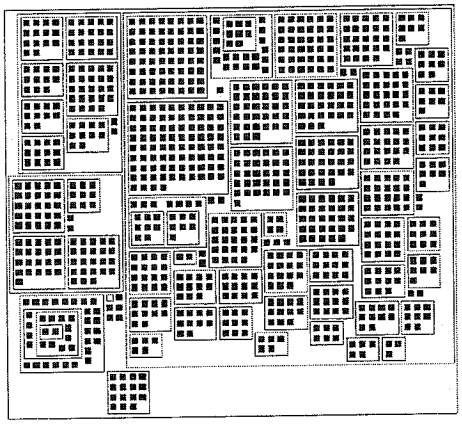

図1「平安京ビュー」を用いた階層型データの視覚化例。

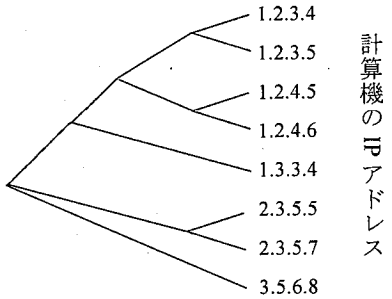

図2 IPアドレスを参照して8台の計算機を階層型データに格納した例

平安京ビューを用いて各計算機をアイコン表示することにより、個々の計算機に関するデータを画面上で重なりなく表現できるとともに、IPアドレスの上位バイトが同一である計算機群（多くの場合、同一組織に属する計算機群）を画面上で一集団として認識することができる。このような表現は、本報告の目的、たとえば

- 数千・数万ものIPアドレスをもつ大規模ネットワーク環境を対象として、個々のIPアドレスに関する不正アクセスの統計的傾向の全貌を一画面に表現したい
- 同一組織に属する計算機群（=IPアドレスの上位2,3桁が同一である計算機群）の不正アクセス傾向を、「不正アクセス群」として概略的に理解したい

というような目的において非常に適切な技術であるといえる。

4. 不正アクセスデータの視覚化

4.1 不正アクセスデータ

本報告で入力データとする不正アクセスデータは、Cisco社のCisco Secure IDS 4320 [Cis]が出力するIDSログデータファイルに基づくものである。当システムではシグネチャ(signature)と呼ばれる典型的なパターンに基づいて不正アクセスを検出するものである。検出された1回の不正アクセス検出に対してログファイルに出力するデータのうち、本手法では、以下のデータを参照する。

- 送信元IPアドレス。
- 受信先IPアドレス。
- 発生時刻。
- 不正の種類を表す正整数ID。
- 危険度を表すレベル(5段階)。

4.2 平安京ビューによる視覚化

本手法では、ログファイルに記述された多数の不正アクセスに対して、まず送信元・受信先に記述されたIPアドレスを列挙し、IPアドレスの各バイトの値を参照して階層構造を構築する。

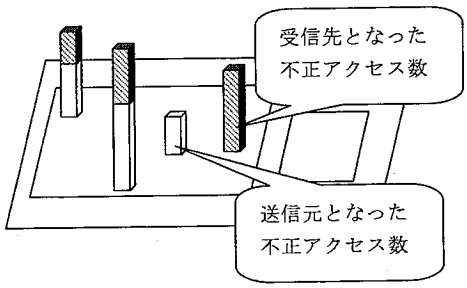

図3 送信元および受信先となった不正アクセス数の表現例。

続いて各々のIPアドレスに対して、送信元となった不正アクセス数、受信先となった不正アクセス数を集計する。このとき、発生時刻の範囲、不正アクセスの種類や重要度、などの条件をユーザーに設定させ、条件を満たす不正アクセス数だけを集計することもできる。

続いてIPアドレスによって構成される階層型データを「平安京ビュー」で表示する。このとき、各々のIPアドレスに相当する葉ノードに高さを

与えることで、各々のIPアドレスの送信元および受信先となった不正アクセス数を表現する。図3に示すように、送信元となった不正アクセス数、受信先となった不正アクセス数、に別個の色を割り当てて表現することで、棒グラフを重ね合わせるように不正アクセス数を表現する。

4.3 拡張機能

不正アクセス視覚化の効果を高めるために、本手法は以下の拡張機能を持つものとする。これらの拡張機能によって、文献[Saw03]に提案されているログ監視支援システムの方針の大半を満たすことができる。具体的には、拡張機能(1)(2)によって、不正アクセスの全体的な統計分布を俯瞰するだけでなく、その詳細分析をも可能にする。また拡張機能(3)によって、管理者による手動操作の手間を低減し、Webブラウザを用いて汎用的な閲覧手段を提供することができる。

(1) 特異な不正アクセスのハイライト

本手法は、数値的に特異な傾向のある不正アクセス群をもつ棒グラフを、特に明るい色で表現する機能をもつ。例えば、

- 非常に多くの受信先に同一送信元から一斉発信している不正アクセス、
- 非常に多くの送信先から同一受信元が一斉受信している不正アクセス
- 直前の時間帯よりも不正アクセス数が急増しているIPアドレス

をハイライトすることで、ネットワーク管理者に注意を促すことができる。

(2) 特定IPアドレスに注視した視覚化

本手法は、特定IPアドレスに相当する棒グラフを画面上でクリックすると、そのIPアドレスを送信元あるいは受信先とする不正アクセスだけを再集計して表示する機能、およびその送信先と受信先を連結する線分を表示する機能をもつ。この機能によりネットワーク管理者は、特定の計算機に関わる不正アクセスの詳細を探索することができる。(ただしこの機能は、下記(3)に紹介するWebブラウザ上でのレポート機能では実現しない。)

(3) 一定時間ごとのレポート機能

筆者らの実装では、一定時間ごと(例えば5分ごろ)に不正アクセス数を集計して視覚化し、その結果をJPEG画像に保存する、という機能をもつ。そのJPEG画像群に対する目次となるようなWebページを自動生成することで、本手法特有のユーザインタフェースを用いることなく、Webブラウザだけで不正アクセスの出現傾向を観察できる。

5. 実行例

本手法を実装した結果を示す。筆者らは、本手法をJava1.4の上で実装し、IBM ThinkPad A31p (CPU 1.7GHz, RAM 756MB)およびMicrosoft Windows 2000の上で実行した。GUIはJava Swingライブラリを用いて実装し、画像生成機能はJava AWTライブラリを用いて実装した。また危険度レベルや発生時刻の範囲などで条件をつけて不正アクセス数を集計するためのGUIを開発した。

以下、図4〜8にて、実在するネットワーク環境で記録したIDSデータを視覚化した例を示す。この例では、各々の計算機について、送信元となった不正アクセス数を青で、受信先となった不正アクセス数を赤で表示している。送信先・受信元ともにゼロであった計算機は、高さゼロの灰色のアイコンで表示している。なお実行例の一部は、白黒印刷された紙面上では効果が認識できないことを断っておく。

図4〜6は、6時間にわたって記録した61822行の不正アクセス記録をもち、3984個のIPアドレスにわたって不正アクセスが検出されたIDSの例である。筆者らの測定では、IDSログファイルの読み込みに120秒、階層型データ構築および平安京ビューの適用に合計約0.6秒、アクセス数の再集計およびJPEG画像生成に平均7.1秒を要した。

図4は、ある時刻から5分間の不正アクセスの集計結果、図5はその直後5分間の集計結果、図6はその2時間後の5分間の集計結果、をそれぞれ視覚化した例である。これらの視覚化結果から、以下のような不正アクセス群を観察できた。

図4の対象時点では、多数の送信元計算機から、多数の受信先計算機に不正アクセスを試していたと思われる傾向が見られる。図5の対象時点では、不正アクセスの発生源をうまく起動できた計算機に、特定の計算機に向けて集中的に攻撃させていることがわかる。図6の対象時点では、図5で送信元となった計算機は既に対処されているものの、他の計算機にも不正アクセス発生源が仕掛けられ、新たに別の計算機も集中的に攻撃されていることがわかる。なお図4〜6において、送信元はネットワーク外の計算機、受信先はネットワーク内の計算機である。

図7,8は、別のIDSデータを視覚化した上で、

特定の棒グラフを画面上でクリックして、その棒グラフに対応する IP アドレスが送信元または受信先になっている不正アクセスを線分で表現した例である。

図 7 に示す例では、1 台の計算機が多数の計算機から不正アクセスを受信しているが、その送信元となっている計算機群の大半では、1,2 回の不正アクセスを受信した後に、多数の不正アクセスを同一計算機に向かって送信していることがわかる。この視覚化結果から、多数の計算機が同様に乗っ取られた結果として、同一計算機が集中攻撃されている可能性がある、と判断して詳しく調べたところ、この不正アクセス自体が一種の誤報であることがわかった。

図 8 に示す例では、1 台の計算機が同一組織に属する複数の計算機から不正アクセスを受信している。この結果は、組織ぐるみで不正アクセスを発生している可能性を示唆している。

6. むすび

本報告では、「平安京ビュー」を用いて、大規模ネットワークへの不正アクセスの統計的傾向を一画面に表示する手法を提案した。

今後の課題として、ネットワーク管理者に本手法を一定期間利用してもらい、ネットワーク管理業務をどのように効率化できるか実証する予定である。また、以下の機能拡張を検討している。

(1) データマイニングやデータベース技術などとの連携による、悪意性の高い不正アクセスや被害の大きい不正アクセスを効果的に警告できる情報視覚化技術の開発。

(2) 5〜10 分といった短期間ではなく、1 週間、1 ヶ月といった長期間を対象とした不正アクセスの傾向分析に向いた情報視覚化技術の開発。

(3) 不正アクセス傾向の時系列変化を効果的に表現する情報視覚化技術の開発。

謝辞

本研究に関して貴重な意見を賜りました京都大学大学院情報学研究科熊谷賢氏に感謝します。

参考文献

[Axe03] Axelsson S., Visualization for Intrusion Detection Hooking the Worm, European Symposium on Research in Computer Security 2003, pp. 309-325, 2003.

[Cis] Cisco Secure IDS. http://www.cisco.com/japanese/warp/public/3/jp/product/security/ids/index.html

[Ito03] 伊藤, 小山田, 平安京ビュー〜階層型データを碁盤状に配置する視覚化手法, 可視化情報学会第 9 回ビジュアリゼーションカンファレンス, 2003.

[Miy02] 宮本, 泉, 田村, 福永, ネットワーク・サーバ運用監視支援システム, システム制御情報学会論文誌, 15, 6, pp. 279-287, 2002.

[Ohy02] 大谷, 桑田, 小迫, 井上, 統合データベースを用いた不正アクセス検出情報の分析および意思決定支援システム, 第 13 回データ工学ワークショップ, A1-6, 2002.

[Saw03] 沢田, 高倉, 岡部, 開放型大規模ネットワークのための IDS ログ監視支援システム, 情報処理学会論文誌, 44, 8, pp. 1861-1871, 2003.

[Tak02] 高田, 小池, 見えログ：情報可視化とテキストマイニングを用いたログ情報ブラウザ, 情報処理学会論文誌, 41, 12, pp. 3265-3275, 2002.

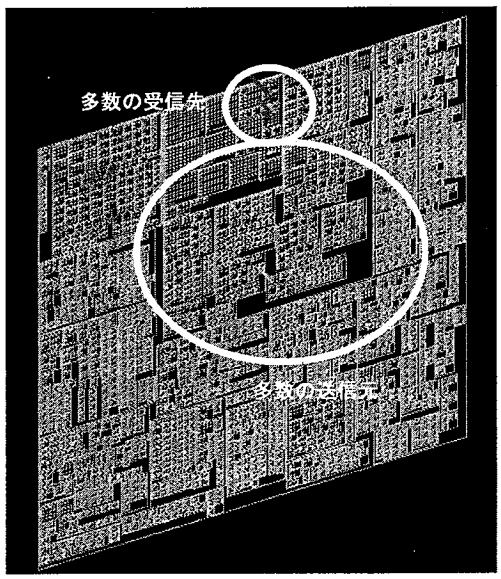

図 4 視覚化例(1)。多数の送信元と多数の受信先の存在が確認できる。

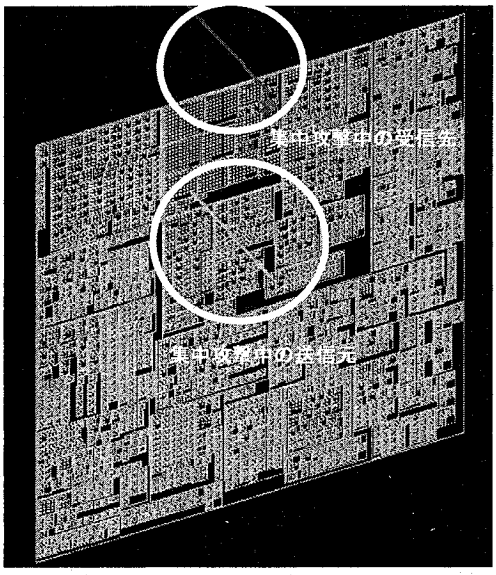

図 5　視覚化例(2)。1 台の計算機が 1 台の計算機を集中攻撃している。

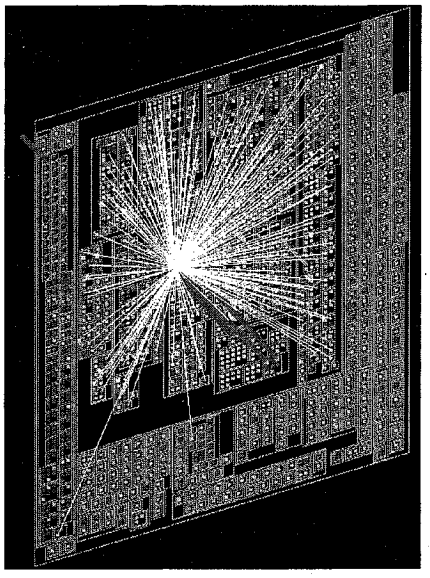

図 7 視覚化例(4)。多数の計算機から 1 台の計算機に不正アクセスを送信しているが、その送信元の大半が送信以前に 1,2 回の不正アクセスを受信していることがわかった。

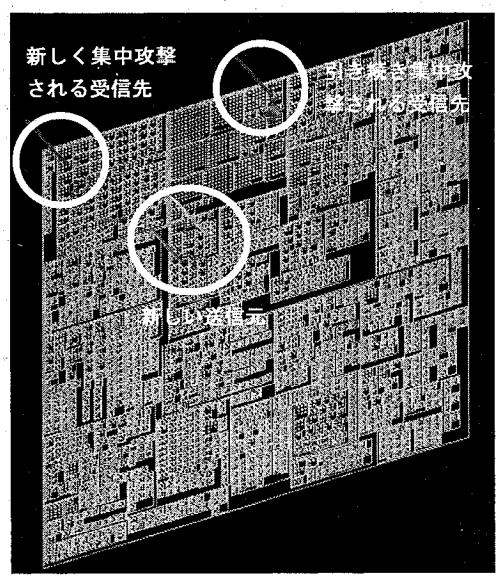

図 6 視覚化例(3)。図 5 の送信元を遮断すると、他の計算機が送信元となって、いくつかの計算機を攻撃しているのがわかる。

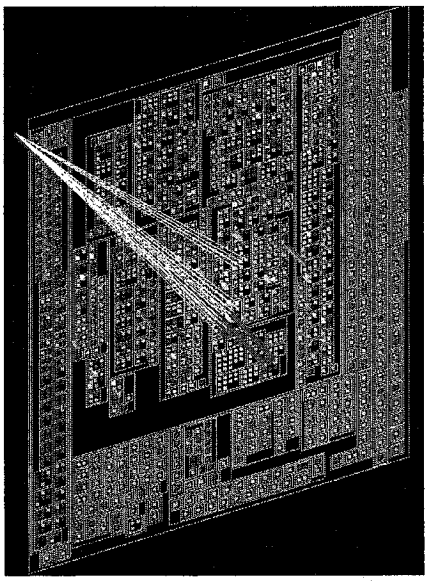

図 8 視覚化例(5)。同一組織の複数の計算機から 1 台の計算機に不正アクセスを集中送信している。