平安京ビューによるIDSデータの視覚化

伊藤 貴之○　高倉 弘喜　沢田 篤史　川原 稔　小山田 耕二（京都大学）

Visualization of IDS data by Heiankyo-View

Takayuki ITOH, Hiroki TAKAKURA, Atsushi SAWADA, Minoru KAWAHARA, Koji KOYAMADA

ABSTRACT

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.

Keywords: Visualization, Heiankyo-View, IDS data

1. はじめに

インターネットのインシデント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。
- 最近のインシデント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）によって一連のインシデント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、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。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 IDS の統計的傾向の理解を支援する手法もいくつか提案されている。高田らの「見えログ」[Tak02]という手法では、画面を水平方向に分割し、左側に IDS 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 IDS の内訳を表示する。また IDS とはデータ構造が異なるが、Axelsson は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]。しかしこれらの手法は、アクセスや計算機などのデータ要素を一次元的に画面に並べるため、数千、数万といった大規模なデータに対しては非常に概略的な傾向しか表現できないことが多い。また「この組織が集中的に狙われている」というような空間的な相関性を理解することが難しい。このことから、データ要素を一次元的にではなく二次元的に画面配置し、大規模な情報を俯瞰するような視覚化手法を適用することが望ましいと考えられる。

本報告では、大規模ネットワークを構成する数千、数万の計算機の IDS による加害、被害の全貌を一画面にすべて表現することを目的とした情報視覚化手法を提案する。本報告の提案手法は、大規模階層型データを対象とした「平安京ビュー」[Ito04]という情報視覚化手法を用いて、数千、数万もの計算機の分布図を表示し、その統計的傾向を一画面に全貌表示する。この手法によって、大量なインシデントの傾向を短時間で把握できると考えられる。

2. 平安京ビュー

前節でも述べたとおり、「平安京ビュー」は大規模階層型データの視覚化手法である。図 1 に「平安京ビュー」

による階層型データの表示例を示す。「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、枝ノードを長方形の枠で表現し、長方形枠の入れ子構造によって階層構造を表現する。

「平安京ビュー」では、まず葉ノードを表現するアイコンを、画面空間上に隙間無く配置する。続いて、この上位階層に属する枝ノードを表現するために、アイコンを包括する長方形を生成する。さらに、上位階層の枝ノードを表現する長方形群を隙間無く配置し、同様にこれを包括する長方形を生成する。以上の処理を、最下位階層から最上位階層に向けて反復することで、データ全体の配置を決定する。図 2 に、「平安京ビュー」による葉ノードおよび枝ノードの画面配置手順を示す。

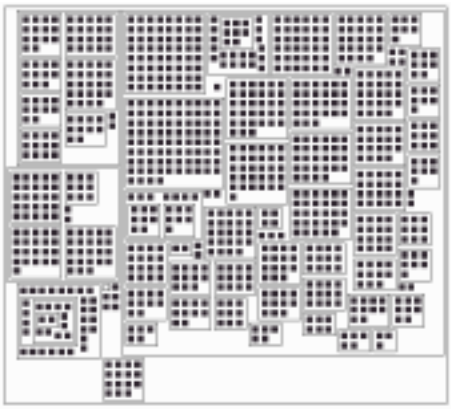

図 1. 平安京ビューによる階層型データの表示例。

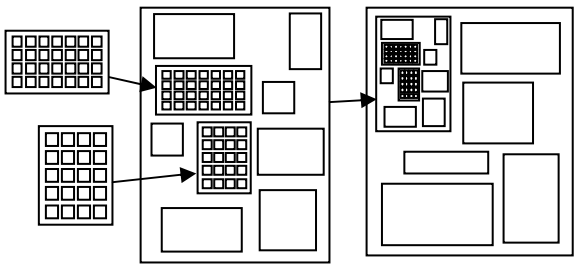

図 2　階層型データの画面配置順。まず最下位階層の葉ノードを配置し、続いて下位階層から上位階層に向かって配置処理を反復する。

このとき「平安京ビュー」では、葉ノードおよび枝ノードを表現する長方形群を、以下の 2 条件を満たすように画面配置する。

[条件 1] すでに配置されている長方形と干渉しない位置に長方形を配置する。

[条件 2] [条件 1]を満たす位置のうち、配置面積の拡大量が最小である位置に長方形を配置する。

この条件を満たすような画面配置を実現することによって、大規模な階層型データの全貌を、限られた画面空間にあますことなく表現することができる。

「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。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 アドレスに関するインシデント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）のインシデント傾向を、「インシデント群」として概略的に理解したい

というような目的において、平安京ビューは非常に適切な技術であるといえる。

3. 「平安京ビュー」による IDS の視覚化

本報告で入力データとする IDS 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インシデントを検出するものである。検出された 1 回のインシデント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ものとする。

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

本手法では、ログファイルに記述された多数のインシデントに対して、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。

続いて各々の IP アドレスに対して、送信元となったインシデント、受信先となったインシデントを集計する。このとき、発生時刻の範囲、インシデントの種類や重要度、などの条件を加え、条件を満たすインシデントだけを集計することもできる。

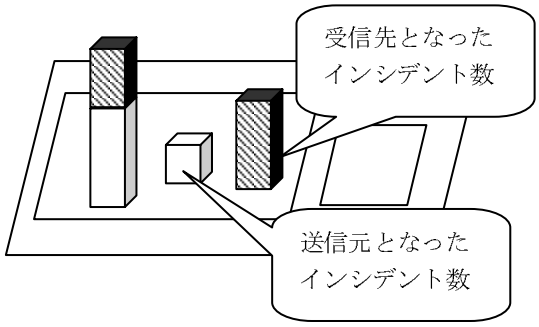

図 3　送信元および受信先となったインシデント数の表現例。

続いて、ネットワークを構成する計算機の 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インシデント数を表現する。図 3 に示すように、送信元となったインシデント数、受信先となったインシデント数、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インシデント数を表現する。

4. 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、3 次元画像生成機能は Java AWT ライブラリを用いて実装した。

筆者らの実験に用いた IDS ログファイルは、実在するネットワーク環境で 6 時間にわたって記録したもので、61822 行のインシデント記録をもち、3984 個の IP アドレスにわたってインシデントが検出されている。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。

以下に提示する画像例では、各々の IP アドレスについて、送信元となったインシデント数を青で、受信先となったインシデント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 IP アドレスに相当する葉ノードは、高さゼロの灰色のアイコンで表示している。

図 4 は、ある時刻から 5 分間のインシデントの集計結果、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、図 7 はさらにその 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。

図 4 の対象時点では、多数の送信元にインシデントの発生源を仕掛けて、多数の受信先にインシデント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、インシデントの発生源をうまく起動できた計算機から、特定の受信先に向けて集中的に攻撃させていることがわかる。図 5 の対象時点では、図 6 で送信元となった計算機は既に対処されているものの、同様なインシデント発生源を他の計算機に仕掛けて、さらに別の受信先も集中的に攻撃していることがわかる。図 7 ではさらに別に、特定のドメイン内にある多数の計算機が横断的にインシデントを受信していることがわかる。

5. おわりに

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インシデントの統計的傾向を一画面に表示する手法を提案した。本手法には以下のような特徴があると考えられる。

- 数千、数万といった大量な計算機を有する大規模ネットワークに対するインシデントの統計的傾向を、一画面に全貌表示できる。
- 計算機空間上の位置と相関性のあるインシデント傾向の理解に向いている。

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、について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。

- データマイニングやデータベース技術との連携により、悪意性の高いインシデントや被害の大きいインシデントを強調提示すること。
- によって提示される情報を効果的に画面表示する情報視覚化技術の開発。
- 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インシデントの傾向分析に向いた情報視覚化技術。

謝辞

「平安京ビュー」の利用方法に関して日本アイ・ビー・エム（株）東京基礎研究所山口裕美氏から貴重な意見を賜ったことを感謝します。

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

[Ito04] 伊藤、山口、小山田、画面空間の格子分割を用いた階層型データ視覚化手法、情報処理学会論文誌、投稿中。

[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、Vol. 15, No. 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インシデント検出情報の分析および意思決定支援システム、第 13 回データ高額ワークショップ、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、Vol. 44, No. 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、Vol. 41, No. 12, pp. 3265-3275, 2002.

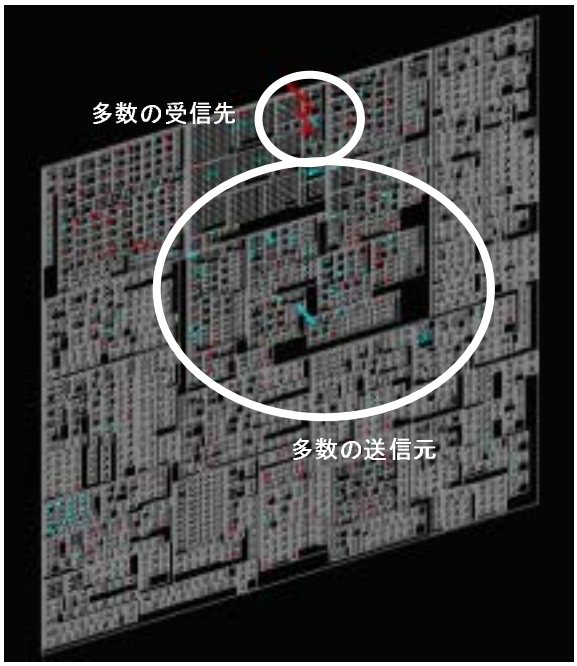

図 4　視覚化例(1)。

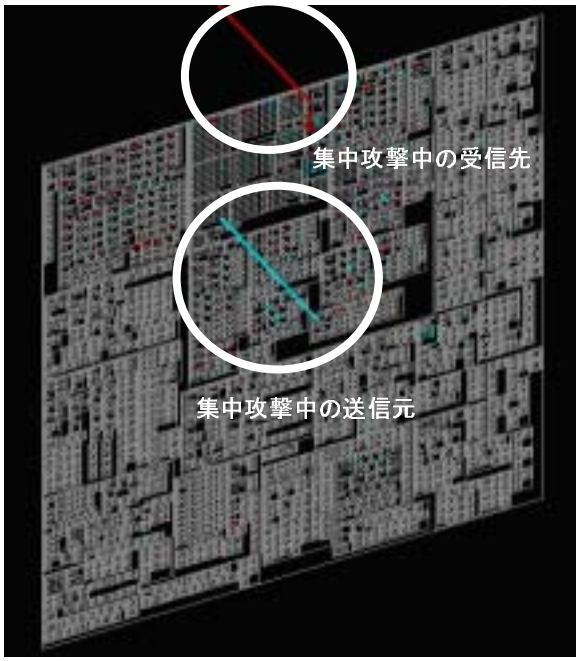

図 5　視覚化例(2)。

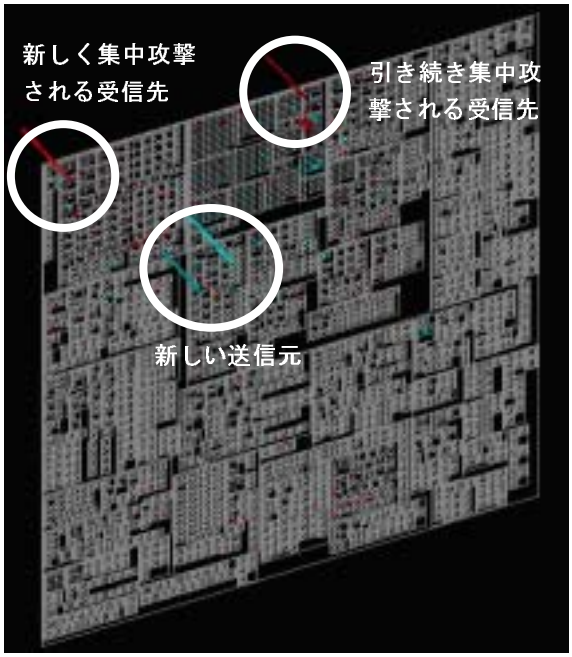

図 6　視覚化例(3)。

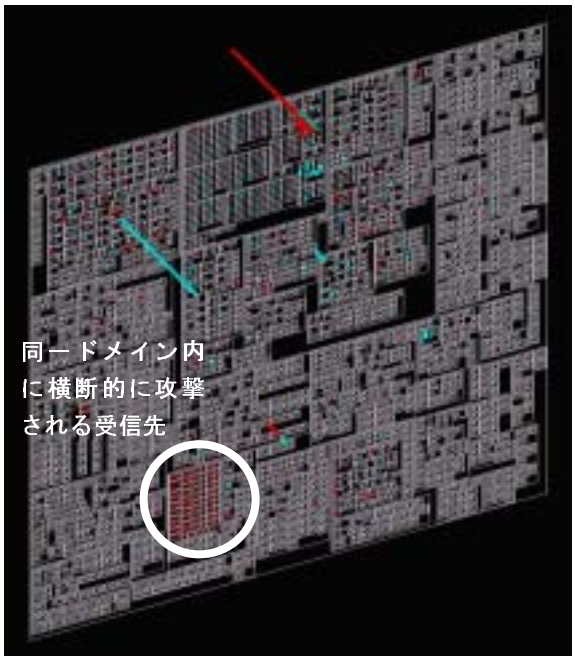

図 7　視覚化例(4)。