平安京ビューによる IDS データの視覚化 ～第 2 報

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）

Visualization of IDS data by HeiankyoView: 2nd report

Takayuki ITOH, Hiroki TAKAKURA, Atsushi SAWADA, 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. We proposed a visualization of IDS data, which constructs hierarchical data according to IP addresses of computers in a target network, and visualizes the data by HeiankyoView. This report presents an extension of the visualization which shows some interesting intrusion behavior. The extension highlights computers which are artificially attacked or seriously damaged. Also, the extension provides a function to represent lines connecting pairs of computers which are senders and receivers of intrusions.

Keywords: Visualization, HeiankyoView, IDS data

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]。

- 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。
- 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、運用上多くの問題がある。

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に、統合管理型データベースを採用した研究[Ohy02]や、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た研究[Miy02]が報告されている。

情報視覚化によって不正アクセスの統計的傾向の理解を支援する手法[Tak02][Axe03]もいくつか報告されている。これらの手法は、不正アクセスの時間的推移や属性を理解するのには向いているが、数千、数万といった膨大な計算機をもつ大規模なネットワークに対して局所的な傾向を理解することが難しい。例えば「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ためには、大規模な計算機ネットワーク空間を俯瞰できるような視覚化手法を適用することが望ましい。

著者らは、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手法[Ito04]を提案している。この手法によって、大量なインシデントの計算機ネットワーク空間上での分布を短時間で把握することができると考えられる。

本報告は[Ito04]の続報として、「平安京ビュー」を用いた IDS 視覚化画面の上で、インシデントの特徴的傾向をさらに強調する手法を提案する。本手法では[Ito04]で視覚化した IDS に対して、以下の機能を追加するものである。

- 特徴的な統計的傾向をもつ計算機をハイライトする機能。
- 特定の計算機を送信元または受信先とするインシデントを計算機間の線分で表示する機能。

本報告ではこの機能が、インシデントに関する以下のような統計的傾向の理解を支援できることを示す。

- 設定ミスや故障等で発生する定期的なエラーと、人為的な集中攻撃インシデントや、ネットワークに影響を与える重度のトラブル、との区別。
- 特定の計算機を送信元とするインシデントの、受信先との空間的相関性。または特定の計算機を受信先とするインシデントの、送信元との空間的相関性。

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。

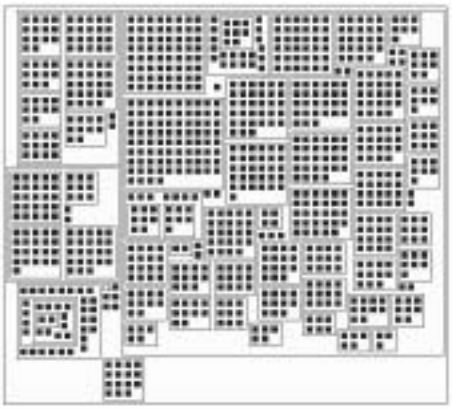

図 1. 平安京ビューによる階層型データの表示例。

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 視覚化

続いて前報で報告した IDS 視覚化手法[Ito04]について述べる。本手法が入力データとする 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。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。
- 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)。

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。

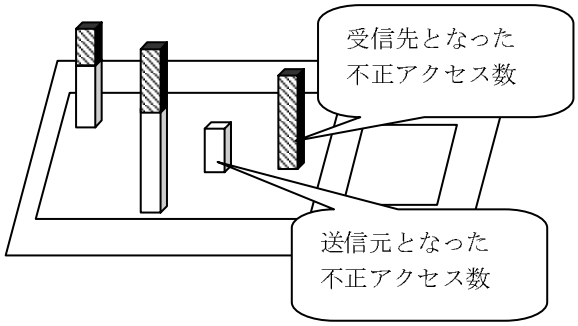

図 2. 送信元および受信先となったインシデント数の表現例。

続いて、ネットワークに接続している計算機の 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 2 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. IDS 視覚化の機能拡張

続いて前章で述べた IDS 視覚化手法に対する新しい機能拡張について述べる。本報告では以下の機能拡張を試みた。

[機能拡張 1] 特徴的な統計的傾向をもつ計算機をハイライトする機能。

本機能では、一定のルールを満たすインシデント群を有する計算機を検出する、という処理を行うプラグイン関数を実装できるようにした。このプラグイン関数が検出した計算機を色濃く表示すること

で、特徴的な統計的傾向をもつ計算機をハイライトする。本報告では、非常に短い時間間隔でインシデントを *n* 回以上連続送信（または受信）した計算機を検出する、というプラグイン関数を実装した。この関数は、人為的かつ執拗な攻撃や、影響の大きい重度のトラブル、などをハイライトする効果があると考えられる。

[機能拡張 2] 特定の計算機を送信元または受信先とするインシデントを計算機間の線分で表示する機能。

本機能では、特定の計算機（例えば画面上でマウスをクリックして指定）を送信元とするインシデントの受信先計算機、あるいは特定の計算機を受信先とするインシデントの送信元計算機、を画面上で線分表示する。この機能により、特徴的な攻撃を受けている（または発している）計算機の攻撃元（あるいが攻撃先）が、広く分布しているのか、特定の組織に集中しているのか、1 台の計算機に特定しているのか、を理解する効果があると考えられる。

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、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。

以下に示す実行例は、実在する計算機ネットワークの 1 時間の IDS ログファイルを用いたものである。このログファイルには、1569 台の計算機に関わる 30306 行のインシデントが記録されている。なお実行例の一部は、白黒印刷された紙面上では効果を認識できないことを断っておく。

図 3 は 1 時間のインシデントの総計を表示したものである。白丸の部分に、受信先となったインシデントを多数もつ計算機が 4 台見られるが、このうち 1 台は棒グラフの色が暗くなっている。この 1 台の受信したインシデントは、執拗な人為的攻撃や、影響の大きい重度なトラブルである可能性が低いので、他の 3 台よりも危険度が低い、と考えられる。

図 4 は、図 3 に白丸で示した 4 台のうち、棒グラフの色が明るくなっている計算機のうち 1 台を指定して、その送信先を黄色い線で示したものである。この結果から、特定の計算機に不正アクセスを送信している計算機は非常に多数あるが、その大半は広い意味での同一組織内の計算機であることがわかる。また、この送信元である計算機のインシデントに着目すると、その多くの計算機は 1,2 回のインシデントを受信するとともに、多数のインシデントを同一計算機に送信していることがわかる。この事実は、送信元である多数の計算機が、インシデントを受信したことがきっかけで送信者に転じた可能性を示唆している。

図 5 は、インシデントを多数受信している別の計算機を指定した例である。この計算機は、狭い意味での同一組織の多数の計算機からインシデントを受信している。詳しく調べたところ、このインシデントは悪意的な攻撃ではなく、組織的に必要に迫られて同一計算機に例外的なアクセスをしていたのが、インシデントとして検出されていた、ということがわかった。

6. おわりに

本報告では、「平安京ビュー」を用いたインシデントの視覚化手法[Ito04]に対して

- 統計的特徴をもつ計算機のハイライト
- インシデントを送受信する計算機間の線分表示

の 2 つの拡張を行った。またこの結果として、以下の効果を得ることができた。

- 設定ミスや故障等で発生する定期的なインシデントと、人為的な集中攻撃インシデントや、ネットワークに影響を与える重度のトラブル、との区別。
- 特定の計算機を送信元とするインシデントの、受信先との空間的相関性。または特定の計算機を受信先とするインシデントの、送信元との空間的相関性。

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の特定を高性能化する手法。
- 悪意性や被害の小さいインシデントの表示を目立たなくする表示技術。
- 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。
- 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

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/secu

rity/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.
[Ito04] 伊藤, 高倉, 沢田, 川原, 小山田, 平安京ビューによる IDS データの視覚化, 第 32 回可視化情報シンポジウム, 2004.
[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.

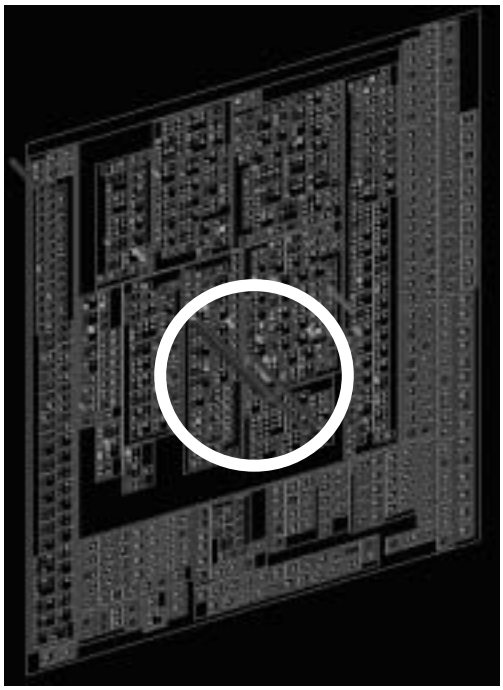

図 3 視覚化例(1)。多数の攻撃を受けた計算機が白丸の中に 4 台見られるが、そのうち 1 台は棒グラフの色が暗いので、他の 3 台より危険度が低いと認識することができる。

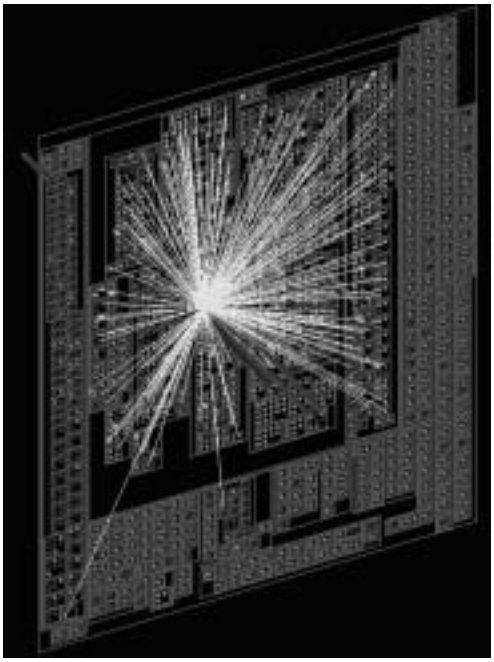

図 4　視覚化例(2)。多数の計算機から 1 台の計算機に不正アクセスを送信しているが、その送信元の大半が送信以前に 1,2 回の不正アクセスを受信していることがわかった。

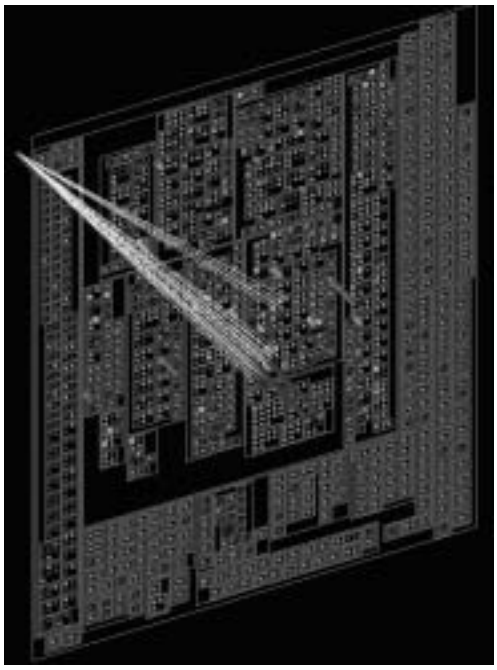

図 5　視覚化例(3)。同一組織の複数の計算機から 1 台の計算機に不正アクセスを送信している。