Network Flow データを用いた不正アクセス監視のための可視化の一手法

菊池愛美　（指導教員：伊藤貴之）

1. 研究背景と目的

近年，侵入検知システム(Intrusion Detection System．以下 IDS)は幅広く普及し，インターネット上の不正アクセスによる深刻な被害のために不正アクセスの検知と解析の研究が進んでいる．ファイアウォールシステムのような一定の安全対策が施されていることを前提に，IDS は不正現象を検知して記録する．IDS のログデータを用いた既存の可視化手法として，伊藤らによる手法[1]や IDGraph[2]などがある．これらの手法は，IDS によって検知できる既知の現象を可視化することはできるが，新手の攻撃を可視化できない．また，すべてのネットワークに IDS が導入されているとは限らないことも，考慮しなければならない．

本論文では，新手の侵入や攻撃を検知することを目的として, Network Flow データを用いた可視化手法を提案する．上述の手法[1]と同様に，大規模階層型データの可視化手法である平安京ビューを用いることで，何千何万というコンピュータを含む大規模な LAN 上の計算機群を一画面表示し，そのネットワーク上の侵入の統計を表現することができる．なお本論文では，実在する大規模ネットワークのログデータを用いた結果を示す．

2. 提案手法の概要

Network Flow データは，悪意のないものを含むすべてのアクセスを記録し，しかもデータの種類は IDS データのものより多いので，Network Flow データから得られる情報は莫大で複雑である．

本論文で用いる Network Flow データの変数は以下のとおりである．ID，年月日，時間，送信者アドレス，受信者アドレス，プロトコル，データサイズ，MD5 値，Buffer Overflow(以下 BO)情報，shellcode 情報，IDS で検知された情報に相当する ID．なお，Network Flow のログに記録される変数はシステムによって異なる．上記の変数はあくまでも，本研究にて使用するネットワーク上のログに記録される変数にすぎない．

提案手法ではまずはじめに, 各 IP アドレスへのアクセスを集計し，コンピュータの階層構造を構築し，平安京ビューを用いて階層構造を可視化する．そして，ログデータに記録されているすべてのアクセスから BO か IDS の情報を持っているアクセスを選別し，その集計結果も平安京ビューで可視化する．IDS を導入していないネットワークであっても，本手法は BO を引き起こすアクセスを可視化することができる．しかし，本手法は BO と IDS の情報の関係性を一画面上に可視化できるため，ネットワークが IDS 情報を持つならばより効果的に可視化することができる．

ここで, shellcode と IDS の signature ID という 2 つの変数に着目する．BO を引き起こしたアクセスに対して，IDS が検知した攻撃の種類の数が，

- 0 個のとき，知られていない攻撃
- 1 個のとき，特定の標的への攻撃
- 2 個以上のとき，攻撃目標の模索

となる可能性がある．また同様に，検出された shellcode と同じ shellcode が過去に何度検知されているかについて，

- shellcode が検出されない
- 初めて検知する
- 過去に検知したことがある

という 3 種類に区別して考える．shellcode が検出されないログは単純な BO であると考えられる．以上 2 種類の変数について，それぞれの攻撃を一画面で観察できるような可視化結果を目指す．また，データ転送時のトランスポート層プロトコル型も，不正アクセスを発見するための手がかりになることがある．そこで各ログのプロトコル型（TCP または UDP）も，可視化の際に参照する．

提案手法では，以下の 3 種類の属性について 1 個ずつ選択することで，全 27 通りの可視化結果を表示できる．

- 送信者側，受信者側，両方
- IDS の signature ID，shellcode，両方
- TCP，UDP，両方

表 1：送信者情報における色分けの詳細

送信者情報	signature ID 0 個	signature ID 1 個	signature ID 2 個以上
shellcode 初めて検知	赤(濃)	緑(濃)	青(濃)
shellcode 過去にも検知	赤(淡)	緑(淡)	青(淡)

表 2：受信者情報における色分けの詳細

受信者情報	signature ID 0 個	signature ID 1 個	signature ID 2 個以上
shellcode 初めて検知	マゼンタ(濃)	黄(濃)	シアン(濃)
shellcode 過去にも検知	マゼンタ(淡)	黄(淡)	シアン(淡)

表 1 及び表 2 に, IDS の signature ID と shellcode の両方を可視化した場合の色分けについての詳細を示す．片方のみを選んだときには,選ばれなかった属性（送信者側または受信

者側，TCP または UDP）を表す棒グラフは，灰色で表示される．また，棒グラフをクリックすると，クリックされた棒グラフと送信者=受信者の関係にある棒グラフを橙色の線で表示する．また，クリックした棒グラフに該当するコンピュータの IP アドレスを，紫色の文字列で表示する．

3. 実行結果

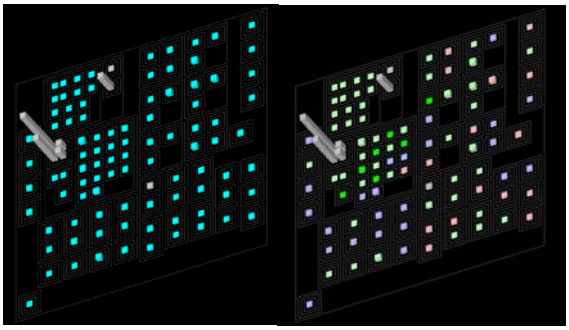

図 1：既存手法との比較
(左)既存手法[1]と(右)提案手法

図 1 は, 既存の可視化手法[1](左)と提案手法(右)をそれぞれ用いて，同一期間での送信者の情報を可視化した結果の一例である．既存手法が送信したアクセス数を水色で描いているだけであるのに対し，提案手法では IDS の signature ID の個数や shellcode の回数によってアクセスを分別し，色分けして表示している．これにより，ユーザは大量のログデータから着目すべきコンピュータを正しく選別できる可能性が高まると思われる．

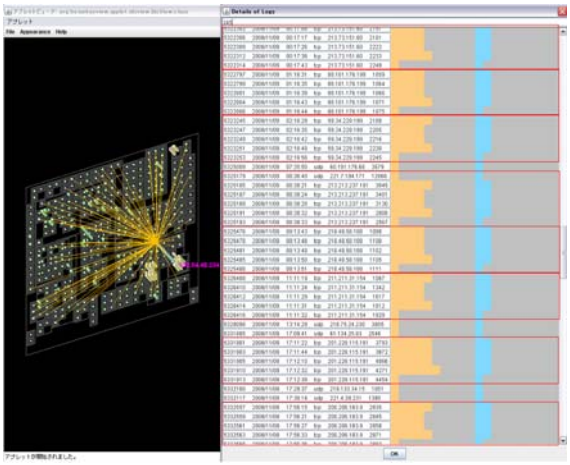

図 2：詳細情報の可視化結果

図 2 は，特定の計算機に関するアクセスについて，IDS の signature ID の個数や shellcode の個数についての詳細を可視化した結果である．左側にはユーザが指定した条件に合致するログの ID，年月日，時間，プロトコル型，アクセス先の IP アドレス, ポート番号を時間経過に沿って表示し，右側にはそのログにおける IDS で検知された signature ID の個数を橙色で，shellcode の個数を水色で，それぞれ帯グラフを用いて表現している．このような詳細情報の可視化により，時間推移に伴うアクセス内容の変化や，そのパターンを発見できると考えられる．実際に図 2 において，送信者の IP アドレスが同一なアクセス群ごとに赤い線で区切ると，2 色の帯グラフ間に連動した動きがあり，かつ同様なパターンが反復されているのがわかる．

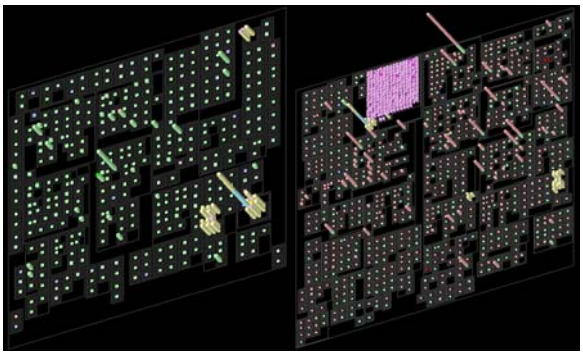

図 3：可視化結果
(左) 2006 年 11 月第 1 週 (右) 2006 年 11 月第 2 週

図 3 は，2006 年 11 月の第 1 週及び第 2 週のアクセス分布を可視化した結果である．左図の第 1 週では棒グラフの多くが緑色で塗られているのに対し，右図の第 2 週では赤色の棒グラフが多くを占めており，アクセスに関わったコンピュータの数も激増している．調べてみると，この時期に WindowsXP の Service Pack 2 にセキュリティ上の脆弱性が存在することが発覚していた．そのため，IDS に検知されない攻撃が多発していたと考えられる．

4. まとめ

本論文では，Network Flow データを用いて不正アクセスを監視するための可視化手法を提案した．本手法では，shellcode と signature ID の 2 種類の変数に着目し，平安京ビューを用いて可視化した．その際，一定期間内のアクセス統計を高さで，同一期間内のアクセスの内容については色分けをして表示した．また，各コンピュータにおいて 2 つの変数の時間推移に関する詳細情報を可視化した．この結果から，ユーザが大量のログデータから疑わしいアクセスを，より短時間で簡単に発見することが可能になると思われる．

参考文献

[1] T. Itoh, H. Takakura, A. Sawada, and K. Koyamada, Hierarchical Visualization of Network intrusion detection data in the IP Address Space, IEEE Computer Graphics and Applications, Vol. 26, No. 2, pp. 40-47, 2006.

[2] P. Ren, Y. Geo, Z. Li, Y. Chen, and B. Watson, IDGraphs: Intrusion Detection and Analysis Using Histgraphs, Workshop on Visualization for Computer Security, pp. 39-46, 2005.