異種アラートの相関に基づくログ分析型IDSの試作と評価

M2005MM032 竹内 孝
指導教員　長谷川 利治

1　はじめに

近年, サイバーテロ, セキュリティ侵害の増加にともない, ネットワークを通過する不正なトラヒックを監視するネットワーク型侵入検知システム (NIDS: Network based Intrusion Detection System) への関心が高まっている.

従来の NIDS は, 監視するパケットと攻撃の特徴が記録されたシグネチャを比較して攻撃の検出を行なっている. この手法は, シグネチャに記録されていない攻撃を検出できないという問題点がある. そのため, [5] や一部の商用 IDS[3] ではシグネチャ型 IDS の弱点を補うために, シグネチャ型 IDS と未知の攻撃を検出できる可能性のある異常検出システムをハイブリッド構成する方式が採用されている. しかし, この方法では 2 台の IDS それぞれが誤検知を大量に含むアラートを出力するため, その中から侵入に成功したことを示す少数のアラートを見つけ出すことは非常に困難な作業となる. この問題を解決するために, これまでにもセキュリティ機器が出力するアラートを分析する手法がいくつか提案されているが, 有効な解決策はまだ見つかっていない.

そこで本研究では, 既知攻撃を検出するシグネチャ型 IDS とシグネチャに依存することなく攻撃を検出する異常検出システムを組み合わせ, 両 IDS が生成したアラートの相関を分析するシステムを提案する. 検出方式が異なる 2 台の IDS が出力した異種のアラートを 1 つのデータベースへ記録する. その後, ログ分析型 IDS はアラートの総量を減少させるために, アラートの類似性に基づいた相関を行なう. つづいて攻撃に成功した可能性の高いアラートを検出するために, 異種のアラートに対して攻撃方法の順番に着目した相関を行なう. そして, 攻撃に成功した可能性の高いアラートに対して高いプライオリティを与える. 本提案システムに基づいてプロトタイプを試作して評価する. 本研究によって, 誤検知を大量に含むアラートの中から, 攻撃に成功した少数のアラートを検出することが可能となる.

2　異種アラートの相関に基づくログ分析型 IDS

本研究では, 異機種の IDS が出力した誤検知を大量に含むアラートの中から, 侵入に成功した疑いの強いアラートを見つけ出すことを目的として, 異種アラートの相関に基づくログ分析型 IDS を提案する. この節では, まずシステム構成について述べ, つづいてシステムを構成する各コンポーネントについて述べる.

2.1　システム構成

近年, アラートの分析に対する研究が活発に行なわれている. Kruegel らが提案した分析手法 [2] は, 主にシグネチャ型 IDS が出力するアラートを対象としている. しかし, 実際のネットワーク環境では対応するシグネチャファイルが存在しない未知の攻撃も存在するため, シグネチャ型 IDS だけでは対応できない. 本提案システムは, シグネチャ型 IDS と未知攻撃を検出できる異常検出システムを組み合わせ, 同一パケットに対して 2 台の IDS で攻撃判定を行なう. そして, 2 台の IDS が出力した異種のアラートに対して相関に基づいた分析を行なうことで, 膨大なアラートの中から攻撃に成功したことを示す少数のアラートを検出する. 本提案システムの構成図を図 1 に示す. 本提案システムをインターネット接続との境界付近に配置し, キャプチャしたパケットを以下の手順で処理する.

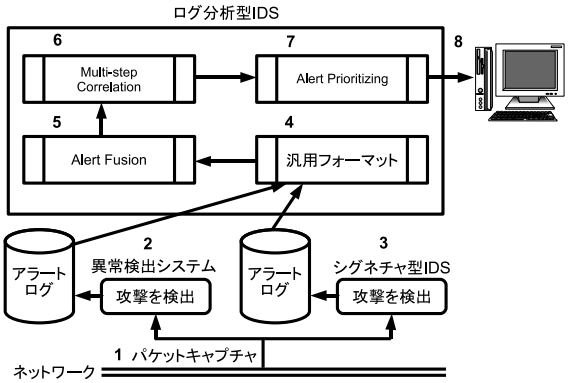

図 1　提案システムの構成

1. ネットワークを流れるパケットをキャプチャする.
2. キャプチャしたパケットをシグネチャ型 IDS で攻撃判定する.
3. キャプチャしたパケットを異常検出システムで攻撃判定する.
4. 2 台の IDS が出力した異種のアラートを統一した汎用フォーマットに整形する.
5. アラートの類似性を基にクラスタリングを行ない, アラートの全体量を減少させる.
6. 攻撃手順に着目した相関を行ない, 膨大なアラートの中から攻撃に成功した可能性の高いアラートを検出する.
7. アラートの重要度に基づいて, アラート毎にプライオリティを割り当てる.

8. 高いプライオリティが割り当てられたアラートを管理者へ通知する

2.2 異常検出システム

本提案システムでは, シグネチャ型 IDS と組み合わせる異常検出システム (非シグネチャ型 IDS) の一例として, ペイロード値の出現頻度分析に基づく, 簡単な IDS を試作した. これまでにも異常検出システムにはさまざまな手法が提案されている. その中で本提案システムは, 検出手法が単純で実装が容易であるという理由で Wang ら [4] の手法を用いた. その検出手法を図 2 に示す. まず, 事前学習処理として, 攻撃を含まない学習データからポート毎のアプリケーションペイロードの長さ別に, 各コードに該当するコードが何文字あるかをコード毎にカウントし, そのコード毎に平均と標準偏差を求める. この求めた値が正常時のプロファイルとなる. 検出フェーズでは, 事前学習処理で学習データに対して行なったのと同様の計算をネットワークトラヒックから新たにキャプチャしたデータに対して行なう. そして特徴ベクトルを求めるために, パケットペイロード毎のコード別の頻度の平均を計算する. 事前学習処理で生成したプロファイルと新たにキャプチャしたデータの距離を計算する. 算出した距離が定められた閾値を越えたら攻撃と判定する.

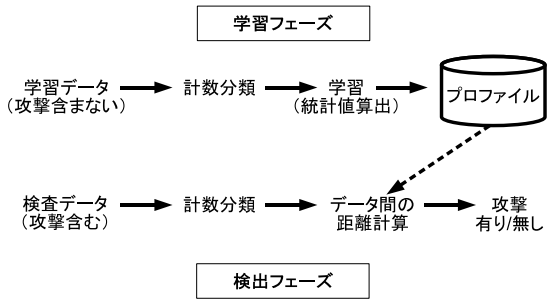

図 2 本提案システムで用いた異常検出システムの検知手法

2.3 ログ分析型 IDS

ログ分析型 IDS は, 異機種の IDS が出力した異種のアラートを相関し, 危険度の高いアラートを取り出す作業を行なう. ここでは, 異種のアラートを統一形式のログフォーマットへの整形を行なう汎用フォーマット, アラートの類似性に基づいて相関を行なう Alert Fusion, 攻撃手法の順序に着目した相関を行なう Multi-step Correlation, そして相関したそれぞれのアラートにプライオリティを割り当てる Alert Prioritizing について述べる.

■汎用フォーマット 汎用フォーマットは, 2 台の検出手法が異なる IDS が出力したアラートを相関し易くするために統一形式の汎用フォーマットに整えることを目的とする. シグネチャ型 IDS は, シグネチャ (ルールファイル) に基づいた攻撃判定を行なうため, アラートに攻撃名を持つ. 一方, 異常検出システムのアラートはパケットヘッダを構成する要素をアラートに持ち, アラートに攻撃名を持たない. 本提案システムで用いる汎用フォーマットを表 1 に示す. これは, この後の相関フェーズでアラートの相関を行ない易くするために, それぞれのアラートから長所を取り出して構成した.

表 1 汎用フォーマットの属性

属性	説明
id	各アラートに対して割り当てられた一意の ID
start_time	最初のパケットを受信した時間
end_time	最後のパケットを受信した時間
time_range	最初のパケットを受信した時間から最後のパケットを受信した時間までの間隔
name	Meta-Alert を生成した際にその旨を記述
attack_name	シグネチャ型 IDS が出力したアラートの攻撃名
src_ip	送信元アドレス
src_port	送信元ポート
dest_ip	送信先アドレス
dest_port	送信先ポート
protocol	IP プロトコルタイプ
sensor	アラートを出力したセンサー
data_length	パケットデータ部の長さ
flag	TCP フラグ
cnt	Meta-Alert に含まれるパケット数
tag	Meta-Alert を構成する参照元の ID を記述

■Alert Fusion 統一形式の汎用フォーマットに整えられたアラートは, Alert Fusion コンポーネントに送られる. ネットワーク型 IDS は, 誤検知を大量に含む膨大なアラートを出力する. Alert Fusion では, アラートの類似性に基づいた相関を行ない, 類似したアラートをクラスタリングすることで, アラートの総量を減らす処理を行なう. 本提案システムでは, 表 2 に示す属性がまったく同じときにアラートをクラスタリングする.

表 2 Alert Fusion で用いる属性

属性	説明
src_ip	送信元アドレス
src_port	送信元ポート
dest_ip	送信先アドレス
dest_port	送信先ポート
protocol	IP プロトコルタイプ

■Multi-step Correlation　Multi-step Correlation は, 攻撃手法の順序に着目した相関を行なうことで, 膨大なアラートの中から攻撃に成功した可能性のある少数のアラートを検出することを目的とする. 攻撃者はまず相手ホストにダメージを加える, または侵入を試みるときに事前の準備としてポートスキャンを行なう. 次に, 攻撃者は相手ホストの弱点を狙うリモート攻撃を行なうことで侵入を企てる. 最後に, 相手ホストへ侵入に成功した攻撃者は, システムの全権を掌握するために root 権限の取得を狙う. Multi-step Correlation は, ポートスキャン, リモート攻撃, root 権限の取得を示す各アラートを 1 つのアラートとして関連づける. 関連づけられたアラートは, 侵入に成功した可能性が高いことを示す.

■Alert Prioritizing　最後に, Alert Prioritizing フェーズで相関済みの各アラートに対してプライオリティを割り当てる. 高いプライオリティが割り当てられたアラート, つまり侵入に成功した可能性の高いアラートは管理者へ通知される.

3　実装

提案システムを実装し, プロトタイプを試作した.

シグネチャ型 IDS は, オープンソースの Snort[1] を利用した. 機械学習を用いた異常検出システム, ログ分析型 IDS は C 言語で実装した. また, データベースにはオープンソースの PostgreSQL を利用した.

4　評価

試作したプロトタイプを用いて, 提案システムを評価した.

4.1　評価用データ

提案システムを評価するために, 大学 LAN ゲートウェイから収集したデータ, 著者の自宅ネットワーク環境から収集したデータを利用した. 前者は, 2004 年 1 月にデータを収集し, 後者は 2007 年 1 月から 2 月にかけてデータを収集した.

4.2　実験結果

■大学 LAN ゲートウェイ　5 日分のデータを用いて本提案システムの評価を行なった. 2 台の IDS は合わせて 386,960 アラートを生成した. その内訳は, 不正検出システムの Snort が 80,410 アラート, 機械学習を用いた異常検出システムが 306,550 アラートを出力した. これらの 386,960 アラートから本提案システムは Alert Fusion により 11,709 の Meta-Alert を生成した. それから, Alert Fusion で生成した Meta-Alert と残りの他のアラートを Multi-step Correlation で相関し, 新たに 9 個の Meta-Alert を生成した. 本提案システムは, 最終的にこの 9 個のアラート全てに対してプライオリティ High を割り当てた.

本提案システムがプライオリティ High を割り当てた 9 個のアラートに基づき, アラートログ及び tcpdump データに対して目視で調査を行なった結果, 大学ネットワークに対する攻撃は検出されなかった.

■自宅環境　2 月 1 日から 2 月 2 日にかけて収集したデータを用いて提案システムを評価した. アラートは全部で 181 アラートあり, その内訳は Snort が 12 アラート, 機械学習を用いた異常検出システムが 169 アラートを出力した. これら 181 のアラートから提案システムは Alert Fusion により 5 個の Meta-Alert を生成した. それから Multi-step Correlation を行なった結果, 提案システムは新たに 3 個の Meta-Alert を生成した. 提案システムはこの 3 個の Meta-Alert に対して プライオリティ High を割り当てた.

プライオリティ High が割り当てられた 3 個の Meta-Alert は, すべて同一のタイムウィンドウ内で集中的に検出された. そして本提案システムがプライオリティ High を割り当てたこれら 3 個のアラートに基づいて, アラートログ及び tcpdump データを目視で調査した結果, これら 3 個のアラートは, Windows の脆弱性として公開されている MS04-007 に対するコンピュータウィルス (以下, ウィルスと呼ぶ) による攻撃であった.

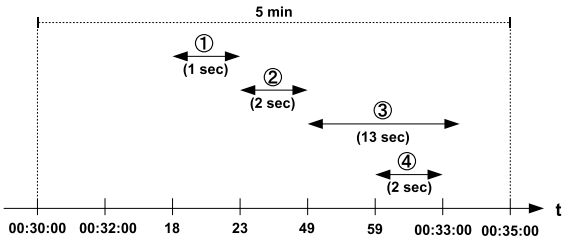

図 3　コンピュータウィルスによる攻撃活動の推移

本提案システムは, 2007 年 2 月 2 日 0 時 32 分から 33 分にかけてウィルスの活動を検出した. 図 3 は, ウィルスに感染したホストから著者が自宅ネットワーク環境でトラヒックデータ収集のために用意したマシンへの攻撃活動の推移を時系列にまとめたものである. 以下に攻撃の分析の結果をまとめる.

1. 00 時 32 分 18 秒 〜 : RPC へのクエリ (1 秒間)
2. 00 時 32 分 23 秒 〜 : Kerberos へコネクション (2 秒間)
3. 00 時 32 分 49 秒 〜 : HTTP への攻撃活動 (13 秒間)
4. 00 時 32 分 59 秒 〜 : ポート 31337 へコネクション (2 秒間)

(1) の期間にウィルス感染元ホストから著者が設置したトラヒック収集マシンの RPC サービスへクエリが発生した. つづいて (2) で, Kerberos へ接続した. そして (3) の期間に, ウィルス感染元ホストは著者のトラヒック収集マシンの HTTP サービスへ約 13 秒間の攻撃を行なった. この期間の tcpdump データを調査した結果, Windows の脆弱性を狙ったウィルスによる攻撃活動であることがわかった. この攻撃活動に対して, 本提案シ

ステムは複数のアラートを出力した．機械学習を用いた異常検出システムが算出した通常パケットに対する攻撃パケットのマハラノビス距離を図 4 に示す．図 5 で示すようにウィルスのパケットペイロードは無意味な文字列の繰り返しで構成されているため，通常パケットのマハラノビス距離と比較した場合に明確な距離の違いがみられた．また，Snort は以下の 2 種類のアラートを出力した．

- (http_inspect) OVERSIZE REQUEST-URI DIRECTORY
- (http_inspect) BARE BYTE UNICODE ENCODING

最後に (4) では，ウィルス感染元ホストはトラヒック収集マシンのポート 31337 へ接続した．これは，ウィルス感染元ホストからトラヒック収集マシンへリモートアクセスの試みがあったことを示す．そして，Snort はこの攻撃活動に対して TCP Portscan のアラートを出力した．

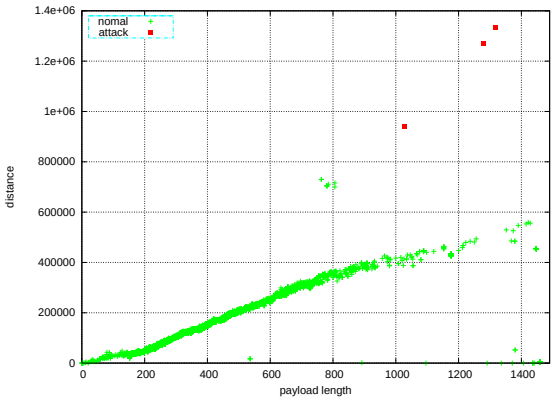

図 4　通常パケットと攻撃パケットのマハラノビス距離

本提案システムは，あらかじめ定められた期間内 (実験では 5 分とした) にアラートの相関に基づく分析を行なうように設計した．このウィルスの検知例では，それが 2007 年 2 月 2 日 0 時 30 分から 35 分にかけての 5 分間である．そして本提案システムはこの期間内に出力したアラートの分析作業を行ない，ポートスキャンを示すアラートと HTTP サービスへのリモート攻撃の活動を示すアラートを相関に基づいた分析によって関連づけることで Multi-step Attack を検出した．そして Multi-step Attack を示すアラートに対して高いプライオリティを割り当てた．これにより，本提案システムは自己感染機能を持つコンピュータウィルスによる一連の攻撃活動を正しく検出した．

5　おわりに

本論文では，NIDS が出力する膨大なアラートの中から侵入に成功した可能性の高いアラートを検出すること

```
00:32:49.844392 IP 203.91.177.53.2521 > 203.91.188.31.http: . 2920:4380(1460) ack 1 win 64240
    0x0000:  000b 9794 5795 0005 31f9 717b 0800 4500  ....W...1.q{..E.
    0x0010:  05dc 79cf 4000 7e06 7940 cb5b b135 cb5b  ..y.@.~.y@.[.5.[
    0x0020:  bc1f 09d9 0050 f36c 0208 2d67 67c0 5010  .....P.l..-gg.P.
    0x0030:  faf0 f522 0000 4141 2f7a 5a6f 4352 4c57  ...".."..AA/zZoCRLW
    0x0040:  592b 6a33 4141 4141 6955 5949 364b 4941  Y+j3AAAAiUYI6KIA
    0x0050:  4141 442f 6467 526f 6139 4172 7975 6a69  AAD/dgRoa9Aryuji

                                （中略）

    0x01e0:  6977 784c 6931 6f63 4165 754c 4249 7342  iwxLi1ocAeuLBIsB
    0x01f0:  3649 6c45 4a42 7868 7767 6741 362f 3544  6IlEJBxhwggA6/5D
    0x0200:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0210:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0220:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0230:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0240:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0250:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0260:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0270:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0280:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0290:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02a0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02b0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02c0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02d0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02e0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x02f0:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0300:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0310:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0320:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0330:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0340:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0350:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0360:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND
    0x0370:  5130 4e44 5130 4e44 5130 4e44 5130 4e44  Q0NDQ0NDQ0NDQ0ND

                          （以下，これの繰り返し）
```

図 5　コンピュータウィルスのパケットペイロードの一例

を目的に，異種のアラートログに対して相関に基づいた分析を行なうログ分析型 IDS について述べた．

そして本提案システムは，大量のアラートの中からコンピュータウィルスによる攻撃活動を示す少数のアラートを検出できることを示した．

一方で，提案システムを実装して実験を行なった結果，アラートの相関を行なう際のタイムウィンドウの設定の難しさ，攻撃パターンおよびアラート分析の良いアイデアを得るための評価データの必要性といった問題点も明らかになった．

参考文献

[1] Caswell, B. and Roesch, M.: Snort. The Open Source Network Intrusion Detection System, http://www.snort.org/.

[2] Kruegel, C., Valeur, F. and Vigna, G.: *INTRUSION DETECTION AND CORRELATION*, Springer (2004).

[3] Symantec Corp.: *Symantec Network Security 7100 Series*. http://www.symantec.com/.

[4] Wang, K. and Stolfo, S. J.: Anomalous Payload-Based Network Intrusion Detection, *Proceedings of the Seventh International Symposium on Recent Advances in Intrusion Detection*, pp. 203–222 (2004).

[5] 山田　明，三宅　優，竹森敬祐，田中俊昭：学習データを自動生成する未知攻撃検知システム，情報処理学会論文誌，Vol. 46, No. 8, pp. 1947–1957 (2005).