マルチメディア通信と分散処理　106－25
コンピュータセキュリティ　16－25
（２００２．２．１５）

ニューラルネットワークを用いた学習型ＮＩＤＳの開発

安藤類央　武藤佳恭
慶應義塾大学　政策メディア研究科
〒252-0816 神奈川県藤沢市遠藤５３２２
TEL 0466-49-1077
E-mail: ruo@sfc.keio.ac.jp / takefuji@sfc.keio.ac.jp

あらまし

ウィルス被害の増大やセキュリティ侵害の事例の多様化によって、ネットワーク管理者の負担が増大している。通常管理者は適宜ログ情報を閲覧するか、IDS(Intrusion Detection System)を用いてネットワーク上を監視する。しかし、日々の管理運営上で発生する全ログ数は非常に大量に記録される場合が多く、ＩＤＳのシグニチャやデータベースは急速に増加している傾向にある。それに加えて、新種のウィルスや不正アクセス手法も登場している。そのため、完全に未知ではないが新しい傾向を持つ不正侵入に関して的確かつ迅速に対処する必要がある。また、最近では監視サービスを提供する場合でも、当該サービスが旨とするネットワークの監視と侵害（要因）除去に関しては、より高次のリアルタイム性と即効性が求められている。本論文では、このようなボトルネックを解消するためにネットワークの状態を視覚的にキャプチャしてリアルタイムでニューラルネットワークに認識学習させる研究の概要と中間報告を記載する。

キーワード 検知作業、ＮＩＤＳ、リアルタイム認識学習

Real-time neural detection with network capturing

Ruo Ando / Yoshiyasu Takefuji
Keio University Shonan Fujisawa Campus
5322,Endo,Fujisawa-city,Kanagawa 252-0816,Japan
E-mail : ruo@sfc.keio.ac.jp / takefuji@sfc.keio.ac.jp

Abstract

The rapid increase of diversity of computer virus and threats to security becomes a great burden to system administrator. Although ideally administrators should investigate all log information or perform intrusion detection by using IDS in terms of security of their network, it is expected that almost all of them cannot complete these tasks.

We assume that it is necessary to capture the network status with providing visual effects. By the intrusion detection system we are developing, administrator will become able to manage security threats with less burden and tensions. Besides, we will give some samples of real-time neural detection to react to the emergence of computer virus or unauthorized access with the new trend more effectively.

Key Words Intrusion Detection, NIDS, Real-time neural detection

１はじめに

本論文では、増加する不正アクセスやウィルスに対処する管理者の負担を軽減させるために、リアルタイムにネットワークをキャプチャしてニューラルネットワークによる認識学習を行う手法を採用した。理由は以下の２点である。

１）現在ネットワークの管理運営上で発生するログはすべて文字と数字の情報の時系列に沿ったリスティングで構成されている。管理者は通常、ネットワークの障害除去も含めて、ログの閲覧を行う。この際、記録された情報すべてを把握した上で、短期間のうちに障害を察知し、原因を特定することになるが、この作業の完遂は熟練したアドミニストレータでも困難なことがある。とりわけ中規模以上のネットワークになると、ログのチェックに関してはその記載量が膨大な量にのぼるため、すべてのログを

いちいち見ていられないというケースがあり得る。

2）また、一定の規則に基づきログ情報を利用して、セキュリティ上のアラートを検知するＩＤＳ（Intrusion Detection System 侵入検知システム）は、概して監視対象のアクセスに関してパターンマッチを行うという方法をとっている。このＩＤＳを使用する際には、新種のウィルスや、誤警報とのトレーオフが生じるＤｏＳ(Denial of Service：サービス不能攻撃)，あるいは完全に未知ではないが新しい傾向を持つ不正アクセスへの対処に関する即効性とそれに付随するコストが不可避の問題になっている。そのため、セキュリティ侵害に関してパターンマッチと比較して緩やかに過去のログを元にして不正を検知するシステムの開発は管理者の負担を軽減すると思われる。

端的に言えば、ＩＤＳはセキュリティ侵害に対する支援を行うツールあり、最近は相関分析などの統計的手法を用いたシステムが公開されているが、監視対象のデータを見た場合、データの採取の仕方と表現にもよるが、異常と正常との間でデータを切り分ける際に、機械的にリニアな分離や明瞭な判別が可能であるケースは少なく、多くの場合種々のトレードオフが発生するので、不正を判断する場合、ネットワーク上を流れるパケットの性質を多角的に判断しなければならない状況に直面することがある。また、統計的手法ではあらかじめクラスタ数などの与件となるパラメータを決定しなければならなく、ネットワーク環境の変化、新たな不正アクセスの出現に応じて管理者がこのような設定を適宜行わなければならなくなる。ＳＯＭ(Self –Organization Model)を用いることによりこのあたりの作業をある程度オートマティックに行うことが可能になると予想される。

管理対象のネットワーク上に障害が生じた際に、ファーストサーチでその原因を特定できることは少ない。そのため、アドミニストレータは特定のマシンのデータから出発して芋づる式に障害の原因をたどっていくことになる。しかし、どの場所に障害が発生するかということと、障害を発生させているセグメント上のマシンの位置特定とのシーケンスの確定には、推定できる範囲ではいくつかの選択肢が存在し、決定と対応には意外と時間がかかることがある。換言すれば、現況のネットワーク管理形態では障害報告が原因となるネットワークトポロジ上のマシンの位置特定にダイレクトに反映されないことが問題となるケースが存在する。また、ウィルス感染などの場合、最近のものは感染したマシンのメーラのアドレス帳などを利用して爆発的に増殖するため、当該ネットワークのあるクライアントに感染したウィルスの被害の防止を最低限に抑えるためにはなによりも即効性が求められる。新種のウィルスが出現した場合、文字と数字情報のログを定期的にチェックしてから、どのマシンが感染源なのか特定するという手法よりも、問題となるコンピュータの検索に関してより効率的な手法が開発できると考えられる。

以上を論拠にして、本論文では特定期間のネットワークの状態をキャプチャし、いくつかの不正アクセス手法に関してはニューラルネットワークに認識学習させ、迅速な対応を支援することに成功した。

2 セキュリティ侵害と検知作業について

インターネットの普及の弊害として、コネクティビティの提供されている計算機へのセキュリティ侵害が急増している。インターネットはそもそも相互に信頼のおける通信主体が自由にアクセスし、情報をシェアして価値を高めていくという理念から起因しており、基本的に開放性や柔軟性を志向したものであると言える。その一方で、インターネットの発展によって我々が享受する恩恵と併行して、ここ数年政府や企業に本格的に導入され重要度の高い情報とサービスが扱われるようになった結果、安全性の追求の後発性をついたセキュリティ侵害の事例が多発している。管轄下にあるネットワークへのセキュリティ侵害には次の３つのタイプが挙げられる。

1）侵入：特定のマシンやシステムに本来認可されていない権限でアクセスする。また、通常の利用形態のみ許されたユーザの権限を、管理者などがもつ特権を所持するように工作すること。
2）サービス妨害：インターネットはコネクティビティの提供されているマシンに幅広いサービスが提供されるように設計されているが、これを逆用して意図せざる入力などをすることによってリソースを浪費させる、またトラフィックを大量に発生させてネットワークやシステムに負荷をかけ、レスポンスの低下、サービスの停止などを引き起こすことなどを言う。
3）情報の改竄と盗用：システムへの侵入を行った攻撃者が、重要な情報が格納されたファイルを不当に参照、削除、あるいは改竄する。また、ファイルを検出が困難な形で変更しておくことで、システムの機能をクラッカに都合のよいように改変すること。あるいは、侵入に成功したマシンに特定のプログラムをインストールすることで、ネットワーク上のクリティカルな情報を不正に入手するなどの行為を行うこと。

上のようなセキュリティ侵害への対応に関しては、受動的対応と能動的対応の２つのものに分類することができる。前者は暗号、認証技術などがその代表としてあげられる。これは不正アクセスなどについて個々のパーツレベルのプロテクトを高めることを旨とする。後者は積極的に侵害や攻撃を検出してクライアントや管理者による対応を可能にする術を指し、ここではログ閲覧とＩＤＳの構築運営を検討する。

これらの積極的対応策の一つとして検知作業があるが、これは管理対象ネットワークを構成するクライアントとサーバが記録する情報から、不正と見なされる情報を検出し、かつシステム管理者に迅速に通知する作業である。当該セグメントへのセキュリティ侵害は各マシンのログのうちに何らかの形跡を残すのが普通である。そのため、セキュリティ侵害を除去あるいは予防するためにはアドミニストレータが適宜ログをチェックする必要がある。組織の目的や状況によってネットワーク環境は多様化するが、管理者が定期的にログを閲覧する作業が必須なのに変わりはない。

ログの閲覧作業はセキュリティ侵害において欠かすことのできないものであるが、その一方でこの作業を通常の業務に加えて日々完遂している管理者は少ないと思われる。その理由として、中規模以上のネットワークになると閲覧対象となるログの情報は膨大な数になり、記載されている情報すべてを把握した上で総合的な判断を下すという一連の作業はかなりの負担になるということが挙げられるだろう。

今回は、ネットワークの状態をキャプチャしてニューラルネットワークに認識学習させる攻撃手法として、DoS(Denial of Service)をチョイスした。その理由は以下の３点である。

1）ＯＳのバグをつくセキュリティ侵害などと比較して、ネットワーク上で攻撃が展開されるとき、パケットのフロー、あるいは各パケットの特徴が謙虚に出るため、ＤＯＳ攻撃は視覚的に把握しやすい傾向がある。

2）ＤＯＳに関しては、侵害と判断する基準をきつくすると、正常なアクセスなのにハングアップする、誤警報が生じてしまう確率が高くなるなどの、トレードオフが生じる。そのため、ログベースのパターンマッチというよりは、ログを集計した結果を閲覧した上で定量的に判断を下す必要がある。

3）ＤＯＳ攻撃の非対称性１
ＤＯＳ攻撃は当分の間なくならないということが予想される。なぜなら、攻撃者の技術力とその効果に非対称性が存在する、端的にいえば簡単に行うことができるからである。不正な特権の昇進、アクセス権の奪取などにくらべるとＤＯＳ攻撃は比較的容易に行うことができる。実際に、現在出回っているＤＯＳ実行用のツールを用いてば、単純なインターフェイス操作で強力なＤＯＳ攻撃をターゲットとするマシンに行うことができる。

4）ＤＯＳ攻撃の非対称性２

DOSは非対称に行える。クラッカは攻撃に使うリソースとは不釣合いな対象マシンの資源を消費し尽くすことができる。つまり、わずか数十行のコードで多少マシンのネットワークやシステムの運用を妨害することが可能である。

5）攻撃対象となる条件の低さ

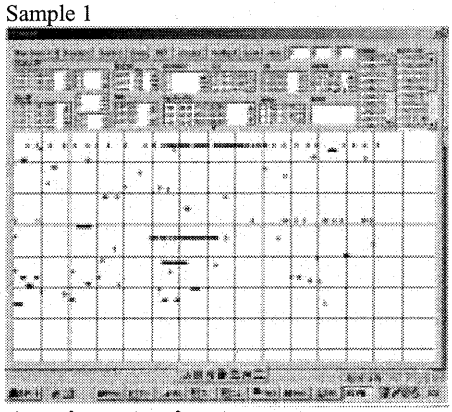
対象となるコンピュータにおいて有限の資源を消費可能であれば、いかなる場合にもDOS攻撃が可能である。広くいえば、TCPIP接続の本質的性質から、インターネットへのコネクティビティが提供されているマシンならばすべてこの条件下に入る。特に、資源への割り当て制限なしに資源を費やされてしまうシステムであればDOSによって多大な損害を受ける。

ＴＣＰ/ＩＰは、インターネットを支える多くの技術と同様に、相互に信用のおける関係のあるオープンなコミュニティで構築されてきたものである。そのため、比較的各ユーザが性善説に基づく行動をとるという前提によって成立している傾向が強く、従来セキュリティ面はあまり考慮されていなかったと言ってよい。ネットワークにアクセスする主体はみな善人で、お互いに信頼関係にあるという理念が、インターネットの発展に大きく寄与したことは間違いないだろう。しかし、一方で多くのＯＳやデバイスは上記の理念や開発上の都合から程度の差こそあれ、多少なりの欠陥を孕んでいるとも見ることができる。また、インターネットは一度コネクティビティが提供されれば、ユーザ間で情報資源をシェアして価値を高めていくというすばらしい利点を持っているが、一方では、これは悪意のあるユーザがリソースを容易に消費できてしまうことを意味する。ＤＯＳ攻撃はこのような弊害をつくことで技術的に洗練されていない者でも容易に実行が可能である。

３ネットワーク情報のビジュアライズについて

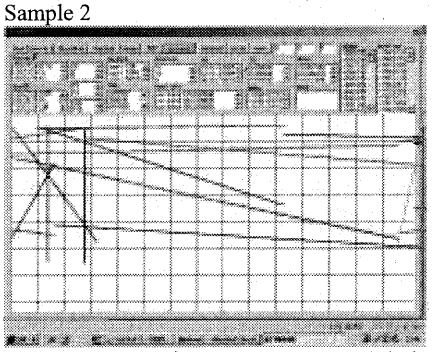
上述したように、ネットワークの状態を文字と数字情報からなるログ情報から把握することは、相当に熟練した管理者でもかなりの負担が伴う。中規模以上のネットワークになると、一日に発生する基幹サーバ系のログだけでも大量に発生することが普通であり、アドミニストレータは通常業務に加えてこれらのトラックを短いスパンで閲覧把握することは事実上不可能であるといえる。しかし、個別かつ膨大に生じるログを特定の視覚情報として表現できれば、情報の把握効率という意味で状況は一変すると思われる。以下は、本研究でこのような目的を達成するために作成したサンプルである。

Sample 1



サンプル１は、ネットワーク上に発生するパケットをリアルタイムにプロットしていくものである。

Sample 2



サンプル２は、ネットワーク上に発生するパケットのフローを描画したものである。

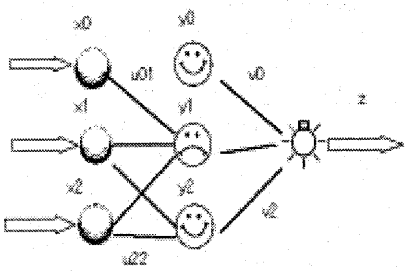
4パーセプトロンによるネットワークのリアルタイム認識学習について

比較的定型的な業務を行っているセグメントをキャプチャしてみても、ネットワークの状態は刻々と変化すると思われる。特に不特定多数からのアクセスを招くネットワークになると、利用形態、ユーティライゼーション、アクセス頻度どれをとってみても動向はかなり激しくなると予想される。このような状況では、どのようなスパンで見れば、ネットワークをキャプチャしたものが再現性を帯びてくるか確定することは難しい。簡単に言うと、ネットワークの状態の変貌が激しくて捉えどころが見つかりにくいケースが多いということである。

通常のログ閲覧とＩＤＳなどによる検知作業においては、管理者の頭の中には正常な基幹サーバ、クライアント、セグメントの状態がある程度記憶されている。それに即して管理者はある時間帯のネットワークのログやＩＤＳのレスポンスの集合をまとめてシステム監査についての判断を下すわけであるが、ネットワークの技術は日進月歩で発展を続けており、日常的にサービスを享受するユーザも数も増えつづけている。加えて、あらたな性質のログが追加されたり、ログが表現するサービスの利用のトレンドも頻繁に変化していくものと思われる。

このように、例えば極端に言えば、１年前の状態を踏まえて、現在の管理下にあるネットワークの状態についてなんらかの判断を下すのは無理というものである。そのため、管理者は意識的にせよ、無意識にせよ、常にネットワーク上の各マシンで発生するログを把握し、異常や不正の検知のためのネットワークのビジョンを更新しつづけていることになる。前述した管理者によるログ閲覧の情報認識負荷の高さも踏まえると、管理者の監視下のセグメントに関するビジョンの定期的かつ頻繁な更新という作業をニューラルに代行させ、かつアドミニストレータよりも迅速にかぎりなくリアルタイムに近く行うことができれば負担大幅に軽減させることができると想定できる。

以下では、本研究で利用したパーセプトロンモデルを用いたバックプロパゲーションについて簡単に紹介する。



バックプロパゲーション（逆誤差伝達法）とは、任意のベクトルを入力層にインプットし、中間層を経て出力層の各パーセプトロンの値を計算し、教師信号（正解）との誤差から各層間の結合係数を変更していくという学習モデルである。

ニューラルネットワークには、大きくわけて１最適化（アプリオリにビルトインされたもの）、２教師付き学習モデル、３教師なし学習モデル（自己組織化）の３つがあるが、今回は２（教師信号によってモデルの構成を変更させ学習を行っていくもの）を採用した。端的に言ってしまうと、学習のポイントは、誤差からいかにモデルの構造（各層の結合係数）を変えるかということであり、出力層から入力層へのフィードバックは以下のように行う。

$$\delta_k = (\hat{z}_k - z_k)\, z_k (1 - z_k])$$

z_k：出力層出力　$\hat{z}_k$：教師信号

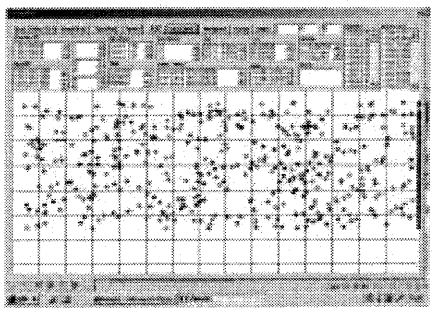
出力層から中間層へのフィードバック

$$v_j = v_j + \eta \sum_{k=1}^{k} \delta_k y_{i,k}$$

中間層から入力層へのフィードバック。中間層に対するデルタ信号を計算しなおす必要がある。

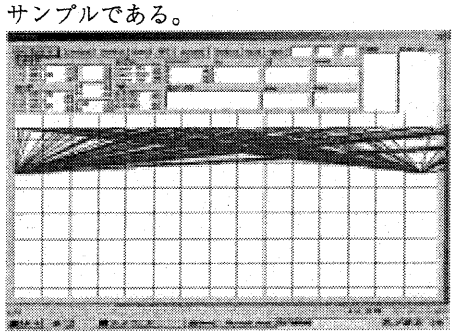
$$u_{i,j} = u_{ij} + \eta \sum_{k=1}^{k} \sigma_k x_{j,k}$$

下は検知対象として選んだSYN FLOODのキャプチャ図の一例である。



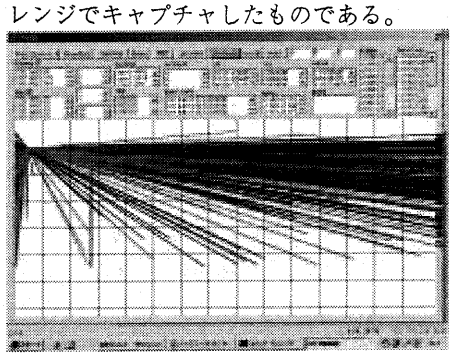
ネットワークをキャプチャした後、プロトタイプに組み込んだニューロによって認識率を上げることに成功した。

次はセキュリティ侵害に係るアクションとしての外部へのスキャンをキャプチャしたサンプルである。



6 補足・今後の課題

5で検討したＳＹＮ　ＦＬＯＯＤを別のアレンジでキャプチャしたものである。



これに対して特定のニューラルネットワークの処理を行うと、ネットワーク上のどのマシンが問題となっているのか人間より早く検索できるようなると想定している。

7 終わりに

以上、ネットワークの情報をビジュアライズし、ニューラルネットワークによって認識学習させるプロトタイプの開発の背景と報告を行った。ネットワークの管理運営上でトラッキングされるログは数字と文字情報から構成されており、日々保守業務に忙殺される管理者にとって、検知作業時のログ閲覧作業にかかる情報把握負荷の高さは無視できないものである。

一方で、互いに信頼関係にあるもの同士がコネクティビティを提供しあい、情報をシェアして高めるというインターネットの理念は、開放性と柔軟性からなっているが、これを欠陥として構築した攻撃手法によるセキュリティ侵害の事例が多発している。

本稿では、このような背景を踏まえて、検知作業をネットワークキャプチャリングした上、ニューラルネットワークに認識学習させることで管理者の負担を軽減し、特定の攻撃への検知作業を支援するためのプロトタイプについて議論した。

参考文献

[1]武田圭史、磯崎宏,"ネットワーク侵入検知",ソフトバンクパブリッシング,June 9, 2000
[2]武藤佳恭、斎藤孝之監修　武藤佳恭研究室編,"応用事例ハンドブック　ニューラルコンピューティング",共立出版,March 30 2001
[3]W.Richard Stevens,"UNIX NETWORK PROGRAMMING",Prentice Hall,Inc,1990
[4] 高田 哲司, 小池 英樹 ,"ログ情報視覚化システムを用いた集団監視による不正侵入対策手法の提案 ".情報処理学会論文誌 Vol.41, No.8, pp.2216-2227, (2000).