様式 C-19

科学研究費補助金研究成果報告書

平成２２年　６月　８日現在

研究種目：若手研究（B）
研究期間：2007～2009
課題番号：１９７００７２
研究課題名（和文）　免疫系に学んだネットワーク異常検出システムの研究

研究課題名（英文）　A study of an immunity-based anomaly detection system

研究代表者
岡本　剛（OKAMOTO TAKESHI）
神奈川工科大学・情報学部・准教授
研究者番号：９０３５０６７８

研究成果の概要（和文）：本研究は、免疫系を参考にして、コンピュータウイルスやネットワーク攻撃などの不正行為を検出する免疫型ネットワーク異常検出システムの理論的枠組みを構築した。シミュレーションでは、ユーザの振る舞いを学習する動的プロファイル更新機能を導入することにより、他人になりすましたユーザの検出精度を、従来と比べて、内部ユーザのなりすましでは、49.72%、外部ユーザのなりすましでは、12.16%改善した。さらに、新種のコンピュータウイルスも検出できることを確認した。

研究成果の概要（英文）：This study proposed a framework of an immunity-based anomaly detection system for detecting computer viruses and network attacks. Simulation results showed that dynamic updating of user profiles improves the detection accuracy of internal masqueraders by 49.72 percent compared to our previous system, and that of external masqueraders by 12.16 percent. In addition, the proposed system could detect some of new computer viruses.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２００７年度	900,000	0	900,000
２００８年度	700,000	210,000	910,000
２００９年度	700,000	210,000	910,000
年度			
年度			
総　計	2300,000	420,000	2,720,000

研究分野：　総合領域
科研費の分科・細目：　情報学　・　計算機システム・ネットワーク
キーワード：ネットワークセキュリティ、異常検出、コンピュータウイルス、なりすまし、免疫系、エージェント、プロファイル、適応

１．研究開始当初の背景

(1) コンピュータウイルスやネットワーク攻撃の検出は、過去に発見された手口を手掛かりにしているため、新種のコンピュータウイルスやネットワーク攻撃の検出は困難であった。

(2) これまで、研究代表者は、生物の免疫系を参考にしたセキュリティシステムを研究してきた。特に、免疫系の多様性を用いた識別戦略を、他人になりすましたユーザ、すなわち、「なりすましユーザ」の検出に応用した研究では、従来手法より優れた検出精度を得た。この研究では、UNIX 互換 OS のコマンド履歴を分析して、ホスト上でのなりすましユーザの検出を対象にしていた。

２．研究の目的

(1) 本研究課題では、ユーザの振る舞いについて、分析の対象をネットワークトラフィックに変更し、既知・未知に関わらず、コンピュータウイルスやネットワーク攻撃を検出できる免疫型ネットワーク異常検出システムの開発を目的とする。なお、想定するシステムは、外部ネットワークからの攻撃ではなく、ネットワーク内部からの攻撃（例えば、DoS 攻撃やコンピュータウイルスの増殖など）を対象とする。

(2) 本研究課題を実現するために、具体的には、以下の目標を設定した。

① 生物の免疫系を参考にして、コンピュータウイルスやネットワーク攻撃を検出する免疫型ネットワーク異常検出システムの理論的枠組みを構築する。

② これまでの研究により、エージェントの多様化が検出精度を改善する可能性があることを確認している。検出精度をさらに改善するために、エージェントを多様化するアルゴリズムを考案する。

③ ネットワーク上でのユーザの振る舞いは時間とともに変化することが予想される。その変化によって検出精度が低下しないようにするために、ユーザの振る舞いに適応するアルゴリズムを考案する。

３．研究の方法

(1) 研究代表者がこれまでに提案したシステムを応用にして、ネットワーク上のユーザの振る舞いを監視し、その異常を検出する理論的枠組を検討する。具体的には、監視の対象を「コマンド履歴」ではなく、「IP アドレスの通信履歴」とする。

(2) 提案システムのイメージを図 1 に示す。スマイルマークのアイコンは、エージェントを表し、各エージェントは、ユーザを特異的に認識する。これはユーザを識別するためのプロファイルを利用することによる。プロファイルとは、ユーザの特徴を表すデータであり、本研究では、学習アルゴリズムのパラメータに相当する。

(3) 各エージェントはネットワーク経由で他のコンピュータへ移動できると仮定する。さらに、エージェントは、各自のプロファイルを変化させることにより、新しいエージェントを生成できると仮定する。つまり、1 台のコンピュータには、複数のエージェントが存在し、それらがユーザの振る舞いを監視する。

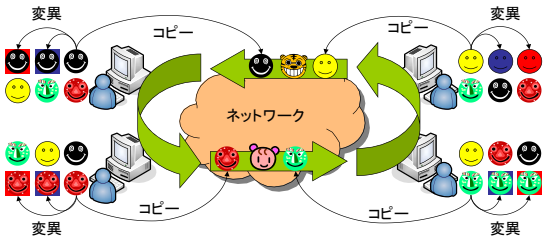


図１　提案システムの概念図

(4) 各エージェントは、ユーザの振る舞いを監視し、一定間隔で、その振る舞いを点数化する。通常、正当なユーザの振る舞いに対して、そのユーザのプロファイルを有するエージェントは高い点数を、それ以外のエージェントは低い点数を提示する。つまり、なりすましユーザの振る舞いに対して、エージェントは低い点数を提示する。各エージェントが互いに提示した点数の相対的な関係に基づいて、提案システムは、ユーザの振るまいが正常な振る舞いであるかどうかを判断する。

(5) これまでの研究で、エージェントの多様性が検出精度と関係していることがわかっている。エージェントの多様性を確保するため、抗体の生成や交配のメカニズムを参考にして、新しいエージェントを生成する。すなわち、エージェントが所有するユーザのプロファイルを自動的に生成するアルゴリズムを考案する。

(6) ネットワーク上のユーザの振る舞いは時間とともに変化することが予想される。ユーザの振る舞いに適応できるように、正当な

ユーザ本人であると認識した通信履歴をエージェントに学習させるタイミングやその方法を考案する。

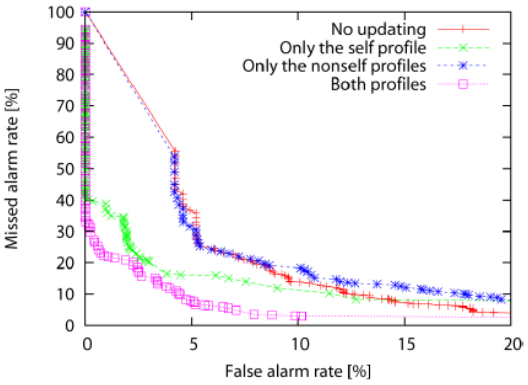
４．研究成果

(1) 研究代表者がこれまでに提案したシステムを応用にして、ネットワーク上のユーザの振る舞いを監視し、その異常を検出する理論的枠組、すなわち、免疫型ネットワーク異常検出システムのフレームワークを構築した。

(2) エージェントが所有するプロファイルの作成方法について検討した。プロファイルは、各ユーザの通信履歴から作成することになるが、IP アドレス空間は膨大な空間であるため、一定期間にユーザが通信した IP アドレスをラベリングすることにより、IP アドレス空間を縮小した。空間の大きさは、経験的に、数十から数百程度の大きさが望ましい。具体的には、最近１ヶ月間に通信したコンピュータの IP アドレスのリストを事前に保持し、IP アドレスをリスト内の通し番号に変換する方法を適用した。IP アドレスがリストになければ、リスト末尾の通し番号＋1 の値に変換する。

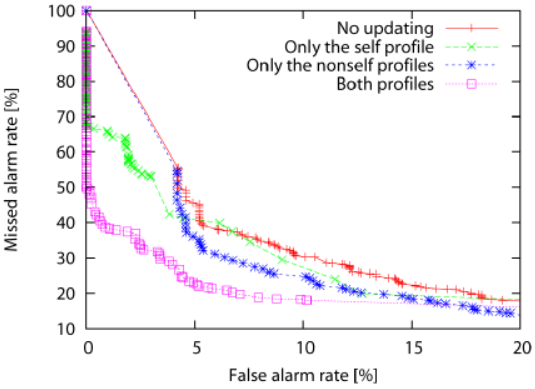
(3) これまでの研究では、隠れマルコフモデルによりプロファイルを作成してきたが、本研究では、これ以外にマルコフ連鎖、ナイーブベイズ分類、ニューラルネットワーク、サポートベクタマシンなどによりプロファイルを作成し、これらの検出精度を調べた。これらを評価した結果、隠れマルコフモデルが最も優れた検出精度を示した。

(4)免疫型ネットワーク異常検出システムのエージェントに適応アルゴリズムを組み込み、その検出精度を評価した。評価には、受信者動作特性曲線（ROC 曲線）の内部面積（AUC）の大きさを用いた。評価した結果、適応アルゴリズムを組み込む前と比べて、内部ユーザのなりすましは、49.72%（図 1）、外部ユーザのなりすましは、12.16%（図 2）改善した。各図は ROC 曲線であり、横軸が誤報率、縦軸が欠報率を表す。赤線は、適応アルゴリズムを導入する前の結果、緑線は、本人のプロファイルのみを適応させた結果、青線は、本人以外のプロファイルのみを適応させた結果、ピンク線は、両方のプロファイルを適応させた結果である。



(a) Internal masquerader detection

図１　内部のなりすましユーザの検出



(b) External masquerader detection

図２　外部のなりすましユーザの検出

(5) なお、上述の検出精度の計算には膨大な組み合わせがあるため、計算速度を改善する必要があった。そこで、研究費補助金で購入した２ウェイの４コアプロセッサーや６コアプロセッサーの性能を最大限に発揮できるように、並列計算を可能にするプログラムを作成した。その結果、検出精度の評価に要する時間を約 8.3～12.5%の時間に短縮した。

(6) コンピュータウイルスの検出では、この適応機能を導入しても検出精度に変化はなかった。すなわち、ランダムに増殖する Slammer などのコンピュータウイルスを、誤報率 0.0%、欠報率 0.0%で検出した。

(7) さらに、インターネットで流行している最新のコンピュータウイルスに対する検出精度を評価するために、高対話型ハニーポットと低対話型ハニーポットを構築した。なお、研究代表者が所属する大学のネットワークにはコンピュータウイルスが流行していないため、商用のインターネットサービスプロバイダーの無線接続サービスを契約した。無線接続サービスでは、NAT ルータなどの中継装置を使わずに、パソコンをインターネットに直接接続する環境が主流であるため、コン

ピュータウイルスが蔓延している。事実、1台のコンピュータで毎分1回ぐらいの割合でコンピュータウイルスと疑われる攻撃を確認し、約1年間に2,198個のコンピュータウイルスを収集した。

(8) ハニーポットにより収集したコンピュータウイルスの中で代表的なコンピュータウイルスについて、適応機能を有する提案システムを評価した結果、他のコンピュータウイルスと同様に検出できることを確認した。

(9) エージェントの多様性を確保するため、エージェント自身がプロファイルを組み替えて、新しいエージェントを生成する機能を検討した。これまでの研究で、ランダムな組み替えは、あらゆるユーザを認識しないという結果を得ていた。そこで、本研究では、免疫系の胸腺での選択機構を参考にして、多数のエージェントを生成し、エージェントの多様性を向上させるアルゴリズムを検討した。さらに、性システムの中心的役割を果たす交配のメカニズムを参考にするアルゴリズムも検討した。その結果、いずれのアルゴリズムにおいても、検出精度を改善できなかった。

(10) プロファイルの組み替えが検出精度を改善できない原因を明らかにするために、上述(9)のアルゴリズム以外に、ヒューリスティックにプロファイルを生成する方法を検討した。その結果、プロファイルを構成する情報空間が巨大であること、さらに、既存のユーザのプロファイルから未知のユーザの特徴を予測することが困難（見ず知らずの他人のお気に入りのホームページを特定することが困難）であることが原因で検出精度を改善できないことがわかった。後者の原因は、ユーザの通信内容を分析すれば可能になると予想されるが、プライバシーの問題がある。この問題を解決できれば、検出精度をさらに改善できると考えられる。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計２件）

① Takeshi Okamoto, Yoshiteru Ishida, Towards an Immunity-Based System for Detecting Masqueraders, 査読有, International Journal of Knowledge-based and Intelligent Engineering Systems, Vol. 13, No. 3, 2009, pp. 103 - 110.

② Takeshi Okamoto, Yoshiteru Ishida, An Immunity-Based Anomaly Detection System with Sensor Agents, Sensors, 査読有, 9(11), 2009, pp. 9175-9195.

〔学会発表〕（計３件）

① Takeshi Okamoto, Yoshiteru Ishida, Evaluations for an Immunity-Based Anomaly Detection with Dynamic Updating of Profiles, 査読有, Proc. of AROB 2010, 2010, pp. 58-61.

② Takeshi Okamoto, Yoshiteru Ishida, Dynamic Updating of Profiles for an Immunity-Based Anomaly Detection System, 査読有, Proc. of KES 2008, LNAI 5179, 2008, pp. 456-464.

③ Takeshi Okamoto, Yoshiteru Ishida, Framework of an Immunity-Based Anomaly Detection System for User Behavior, 査読有, Proc. Of KES 2007, LNAI 4694, 2007, pp. 821-829.

６．研究組織
(1)研究代表者

岡本　剛（OKAMOTO　TAKESHI）
神奈川工科大学・情報学部・准教授
研究者番号：９０３５０６７８