様式 C-19

# 科学研究費補助金研究成果報告書

平成２３年　５月１９日現在

機関番号：１１３０１
研究種目：　若手研究(B)
研究期間：　2009～2010
課題番号：　21700066
研究課題名（和文）
オーバーレイネットワークを利用したインターネットにおける情報の拡散経路推定方式
研究課題名（英文）
Estimating Diffusion Path of Information using Overlay Network
研究代表者
和泉　勇治（WAIZUMI YUJI）
東北大学・大学院情報科学研究科・准教授
研究者番号：90333872

研究成果の概要（和文）：
インターネットを流通する情報の同一性を判断するフローの識別子と同一性評価アルゴリズムを提案し、ある情報がインターネットを移動した経路を推定する方式を開発した。フローの識別子としては、フローに含まれる 8 ビットコードのヒストグラムとメッセージの送受信のタイミングに基づいた識別子を開発した。これらの識別子を利用し、ネットワークで観測された情報が同一のものであるか否かを、プライバシー性を保持しながら判断するアルゴリズムを完成した。

研究成果の概要（英文）：
An estimation method of diffusion path of information on the Internet has been developed. The method consists of extracting module of a flow identifier and identifying module. As flow identifier, 8 bit code of flow payloads and time interval for transmitting messages are proposed. Using these identifiers, an algorithm to detect identity of flows observed in network is developed keeping to privacy.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合　計 |
|---|---|---|---|
| 2009 年度 | 1,600,000 | 480,000 | 2,080,000 |
| 2010 年度 | 1,400,000 | 420,000 | 1,820,000 |
| 年度 | | | |
| 年度 | | | |
| 年度 | | | |
| 総　計 | 3,000,000 | 900,000 | 3,900,000 |

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：情報システム、セキュア・ネットワーク

## １．研究開始当初の背景

近年、インターネットは電子政府などをはじめとする多くの社会・経済基盤に利用され、その役割の重要性は高まる一方である。これに伴い、インターネットには健全で堅固であることが強く要望されている。誰もが情報通信技術の利便性を享受できる安全・安心な情報化社会の発展のためには、インターネットの安全な運用技術の確立は世界的な最重要課題である。最近では、不正アクセスやそれに感染した P2P ソフトウエアにより、情報の所有者の意図しない形でその情報のインターネットへの流出や拡散が生じ、大きな社会問題となっている。

この問題へ対処するために、インターネットワームや P2P トラヒックを検知し遮断する必要があるが、それらを 100%正しく検知することは非現実的である。そこで、検知漏

れとなった通信により送信された情報が、インターネット上をどのように伝搬し拡散して行ったか、また、どこから発生したのか明らかとする方式を開発することにより、インターネットワークの発生の抑止や情報流出などの被害拡大を防ぐことが出来ると考えられる。

２．研究の目的

インターネット上で送受信される情報の伝搬・拡散経路を推定するためには、ネットワーク上で観測された情報が、経路を推定すべき追跡対象の情報であるのか否かを高い精度で判断する必要がある。それ実現するために、あるネットワークにおいてフローの流入と流出を識別するフローの同一性評価を実現するための数値化方式手法と同一性評価アルゴリズムを開発することが研究の目的となる。数値化方式に関しては、利用者のプライバシーを考慮し、通信データ中に含まれる文字列などを直接利用せずに同一の不正アクセスを同定するためにはどのような情報を抽出すべきかについて明らかにする。

移動経路推定アルゴリズムについては、オーバーレイネットワークを想定した交換すべきフローの情報の決定と、その情報に基づく追跡対象か否かの判断アルゴリズムの構築が研究の目的となる。

３．研究の方法

（１）同一フローの流入流出識別方式の開発

この識別方式は、あるノードのワームが流入し感染した後、他のノードへ感染するためにそのノードから流出する通信の判別や、P2P ソフトウエアが他のノードへ転送を行っていることを把握するための方式の開発となる。

研究代表者は、既にインターネットワームの検知のためのフロー識別方式や、P2P 通信の識別方式を開発している。既に提案しているワームフロー識別方式はフローの一致性を評価するもので、同じフローが多数発生している場合にワーム拡散が起こっていると判断するものであるが、ワームがあるノードへ感染し拡散する際の変化を許容した同一性を評価する指標としては設計されていない。また、P2P 通信の識別方式も、他のアプリケーションと P2P ソフトウエアによる通信違いを評価可能であるが、P2P ソフトウエアにより通信される個々のファイルの識別は出来ていない。これらの問題を解決するために、個々のフローに特徴的なデータ列の発見方法と、その数値化方式について検討を行う。これらの方式は、申請者が研究してきたワーム識別方式、P2P トラヒック検知方式の詳細化との解釈が出来る。つまり、同じ種類のワームであっても異なるノードから送信されたフローは別なものとして識別し、P2P ソフトウエアによる通信についても同様に、ファイル毎の違いを識別しそのネットワーク上の移動を捕捉する手法を検討することになる。

（２）情報の移動を表現するグラフの構築方式

情報移動を表すグラフの構築の際に考慮すべき問題は、追跡対象フローの同一性の判定基準、フローの観測時刻の利用方法、誤ったグラフ生成リスクの低減となる。

本研究が生成するグラフは、グラフの点をネットワーク上のノード、辺を同一情報の移動経路を表すものである。グラフのある点に流入する辺と流出する辺を作成することは、ネットワーク上のあるノードで情報が受信され、それが別のノードへ送信されたことを表す。この同一の情報の移動を捕捉するためには、ネットワーク上のノードへの流入フローと流出フローの同一性を適切に判断することが必要となる。フローの同一性の判定基準には、（１）で開発したフローの流入流出識別方式を利用可能であるが、あるネットワーク上のノードに着目した場合、一定時間内に多数のフローが出入りするため、それらの複数のフローから適切に同一性を判断する必要がある。この問題に対しては、フローの送受信 IP アドレスのマッチングを用いた同一性判断の対象となるフローの絞り込みや、観測時刻による流入、流出の順序の評価を併用し適切な同一性判断の基準を検討する。

４．研究成果

（１）フローの同一性判断用識別子の開発

ネットワークを流入流出するフローの同一性を定量的に判断するための数値化方式としてフローを構成するパケットのペイロード部の８ビットコードのヒストグラムを利用した識別子とパケットやメッセージの送受信タイミングをベクトル化した識別子を提案した。この二つの識別子は共にベクトルとして表されており、これらのベクトル間距離を利用することで、フローの同一性を評価可能であることが判明した。しかし、追跡対象の情報に類似したフローによるエラーが発生することも明らかとなった。

（２）尖度を利用した追跡情報受信ホストの絞り込み

（１）で開発したベクトル間の距離を評価することにより、追跡対象の情報があるホストに到達しているか否かを判断出来ることが明らかとなった。しかし、ベクトル間の距離を利用した経路推定では、追跡対象の情報が到達したと判断されたホスト群の中に、誤って到達したと判断されたホストも含まれてしまうことも明らかとなって。

このエラーを解決するために、ベクトルを

構成する要素の平均と分散を計算し、それらから尖度を算出する。同一の情報を転送するフローの尖度は類似した値をとることが明らかとなった。図１には、不正アクセスの一種であるワームの種別毎の尖度を示している。

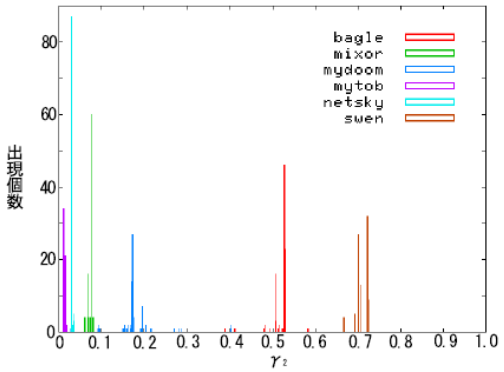


図１：ワーム種別毎の尖度の分布

図１から、ワームの種類が異なると尖度の値もことなり、それぞれの分布の重なりも非常に小さくなり得ることが分かる。

この特性を利用し、図２に示す同一性判断の基準となるベクトル間距離に対する閾値の自動設定アルゴリズムを構築した。

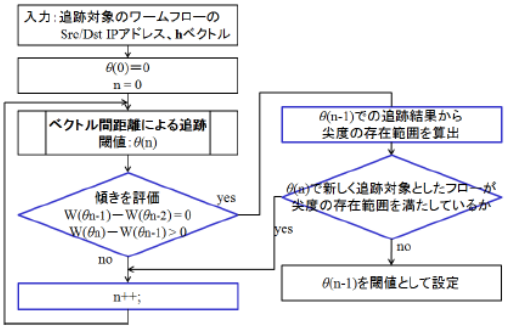


図２：ベクトル間距離の閾値決定アルゴリズム

図２中のWは、経路を推定している追跡対象情報を受信したと判断されたホストの数である。ベクトル間距離に対する閾値θを大きくしていくと追跡対象情報を受信したと判断されるホスト数が増加する。この時、尖度を利用した情報の同一性判断基準を利用し、追跡対象受信ホストの増加が妥当なものであるかを評価する。この評価が図２中の右側の条件分岐に相当する。

このアルゴリズムは、追跡対象の情報と同一のものを受信しているかどうかをベクトル間距離により判断し、その時点での追跡対象受信ホストが増加した場合に、その増加が妥当なものであるかを判断していると解釈出来る。このアルゴリズムで尖度を利用することの妥当性を発見したことが本研究で最も大きな成果であると言える。

（３）総ホスト数に対する経路推定成功率

図３にネットワーク内の総ホスト数に対する経路推定成功率を示す。ネットワーク内の総ホスト数が増加すると経路のバリエーションが増加し、推定ミスもそれに伴い増加すると考えられるため、総ホスト数と経路推定成功率の変化を検証した。

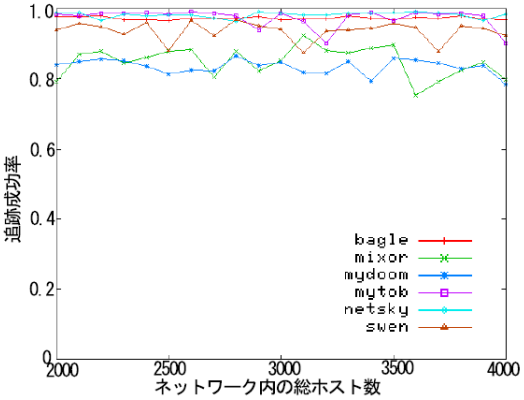


図３から分かるように、総ホスト数が4000まで増加したとしても、経路推定成功率に大きな低下が見られなかった。この結果から、本研究はネットワーク規模への耐性を有し、インターネットレベルでの情報移動の経路推定に適用可能であるということを示唆していると判断出来る。

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４件）


１．　Y. Waizumi、Y. Tsukabe, 、H. Tsunoda, Y. Nemoto 、 K. Tanaka 、 ”Network Application Identification Based on Communication Characteristics of Application Messages” 、 Journal of Communication And Computer、No. 8、2011年、111-119、査読有

２．　和泉勇治、阿部康一、根元義章、”メッセージの遷移パターンに基づくネットワークアプリケーション識別システムの試作”、電子情報通信学会論文誌 D、J93-D、2010年、2257-2267、査読有

３．　K. Simkhada, T. Taleb, Y. Waizumi, A. Jamalipour, Y. Nemoto、“Combating against internet worms in large-scale networks: an autonomic signature-based solution”、 SECURITY AND COMMUNICATION NETWORKS、2009 年、11-28、査読有

４．　H. Tsunoda, K. Ohta, A. Yamamoto, N. Ansari, Y. Waizumi, Y. Nemoto 、“Detecting DRDoS attacks by a simple response packet confirmation mechanism”、Computer Communications、

No. 3、2008 年、3299-3306、査読有

〔学会発表〕(計 4 件)

１． Y. Waizumi, Y. Tsukabe, H. Tsunoda, and Y. Nemoto,”Network Application Identification Based on Communication characteristics of Application Messages”,Proc. of WCSET 2009, No. 6, 708-713、2009 年 12 月 26 日、Bangkok、タイ、査読有

２． Y. Waizumi, T. Sato, and Y. Nemoto, ”A New Traffic Pattern Match-ing for DDoS Traceback Using Independent Component Analysis”,Proc. of WCSET 2009, No. 60,701-707、2009 年 12 月 26 日、Bangkok、タイ、査読有

３． S. Yagi, Y. Waizumi, H. Tsunoda, Y. Nemoto,”A Reliable Network Application Identification Based on Transition Pattern of Payload Length”, Proc. of IEEE Globecom 2008, 2008 年 12 月 2 日、New Orleans, 米国、査読有

４． S.Yagi, Y.Waizumi, H.Tsunoda, A.Jamalipour,N.Kato,Y.Nemoto, ”Netw ork Application Ident -ification Using Transition Pattern of Payload Length”, IEEE WCNC 2008, 2008 年 4 月 2 日、Las Vegas、米国、査読有

６．研究組織

(1)研究代表者

和泉 勇治 (WAIZUMI YUJI)

東北大学・大学院情報科学研究科・准教授

研究者番号：90333872