東北工業大学地域連携センター・研究支援センター紀要 EOS
Vol.35 No.1 p.27 ～ p.32, 2023

令和３年度学内公募研究（萌芽型）
〔研究紹介〕

ホストの通信行動履歴の把握を目的とした トラフィック観測・分析技術の基礎研究

角田　　裕 [1)]，岡部　将也 [1)]

A study on traffic monitoring and analysis for understanding the communication activities of intranet hosts

Hiroshi TSUNODA [1)], Masaya OKABE [1)]

Abstract

For intranet security, an administrator needs to understand the communication activities of each host in the intranet. Here, the communication activity indicates with whom the host communicates, for how long, and for what kind of communication. Such information is helpful for both prevention and post-response to security incidents. This research monitored traffic flow in our laboratory's real-operated network. We analyzed the monitored flows and visualized the flow information using a Sankey diagram to show the communication behavior of the hosts in an easy-to-understand manner. Also, we developed a simple web application to display a Sankey diagram visualizing the flow information of a specified host interactively.

1　はじめに

サイバー攻撃の目的は多様化し，あらゆる組織や個人がその被害者になり得る。攻撃の手口は高度化・巧妙化し続け，攻撃や侵入の検知は容易ではない。そのため，各組織内ネットワーク（イントラネット）では，侵入されることを想定した対応策として，流れるトラフィックを継続的に収集・分析しインシデントの兆候や痕跡を把握するネットワークフォレンジクス（ネットワーク上の科学捜査）が重要となっている。そのためには，イントラネットを流れるあらゆるトラフィックを網羅的に監視し，それに基づいてホストの通信行動を詳らかにすることが望ましい。しかし，現代のイントラネットはユニキャスト通信が基盤のスイッチングネットワークであり，一箇所ですべてのトラフィックの監視はできない。そのため，トラフィックを網羅的に監視するには，すべてのホストとスイッチとの間にトラフィック監視センサをくまなく設置することが必要と

1）東北工業大学　工学部　情報通信工学科
Department of Information and Communication Engineering, Faculty of Engineering, Tohoku Institute of Technology

なり，それはセンサの設置や管理のコストを考えるとほとんどのイントラネットにおいて現実的とは言えない。結果として，センサをイントラネットの出入り口のみに設置し，イントラネット内のホストが外部のホストと通信する際のトラフィックのみを監視するに留まる体制が一般的である。つまり，現在は，出入り口を通過しないイントラネット内部で行われる通信のトラフィックは監視の対象外となっていることが多い。

本研究では，各ホストの通信行動を詳細に把握し分析する基礎研究の起点として，各ホストのネットワークとの接続点にトラフィック観測センサを設置し，送受信するトラフィックのネットワークフロー情報（以下，フロー情報）を観測した。そして，フロー情報をサンキーダイアグラムとしてインタラクティブに可視化できるWebアプリケーションを開発した。通信行動を把握し，わかりやすく可視化することは，管理者やユーザのアウェアネス（意識，気付き）を向上させ，標的型攻撃などいつもと違う不自然な通信に気づく助けとなると考えている。

２　トラフィックのモデル化

イントラネット内のホストは様々な相手との多様な通信を行う。ホストが行う通信は，通信相手の位置と通信形態に基づいて以下の４種に分類できる。

１）同じイントラネットに属するホストとのユニキャスト通信
２）同じイントラネットに属するホスト群とのブロードキャスト・マルチキャスト通信
３）インターネット上（イントラネット外）のホストとのユニキャスト通信
４）インターネット上（イントラネット外）のホスト群とのブロードキャスト・マルチキャスト通信

ここでユニキャスト通信とは送信者と受信者が一対一で行う通信であり，ブロードキャスト・マルチキャスト通信とは１つの送信者に対して複数の受信者が存在する通信である。指定した範囲の全ホストが受信者となる場合はブロードキャスト，指定した範囲の特定のホスト群が受信者となる場合はマルチキャストと呼ぶ。

また，各通信は使用するプロトコルによって更に細分化できる。本研究では，主要となる３種の通信プロトコルとして

（ａ）TCP（Transmission Control Protocol）
（ｂ）UDP（User Datagram Protocol）
（ｃ）ICMP（Internet Control Message Protocol）

に分けて考えることとした。ホストの通信によって発生するトラフィックは図１に示すように，４種の通信のそれぞれで３種類のプロトコルが利用されるため計12種類にモデル化できる。ホストの通信行動を把握するには，この12種類のトラフィックを網羅に観測することが望ましいが，広く行われている観測方式であるトラフィックセンサをイントラネットの出入り口のみに設置した場合では，６種類のトラフィックしか観測できないことがわかる。このことから，本研究では，各ホストの通信行動を詳細に把握し分析するために，各ホストのネットワークとの接続点にトラフィック観測センサを設置し，送受信するトラフィックのフロー情報を観測し，その分析を実施する体制を構築することとした。

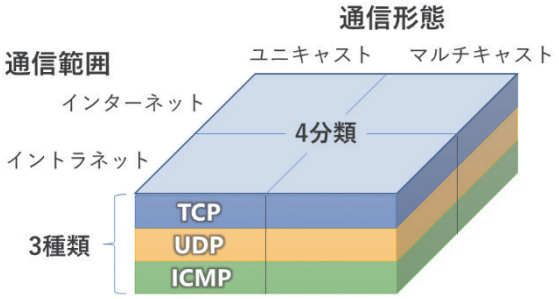

図 1　ホストが送受信するトラフィックのモデル

3　NetFlow を活用したトラフィック観測

本研究では，シスコシステムズ社が開発したネットワークフロー観測技術 NetFlow［1］［2］を使用して，研究室ネットワークのトラフィック観測体制を構築した。ネットワークフローとは，送信元 IP アドレス，宛先 IP アドレス，送信元ポート番号，宛先ポート番号，トランスポートプロトコルの 5-tuple が等しい一連のパケット群である。本研究で使用した NetFlow version 5 では 5-tuple に加えて入力インターフェースと IP ヘッダの ToS（Type of Service）フィールドも等しいパケットを同一のフローに属するものと判断する。ホストが送受信するネットワークフローを観測することで，そのホストが「いつ」「どのホストと」「どのようなアプリケーションで」「どのくらいのトラフィック量や継続時間の」通信を実行したか把握できるようになり，ホストの通信行動の把握につながる。従来，NetFlow による観測は大規模ネットワークやネットワークの境界地点において広く行われているが，本研究ではイントラネット内部での観測にこの技術を活用することとした。

本研究では，図 2 に示すネットワークフロー観測環境を研究室ネットワーク内に構築した。NetFlow による観測は，プローブとコレクタによって行われる。プローブは，ネットワークを流れるパケットからフロー情報を取り出す機能を持つ装置であり，コレクタはフロー情報を収集し蓄積する装置である。プローブは取り出したフロー情報をコレクタに送信し，コレクタではフロー情報を受信して蓄積するとともに統計処理などの前処理も実行する。構築した環境では，注目対象のホストがイントラネットに接続するスイッチとの間に，ポートミラリング機能を持ったスイッチ（ミラーリングスイッチ）［3］を設置し，ホストが送受信するパケットをすべてコピーして NetFlow プローブへ転送するようにした。NetFlow プローブのハードウェアにはシングルボードコンピュータの Raspberry Pi 4 Model B［4］，ソフトウェアにはオープンソースの fprobe［5］を使用した。また NetFlow コレクタには，汎用のデスクトップ PC 上にインストールした nfdump［6］を使用した。

なお，プローブが取り出すフロー情報にはパケットの送信元と宛先については IP アドレスの情報しか含まれていない。そこで，本研究独自の工夫として，フローとともにホストが送受信する DNS（Domain Name System）のパケットもキャプチャし，通信相手のホスト名も特定できるようにしている。

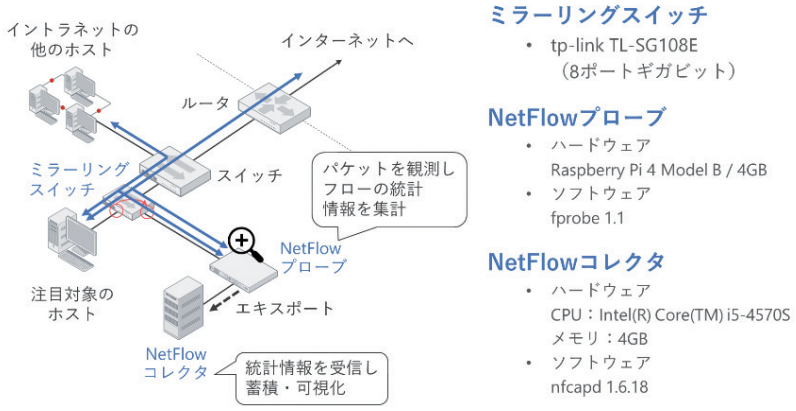

図２　ネットワークフロー観測環境

4　トラフィック観測結果の分析とホストの通信行動の可視化

観測したフローに基づいてホストの通信行動をわかりやすくネットワーク管理者に提示するために，図３のようにサンキーダイアグラム［7］として可視化した。サンキーダイアグラムは，２点間のモノの流れとその量を視覚的に表現した図であり，物流管理等の工程間の流量を表現するために使用される。本研究では，送受信する２つのホスト間の通信データの流量を可視化した。図３では，左端が注目対象のホスト，右端がそのホストの通信相手となるエンドポイント（IP アドレスとポート番号の組），両端を結ぶ帯の幅がホストとエンドポイント間の通信量を表している。この図からは，2022 年２月 14 日 10 時から 11 時の１時間の間に，注目しているホスト（IP アドレス 192.168.1.71）が，ビジネスチャットである Slack のサーバ（wss-primary.slack.com）と HTTPS による通信が行われ，約 500Kbyte のデータの送受信が発生したことがわかる。このようなフロー情報とその可視化結果の蓄積はホストの通信行動を把握する上で重要であり，今後は研究室ネットワーク内での観測を継続して実施する予定である。

次に，任意の時刻のホストの通信行動を効率的に可視化するために，指定時刻のサンキーダイアグラムをインタラクティブに生成する Web アプリケーションを開発した

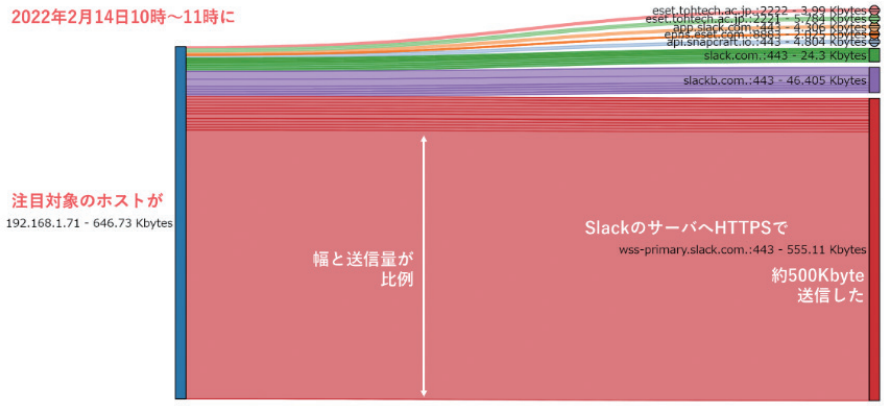

図３　サンキーダイアグラムによる可視化の例

［８］。開発言語には Python 3.8.2，Web フレームワークとして Streamlit 1.7.0［９］，サンキーダイアグラムの生成には可視化ライブラリ Holoviews 1.14.8［10］と Bokeh 2.4.1［11］を使用した。

開発した Web アプリケーションのスクリーンショットを図４に示す。アプリケーションのユーザインターフェースは可視化条件を指定するサイドバーとサンキーダイアグラムを表示するメイン画面に分かれており，可視化するフローを一定の条件で動的に選別するフィルタリング機能と，宛先ドメイン名に基づく集約機能がある。フィルタリング機能は，指定した条件に合致したフローとそれに関係するエンドポイントのみを描画する機能である。条件として，２節で検討したモデルに基づいて，通信形態（ユニキャスト，マルチキャスト），通信範囲（インターネット，イントラネット），プロトコル（TCP，UDP，ICMP）を指定できるようにした。集約機能は，フローの宛先のエンドポイントを IP アドレスではなくドメイン名で識別し，同一のドメイン名を持つエンドポイントを集約する機能である。IP アドレスが異なるエンドポイントでも，アドレスに対応するドメイン名が等しければ，同じとみなして集約するため，結果として描画されるエンドポイント数が減少する。この時サンキーダイアグラム上には，エンドポイントのドメイン名を表示させる。

アプリケーションの動作検証として，研究室ネットワーク内のある１台の Linux 端末が 2022 年２月 15 日の４時～５時に送受信した 48 個のエンドポイントとの間の 466 本のフローを可視化した。フィルタリング機能によって，通信範囲がインターネットで TCP のユニキャストのフローのみを残したところ，エンドポイントが 17 個，フローが 80 本まで減少した。さらに同一のドメイン名を持ったエンドポイントを集約したところ，エンドポイントは 15 個まで減少した。このように，開発したアプリケーションを使用することで所望の通信を絞り込んで可視化することができ，通信行動把握の効率化

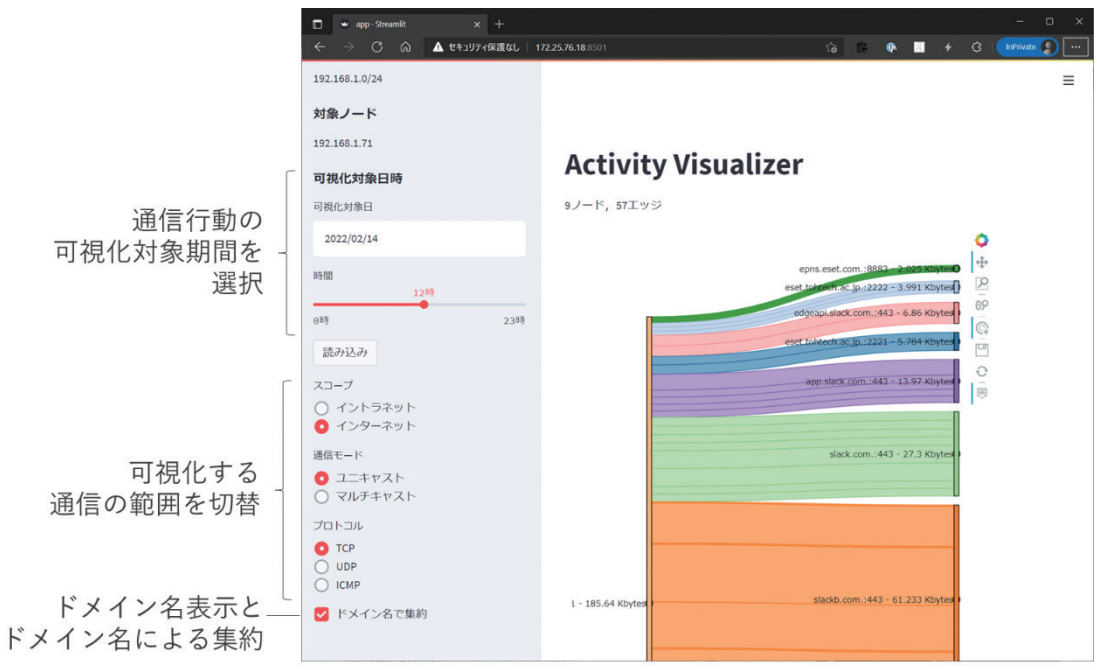

図４　開発した可視化アプリケーションのスクリーンショット

につながる。今後は，通信量が一定以上のフローだけを表示するフィルタリングや，考慮するドメイン名の範囲を制限したエンドポイントの集約など，さらにわかりやすい可視化の方法を検討する。

まとめ

本研究では，イントラネット内のホストの通信行動を把握し，セキュリティインシデントの未然防止や事後対応に役立てることを目指し，基礎的検討を行った。具体的には，NetFlow 技術を用いたイントラネット内でのフロー観測を実施し，観測結果をサンキーダイアグラムとしてインタラクティブに可視化できる Web アプリケーションを開発した。フロー情報をサンキーダイアグラムとして可視化することで，ホストが「いつ」「誰と」「どのような通信を」「どのくらいの量」行ったかをわかりやすく把握できるようにした。今後は，開発した Web アプリケーションの機能強化を進めるとともに，把握した通信行動の中から不正な通信の発見するための手法を検討する。

謝辞

本研究は，東北工業大学学内公募研究（2021-01）の援助により行われたものである．ここに記して謝意を表する．

参考文献

［1］Michael W. Lucas 著，株式会社クイープ訳，ネットワークフロー解析入門 flow-tool によるトラブルシューティング，株式会社アスキー・メディアワークス，2011
［2］Rick Hofstede, Pavel Celeda, Brian Trammell, Idilio Drago, Ramin Sadre, Anna Sperotto and Aiko Pras,“Flow Monitoring Explained: From Packet Capture to Data Analysis with NetFlow and IPFIX,”IEEE Communications Surveys & Tutorials（Volume: 16, Issue: 4, Fourthquarter 2014
［3］tp-link,“TL-SG108E 8 ポート ギガビット アンマネージ プロ スイッチ”https://www.tp-link.com/jp/business-networking/easy-smart-switch/tl-sg108e/（参照 2022-10-01）
［4］“Buy a Raspberry Pi 4 Model B – Raspberry Pi,“ https://www.raspberrypi.com/products/raspberry-pi-4-model-b/（参照 2022-10-01）
［5］fprobe，SOURCEFORGE，https://sourceforge.net/projects/fprobe/（参照 2022-10-01）
［6］nfdump，Github，https://github.com/phaag/nfdump（参照 2022-10-01）
［7］矢崎裕一，サンキー・ダイアグラム（Sankey Diagram），visualizing.jp，2020-8-2，https://visualizing.jp/sankey-diagram/（参照 2022-10-01）
［8］岡部 将也，角田 裕，“ホストの通信行動把握を支援する通信フロー情報の可視化システム”，2022 年度電気関係学会東北支部連合大会 4F03，2022 年 8 月
［9］Streamlit・The fastest way to build and share data apps, https://streamlit.io/（参照 2022-10-01）
［10］Installation – HoloViews v1.15.0, https://holoviews.org/（参照 2022-10-01）
［11］Bokeh, http://bokeh.org/（参照 2022-10-01）