6F-01

Tableau Desktopを用いたIPアドレスのネットワーク部に着目した侵入検知システムアラートの可視化

輪島　幸治†　高橋　健志†　井上　大介†

† 国立研究開発法人 情報通信研究機構

1 はじめに

今日，侵入検知システム (IDS) を対象とした研究では，不正通信検知を目的として，多くのセキュリティ対策の研究が提案されている [1]．各種通信機器におけるマルウェア感染の検知や，通信機器に対する DDoS 攻撃イベントの検出を目的とする場合，特定日時における通信状況から，異常値となる IP アドレスを検知して，フィルタリングを適切に処理する必要がある．そこで本研究では，各月に分割された IDS のアラートデータを対象に，通信状況およびアラートデータを可視化して，異常値となっている IP アドレスなどの検知や，時系列データの可視化からスパイクポイントなど，異常値が検出した日時の特定を試みる．本研究では，アラートデータの可視化アプリケーションに，パッケージソフトウェアである Tableau を用いて実験を実施した [1]．Tableau Desktop を用いて，IP アドレスを可視化した例を図 1 に示す．

図 1　Tableau Desktop によるアラート可視化の例

ところで，可視化アプリケーションを用いた分析だが，クラウド型アプリケーションを用いて，アラートデータを可視化できることは，既に示されている [2]．しかし，アラートは組織内部の通信情報を含む機微なデータであり，クラウドアプリケーションで機微なデータを管理することが，セキュリティポリシ違反となる組織も少なくない．そこで，オンプレミスで用いられる可視化アプリケーションである Tableau Desktop を用いた．発電所やプラントにおけるシステム環境 [3] や，プラントのタンク設備を監視する制御システムを対象としたシステムなど，重要インフラではインターネット接続が制限されることも多く，オンプレミスでの分析が必須となる環境は多い [2]．

Visualization of Intrusion Detection System Alerts Focused on the Network Part of IP Address using Tableau Desktop

Koji Wajima†,Takeshi Takahashi†,Daisuke Inoue†

†National Institute of Information and Communications Technology

[1]Tableau Software:https://www.tableau.com/

[2](参考) 横河電機 ライブラリ：https://www.yokogawa.co.jp/library/

2 IPアドレスのネットワーク部

本研究の提案では，アラートの可視化軸に IP アドレスのネットワーク部を用いることに特徴を持つ．一般に，組織における IP アドレスの割り当ては，部署や研究室，サーバ室，敷地内の建築物など場所ごとに異なる．本研究では，異常が発生しているネットワーク部を明らかにして，特定の IP アドレスのみを可視化することや，特定の IP アドレスやポート番号をフィルタリングして，通信が少ない宛先を分析することを目的に，IP アドレスをネットワーク部とホスト部に分割して，分析することを試みた．異常値となっているネットワーク部が明らかとなることで，新たな IDS ルールの導入や，詳細分析，セキュリティ対策などを迅速に実施可能となることが期待できる．本研究で用いる IP アドレスの分割箇所を下記に示す．

$$\underbrace{XXX.XXX}_{(1)}.\underbrace{XXX}_{(2)}.\underbrace{XXX}_{(3)}$$

(1)　IP アドレス ネットワーク部：16 ビット

(2)　IP アドレス ホスト部 (1)：8 ビット

(3)　IP アドレス ホスト部 (2)：8 ビット

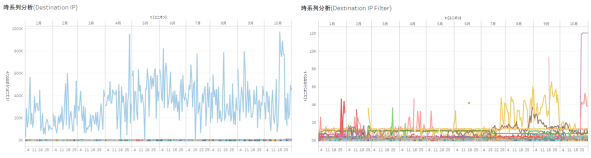
IP アドレスのネットワーク部とホスト部の設定には，クラスフルと呼ばれる 8 ビット単位に固定で設定する方法と，クラスレスと呼ばれる CIDR（Classless Inter-Domain Routing）を利用して，可変長で設定する方法がある．本研究における，ネットワーク部の取り扱いでは，ネットワーク部は 16 ビットに設定した．そして，ホスト部は，残りの 16 ビットを 8 ビットづつ 2 個に分割して設定した．

3 評価実験

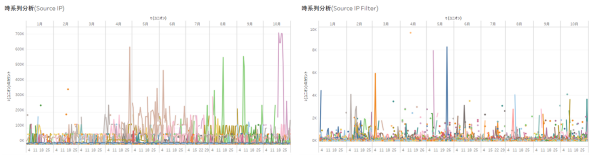
本研究の評価対象として，情報通信研究機構の LAN で 2017 年 1 月 1 日から 2017 年 10 月 31 日までの 10ヶ月間に発報された IDS アラートのデータセットを用いた．本研究で評価対象としたアラートデータは，プログラミング言語である Python を使用して，各月で CSV ファイルに分割する前処理を行っている．評価実験に用いた CSV ファイルの総レコード数（アラート数）は 131,902,019 件，合計サイズは 27.9GB であった．Tableau Desktop では，各月に分割された CSV ファイルを読み込み処理し，データのユニオンを作成してから可視化を行った．検証環境として，Windows 10 Pro の Hyper-V の仮想環境上で，Windows Server 2019 環境を構築した．Tableau Desktop をインストールした Windows Server 2019 環境のメモリサイズは，24GB を設定して，評価実験している．

まず，送信先 IP アドレスおよび送信元 IP アドレスのネットワーク部に基づき，可視化した結果を，図 2 および図 3 に示す．

図 2 および図 3 は，縦軸がアラート件数，横軸が日時，各折れ線グラフの種別は IP アドレスのネットワーク部である．図 2 および図 3 における (1) および (2) は，可視化グラフにおけるフィルタの適用前およびフィルタの適用後である．図 2 のフィルタには，異常値 IP を使用しており，フィルタの適用後のグラフである (2) は，(1) における異常値 IP を除外した結果である．また，図 3 のフィルタには，アラートレコード数を使用しており，フィルタの適用後のグラフである (2) は，(1) におけるアラート件数が多い IP アドレスを除外した結果である．



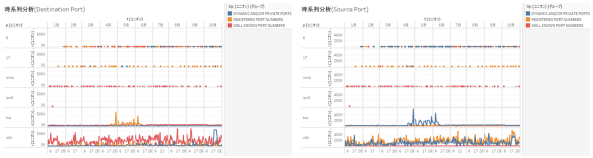
(1) 異常値 IP フィルタ前　　(2) 異常値 IP フィルタ後
図 2　送信先 IP アドレス ネットワーク部に基づいた時系列可視化



(1) 件数上限値 フィルタ前　　(2) 件数上限値 フィルタ後
図 3　送信元 IP アドレス ネットワーク部に基づいた時系列可視化

結果，図 2 の (2) から，異常値となっているネットワーク部を持つ IP アドレスをフィルタすることで，異常値でないネットワーク部のスパイクポイントが明らかとなり，時系列分析に効果的であることが明らかとなった．また，図 3 の (2) から，アラート数が多いネットワーク部を持つ IP アドレスをフィルタすることで，アラート数が少ないネットワーク部における異常値を評価するのに効果的であるのが明らかとなった．

次に，プロトコルとポート番号で時系列分析した結果を図 4 の (1) および (2) に示す．図 4 におけるポート番号は，“WELL KNOWN PORT (0 番-1023 番)”，“REGISTERED PORT (1024 番-49151 番)”，“DYNAMIC AND/OR PRIVATE (49152 番-65535 番)” の 3 グループにグループ化している．縦軸は，プロトコルの種類とレコード件数，横軸は，アラート発生日時である．

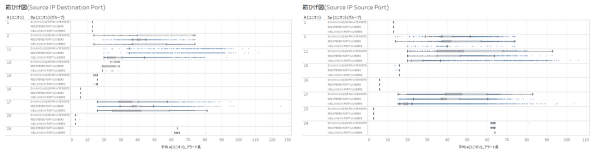


(1) 送信先ポート　　(2) 送信元ポート
図 4　ポート番号に基づいたアラートデータの時系列可視化

結果，ポート番号で時系列を可視化することで，通信が多い日時が明らかとなった．また，図 4 では，(1) および (2) における TCP プロトコルで 2017 年 4 月 30 日付近に，スパイクポイントがあることが明らかとなった．可視化したスパイクポイントだが，図 4(1) は “REGISTERED PORT”，図 4(2) は “DYNAMIC AND/OR PRIVATE PORTS” で得られた．ネットワーク部で可視化した図 2 および図 3 と比較した場合，スパイクポイントは，フィルタ対象としたネットワーク部を持つ IP アドレスによる通信であることが明らかとなった．

最後に，箱ひげ図で IP アドレスのネットワーク部ごとに，受け取るアラートのメッセージ長の平均値にバラつきを可視化した．結果を図 5 および図 6 および表 1 で示す．図 5 および図 6 の縦軸は，アプライアンスとポート番号のグループ，横軸はアラートのメッセージ長である．表 1 の値は，各箱ひげ図で算出された五数要約した値を “WELL KNOWN PORT” のみを対象にして，算出した値である．

(1) 送信先ポート　　(2) 送信元ポート
図 5　送信先 IP アドレスに基づいた各アプライアンスでの箱ひげ図



(1) 送信先ポート　　(2) 送信元ポート
図 6　送信元 IP アドレスに基づいた各アプライアンスでの箱ひげ図

表 1　五数要約による要約統計量（WELL KNOWN PORT）

	Destination IP		Sorce IP	
	Destination Port	SorcePort	Destination Port	SorcePort
ヒゲの上端	22.00	72.37	75.6	35.45
上部ヒンジ	22.00	42.25	45.4	26.10
中央値	22.00	22.48	38.5	20.53
下部ヒンジ	22.00	22.00	25.1	19.87
ヒゲの下端	22.00	3.00	3.00	10.78

可視化の結果，IP アドレスのネットワーク部および各アプライアンスで，受け取るアラートメッセージの長さにおける平均値にバラつきがあることが明らかとなった．また，表 1 から，長いアラートメッセージばかり受け取っているネットワーク部を持つ送信元 IP アドレスがあることも明らかとなった．

4 ま と め

本研究では，侵入検知システムのアラートデータを対象に，IP アドレスのネットワーク部に着目して，可視化を行った．結果，フィルタすべき異常値 IP アドレスや，異常値が発生している日時，長いアラートメッセージばかり受け取っているネットワーク部を持つ送信元 IP アドレスを可視化することができた．今後は，より詳細な可視化分析を行い，マルウェア発生日時や，DDoS 攻撃イベントの日時と比較して，分析したい．

謝　　辞

情報通信研究機構の IDS データセットを提供し，研究の議論の機会を与えてくれた研究者である班 涛氏に感謝する．

文　　献

[1] Martin Husák, Jana Komárková, Elias Bou-Harb, and Pavel Čeleda. Survey of attack projection, prediction, and forecasting in cyber security. *IEEE Communications Surveys & Tutorials*, Vol. 21, No. 1, pp. 640–660, 2018.
[2] 輪島幸治. セキュリティベースのデータ変換プロセスに基づく salesforce einstein analytics を用いたデータ解析. インターネットと運用技術シンポジウム論文集, pp. 95–96, nov 2020.
[3] 統合生産制御システム CENTUM VP. 横河電機株式会社, 閲覧日 2020 年 12 月 30 日.
[4] 製造情報活用パッケージ TriFellows Ver.5. 横河ソリューションサービス株式会社, 閲覧日 2020 年 12 月 31 日.