リアルタイムトラフィック可視化システムとその運用経験

吉田和幸[1]　宇野秀亮[2†]　池部実[2]　吉崎弘一[1]

概要： インターネットトラフィックの増大に伴い，サーバやそこで動作しているサービスを調査するスキャンも増加している．大量のパケットが来るスキャン等の傾向を把握するには，個々のパケットやフローの解析では困難である．我々は、スキャンの状況をリアルタイムに表示するシステムを作成し，運用している．本システムでは、外部から LAN 宛てに来るパケットについて、IP ヘッダ、TCP ヘッダの内容により分類、抽出して、宛先 IP アドレスの下位２オクテット，宛先ポート番号など 16 ビット空間でそれぞれに対応するパケット数を計数し、256×256 のマップに展開して全体の分布を表示している．これにより，ネットワークトラフィックの概要の把握が容易になる。本論文では、システムの構成、実行例を示し、可視化した分布図から読み取れるスキャンの傾向などについて考察する。

Realtime Network Traffic Visualizer and Its Operational Experience

KAZUYUKI YOSHIDA[1]　HIDEAKI UNO[2†]　MINORU IKEBE[2]
KOICHI YOSHIZAKI[1]

1. はじめに

インターネットの普及に伴い，インターネットのトラフィックも増大している.「我が国のインターネットにおけるトラヒックの集計・試算」[1]によると，2020 年 11 月時点で日本における固定系ブロードバンド契約者の総ダウンロードトラフィックは約 19.8Tbps であり，前年同月比で 56.7%増加している．また，ブロードバンドサービス契約者の総アップロードトラフィックは約 2.4Tbps であり，前年同月比で 51.1%増加している．これに伴いサーバやそこで動作しているサービスを調査するスキャンも増加している．個々のパケットや，サーバ・クライアント間の一連のパケット(フロー)を詳細に調べることは，不正侵入等の検知には重要であるが，大量のパケットが来るスキャン等の傾向を把握するには，個々のパケットやフローの解析では困難である．宛先 IP アドレスの第 3，第 4 オクテットやポート番号など，パケットヘッダ中の 16 ビットの箇所について 65536 に分類し、それぞれを計数して 256×256 に展開し、リアルタイムに表示することにより、ネットワークトラフィックの分布をリアルタイムに表示するシステムを作成した．

本論文では，2 章で関連研究について述べ，3 章でシステム構成について述べる.4 章で運用例を示して，5 章でまとめと今後の課題について述べる．

2. 関連研究

NICTER(Network Incident analysis Center Tactical Emergency Response)[2]は，サイバー攻撃観測・分析・対策システムである．ダークネットを観測することで収集したパケットデータを解析している．ダークネットとは，インターネットに向けて公開されている IP アドレスのうち，未使用のアドレスによって構成されたネットワークであり，通常の通信では，存在するホストに対してパケットを送信するため未使用のアドレス空間宛に通信が発生することはない．そのため，ダークネットで観測されるすべてのトラフィックは不正な通信である可能性が高い．ブラックホールモニタリングと呼ばれる，受信されるパケットに対して一切応答せず，記録する手法を取っている．
NICTER による観測・分析結果の一部は，NICTERWEB 2.0[2]によって一般公開している．NICTERWEB 2.0 ではダークネットトラフィックを，立方体にマッピングした「Cube」や，地図上に表示した「Atlas」によって，リアルタイムに可視化している．

新川ら[3]は IP アドレスの下位 8bit，ポート番号の下位 8bit を用いた 2 次元マトリクスを表示する手法を提案，実装した．2 次元マトリクス上の点をクリックすると，その通信に関する IP アドレス, ポート番号, 時間, プロトコル, 方向が表示される．しかし，下位 8bit が重複するホストの通信が複数ある場合，重なって表示されてしまうため，ホストごとの正確な通信状況を把握できない．

宇都木ら[4]は TCP コネクションごとのトラフィックを

1 大分大学 情報基盤センター
Information Technology Center, Oita University
2 大分大学 理工学部
Faculty of Science and Technology, Oita University

† 現在，（株）九州テン

帯状に表示し，リアルタイムなトラフィック量によって色のグラデーションで示すことでトラフィックを可視化した．ポート番号や IP アドレスによってトラフィックをソートしたり，フィルタしたりできる．トラフィックの可視化に特化しており，パケットの宛先の分布などを把握できない．

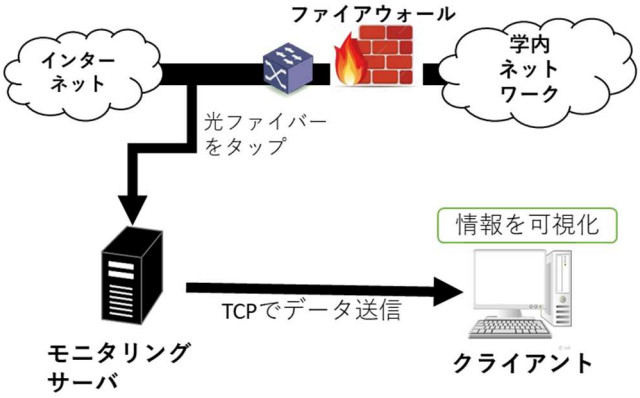

図 1 システム構成

3. リアルタイムトラフィック表示システム

3.1 システム構成

図 1 にシステム構成を示す．インターネットへのアクセス線(10Gbase-SR)×2 の光ファイバをタップにより分岐して，モニタリングサーバに送る．モニタリングサーバでは，tcpdump コマンド[5]によりインバウンド（外から内）パケットのヘッダ情報を抽出・収集し，ncat コマンド[6]を用いて，クライアント PC へ送る．クライアント PC では，Java プログラムにより，データを読み込み，必要な部分を抽出・計数し，結果を表示する．Java プログラムについては次節で詳述する．

3.2 プログラム構成

プログラムの構成を図 2 に示す．モニタリングサーバから毎秒 30000 個ほどのパケットヘッダが送られてくる．入力部は，独立した thread とし，入力したパケットヘッダ情報をパラメータに抽出＆計数部のメソッドを呼び出し，計数させる．

図 3 に，抽出＆計数部の計数するメソッド（cnt）と，結果を表示部へ渡すメソッド（getd）の例を示す．この例では ICMP パケットの宛先 IP アドレス（学内ネットワーク側のアドレス）の第 3、第 4 オクテットを抽出し、個々のアドレスごとにパケット数を計数する。また、計数結果がオーバフローしないように，指数平滑移動平均を計算するため thread で 10 秒ごとに計数結果に 1/2 を乗じている．

cnt メソッドは，パケットの先頭 94byte（イーサヘッダ 14byte, IP ヘッダ 20～60byte, TCP ヘッダ 20byte） を受け取り，計数したいパケットを抽出し，計数する．getd メソッドは，2 つのインデックス値(0～255)をもらい，対応する計数結果を返している．

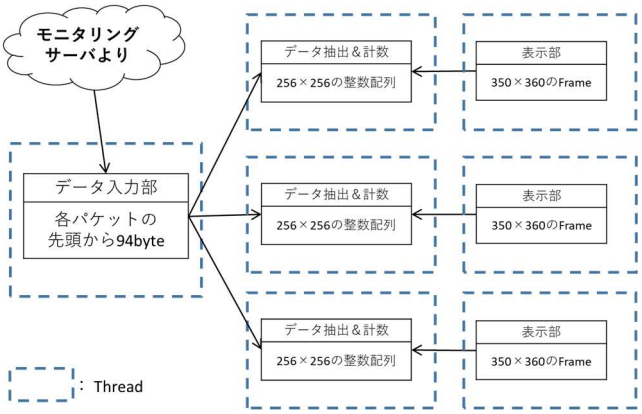

図 2 プログラム構成

```
protected int d[][]=new int[256][256];
public boolean cnt(byte b[]) {
        if(b[23]==1)                    // protocol=ICMP
            d[b[32]&0xff][b[33]&0xff]++;
                // 宛先IPアドレスの３，４オクテット
        return r;
}
public int getd(int x,int y) { return d[x][y]; }
```

図 3 抽出・計数部のプログラム例

図 4 カラーマップ

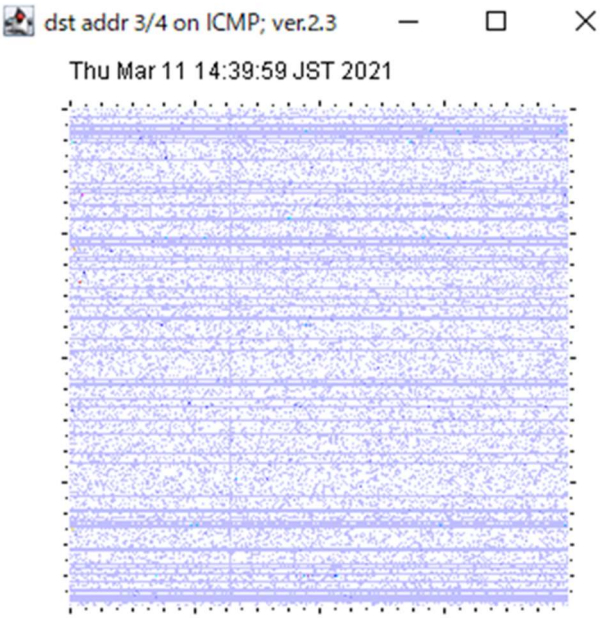

図 5 ICMP パケットの宛先 IP アドレスの分布

表示部は，抽出＆計数部から計数結果の 256×256 の整数配列をもらって，計数結果を色で表して，256 dot×256 dot の散布図として表示する．図 4 に計数結果から色への対応を示す。計数値を白(0)～青(4)～水色(8)～黄色(16)～

赤(32)～マゼンタ(64)～紫(128)～黒(255)へマップしている(()内の数は計数値)．値が小さいところでは，早く変化し，値が大きいところでは，ゆっくり変化するようにしている．図 5 に表示例を示す。抽出している 16 ビットのうち上位 8 ビット(0～255)を縦軸に、下位 8 ビットを横軸にして 256×256 のマトリックスで表示し、上下端、左右端に 8 ごとにメモリを表示している。この例では、インバウンドパケットの宛先 IP アドレスの第 3 オクテットを縦軸に示し、第 4 オクテットを横軸に示す。なお、縦軸では、上端が 0、下端が 255、横軸では左端が 0、右端が 255 である。

表示部は，thread とし、他の部分とは独立に自身のタイミングで描画させている．

4. 運用例

運用例を図 5 のほかに図 6～図 13 に示す．これらは，本システムを 52 分間実行した結果であり，モニタリングサーバ上で動かしている tcpdump コマンドによると 84,750,410 パケット受信し，233,253 パケット(0.3%)取りこぼしている．10 秒ごとに値を半分にする指数平滑移動平均値を計算しているため，おおむね 20 秒間のパケット到着数を示している。

図 6 は，インターネットから学内ＬＡＮに入ってくるすべてのパケット(インバウンドパケット)の宛先 IP アドレスの第 3，第 4 オクテットの分布である．図 7、図 8 は、それぞれインバウンドパケットの内、TCP のコネクション確立要求である SYN フラグのみ立っているパケット、UDP パケットの宛先アドレスの分布である。

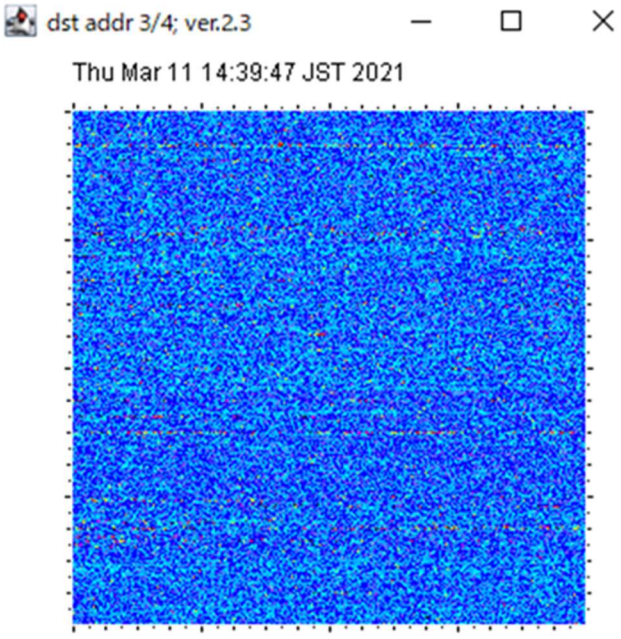

図 6 宛先 IP アドレスの分布

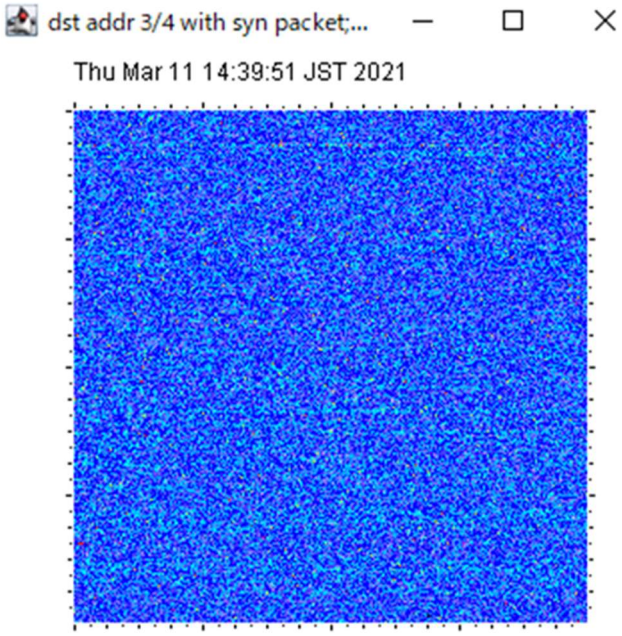

図 7 SYN パケットの宛先アドレスの分布

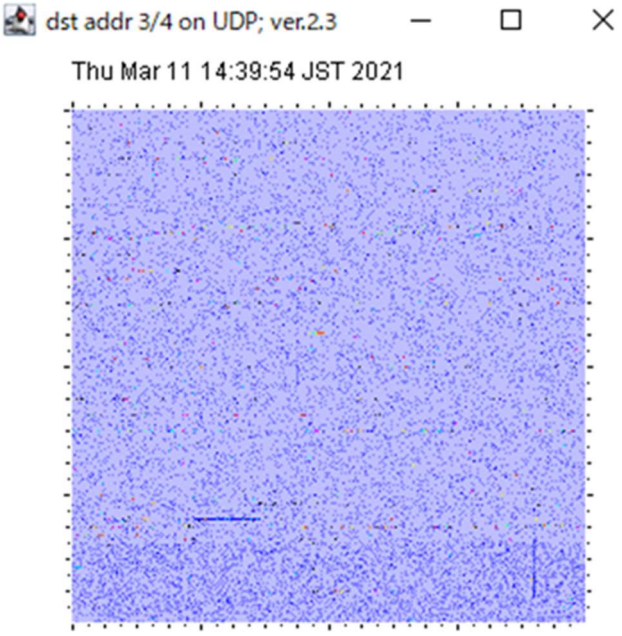

図 8 UDP パケットの宛先 IP アドレスの分布

図 6 と図 7 とは，図 5、図 8 に比べて計数値が大きい(色が濃い)．スキャンの大部分は，TCP スキャンであることがわかる．この 2 つの図からはスキャンが多すぎて、どのようにスキャンが行われているかは不明である。.図 5 は，横線が目立つので，第 3 オクテットを固定して，第 4 オクテットを変化させるスキャンが、頻繁に行われていることがわかる。左 1/3 あたりに縦線も確認でき、第 4 オクテットを固定して，第 3 オクテットを変化させるスキャンが行われたことがわかる．図 8 については、アドレス空間全体

をスキャンされてはいるが，左下に短い横棒、右下に短い縦棒が確認できる。

図 9 は, すべてのインバウンドパケットのポート番号の分布，図 10 は TCP の SYN フラグのみ立っているパケットのポート番号の分布である．図 9 では，下 1/4(ポート番号の上位バイトが 192 以降)の部分が多くなっている．これば, 学内の PC が外部のサーバにアクセスするときの PC 側

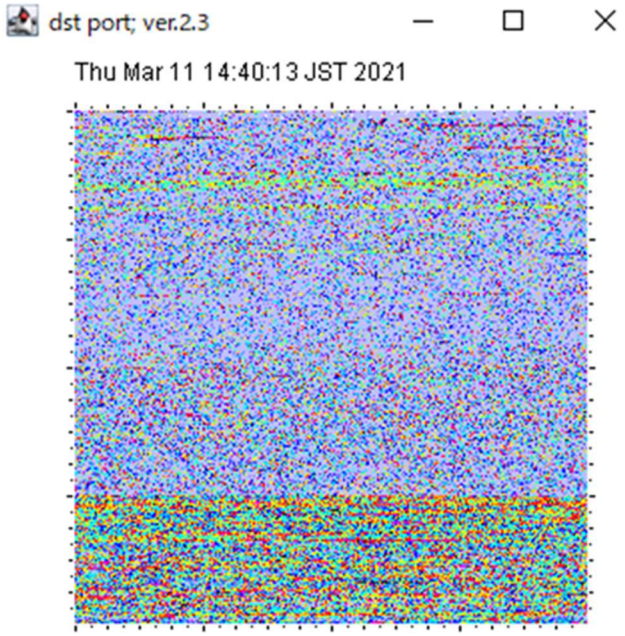

図 9 宛先ポート番号の分布

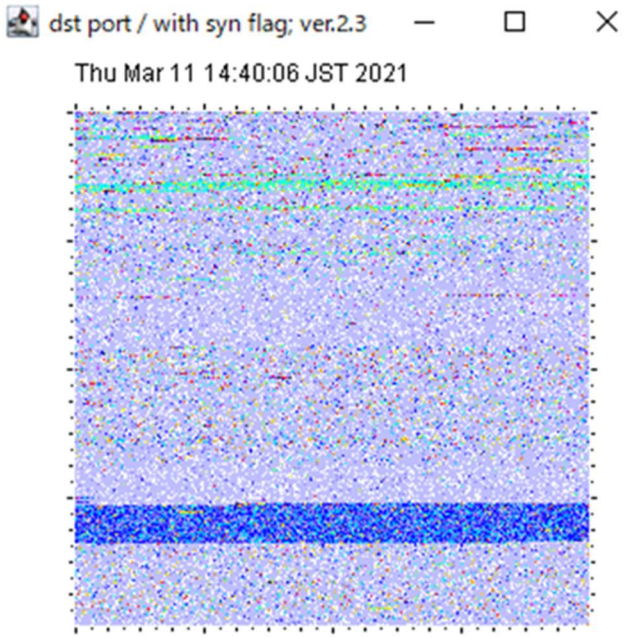

図 10 SYN パケットの宛先ポート番号の分布

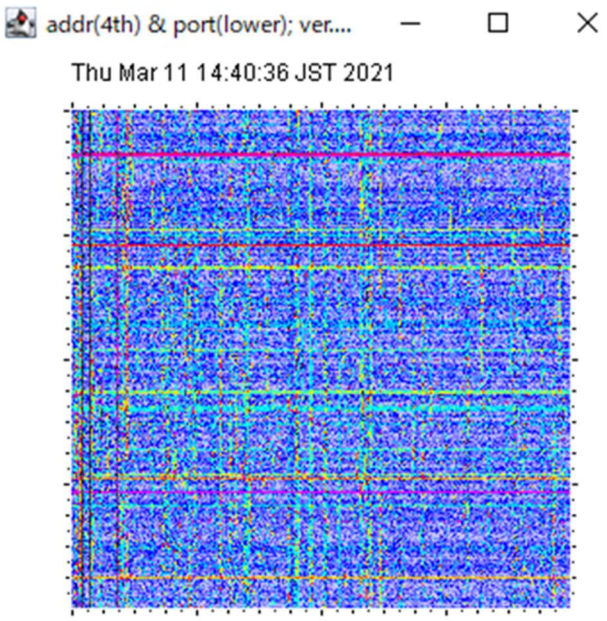

図 11 宛先アドレスとポート番号の分布

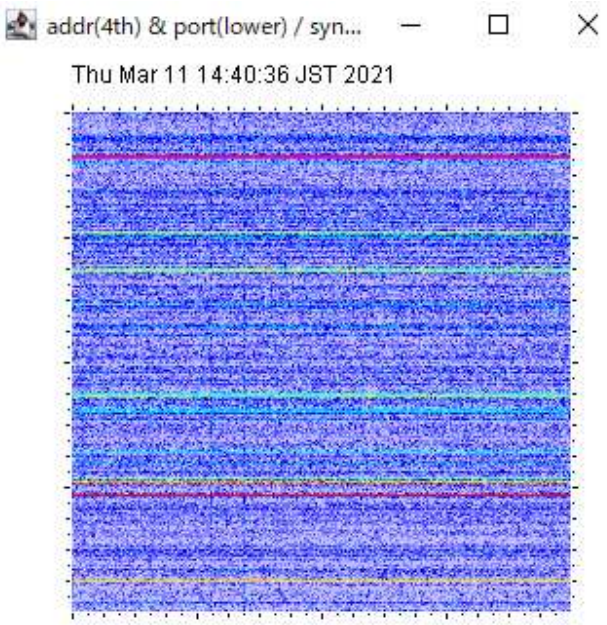

図 12 SYN パケットの宛先アドレスとポート番号の分布

のポート番号(エフェメラルポート)であろう．この影響を排除するため, 外部からの TCP コネクション要求に絞ると図 10 になる．上位バイトが 32～40 付近(ポート番号 8000～10000)と 192～216 付近(ポート番号 49000～55000)に，スキャンが多いことがわかる．

文献[3]にならい，図 11 に，すべてのインバウンドパケットの宛先 IP アドレスの第 4 オクテットとポート番号の下位バイトの分布を，図 12 に，TCP の接続要求(SYN フラグのみ立っている)パケットの宛先 IP アドレスの第 4 オクテットとポート番号の下位バイトの分布を示す．この 2 つ

の図では、縦軸はポート番号（の下位バイト）を、横軸はIP アドレス（の 4 バイト目）を表す。図 11 で縦線が目立つのは，図 9 と同様にエフェメラルポートの影響である．それを排除した図 12 では，横線(ポートを固定して，IP アドレスを変化させる)が目立っている．特定のサービスが立ち上がっているサーバを探していると思われる。

図 13 にパケット長の分布を示す．イーサネットの最大パケット長は 1500 であり、大部分のパケットの長さはその範囲に収まっているが，モニタリングサーバのインターフェースカードにおいてハードウェアオフロード機能により，TCP パケットが最大 window size になるまで結合された長大なパケットを生成できるので、本システムではそれを観測している．長大なパケットの長さが，8 の倍数に集中しているので，この時は，Window size 拡張オプションが 3 (8 倍)の TCP コネクションで，大量のデータの受信を行っていたことがわかる．

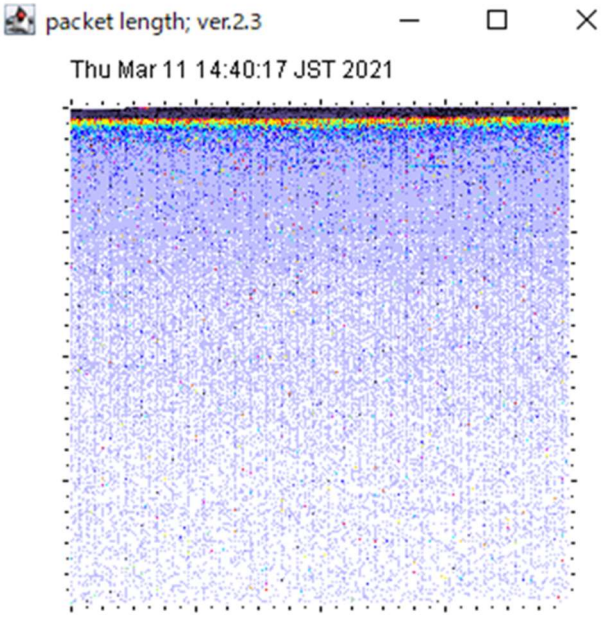

図 13 パケット長の分布

5. おわりに

インターネットから LAN に向かって流れてくるパケットをすべてキャプチャし、ヘッダ情報を解析し、パケット数を計数することにより、スキャン活動の宛先 IP アドレスの分布、ポート番号の分布等をリアルタイムに可視化するシステムを開発し、運用してきた。

本システムにより、スキャン活動の実態を可視化できるようになった。特にポート番号の分布では、8000～10000、49000～55000 の 2 か所に集中していることがわかった。UDP, ICMP のスキャンは、IP アドレスの第 4 オクテットを変えながらスキャンしている等の傾向がわかるが TCP のスキャンは、大量に来ているため、すべて埋まってしまい、傾向を可視化できていない。わかりやすい可視化ができるよう本システムのパラメータを調整していきたい。

参考文献

[1] 総務省，“我が国のインターネットにおけるトラヒックの集計・試算”，https://www.soumu.go.jp/menu_news/s-news/01kiban04_02000182.html， 2021.2
[2] nicter，https://www.nict.go.jp/，国立研究開発法人情報通信研究機構,閲覧日: 2021 年 2 月 4 日.
[3] 新川拓也,山之上卓,IP アドレスとポートによる二次元平面を用いた通信トラフィックの可視化について,情報処理学会研究報告, Vol.2006,No97,pp.31-36,2006 年 9 月 15 日.
[4] 宇都木進,渡邉晶,TCP コネクション単位でトラフィックの可視化を行うツールの開発,情報処理学会研究報告,Vol.2009-IOT-7,No.4, pp.1-5,2009 年 10 月 9 日.
[5] tcpdump コマンド, https://www.tcpdump.org
[6] ncat コマンド, https://nmap.org/ncat/