インターネットトラフィックの特徴抽出に向けて

4H−05

小出和秀[†]　齋藤武夫[†‡]　Glenn Mansfield Keeni[‡]　白鳥則郎[†1]

[†]東北大学電気通信研究所/情報科学研究科[2]
[‡]株式会社サイバー・ソリューションズ[3]

1　はじめに

近年のインターネットの拡大に伴い，異常事態の早期検出や安定したサービス提供・運用の必要性から，インターネットトラフィックの定常的なモニタリングの重要性が増している．トラフィックモニタリングを行う際重要になるのは，トラフィックの特徴抽出のためのパケットの分類手法である．一般的なのはトラフィックをパケット上の情報，例えばIPアドレスやポート番号を用いてフローに分類する方法 [1] であり，広く用いられている．ここで，本論文において，トラフィックを分類することをトラフィックプロファイリング，分類に用いたパケットの特性をプロファイルと呼ぶこととする．上記の例ではIPアドレスやポート番号がプロファイルに相当する．

IPアドレスを用いてトラフィックプロファイリングを行う場合，プロファイル数の削減の必要性や，よりトラフィック量の多いプロファイルを重視する観点から，IPアドレスの階層性に注目してプロファイルの集約を行う手法が知られている [2]．図 1 にこの手法を示す．丸印はプロファイルと対応し，丸の大きさがそのプロファイルのトラフィック量を示している．IPアドレス空間を 2 分木で捉えるとIPアドレスはツリーのリーフに対応する．隣接するリーフノードを集約し，ネットワークプレフィックス形のプロファイルを得ることができる．この例では，トラフィック量の少ないノードを集約している．しかし，この方法はIPアドレス空間を極めて機械的に扱っており，現実世界のネットワーク構成に即した結果が得られるとは限らない．

本論文では，IPアドレスがどのような組織に属しているのかを示す情報を利用してプロファイルを集約し，得られた結果を上記手法の結果と比較・考察する．

2　IPアドレス所属組織情報の利用

IPアドレスは，それ自身からそのIPアドレスを利用している組織の情報を得ることが可能な場合が多い．例えばDNSを利用することで，IPアドレスからそのアドレスが所属する組織の階層的な構造を知ることができる．この情報を利用し，トラフィックプロファイリングを行う際IPアドレスの所属する組織の情報を用いて複数のIPアドレスプロファイルを集約すること

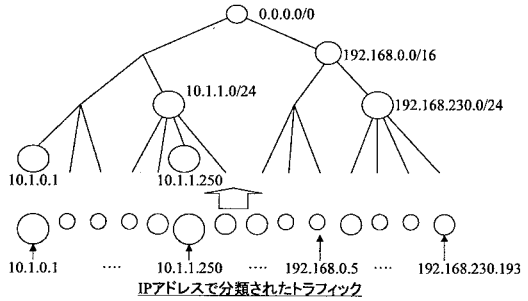

図 1: IPアドレスの階層性に注目したプロファイル集約

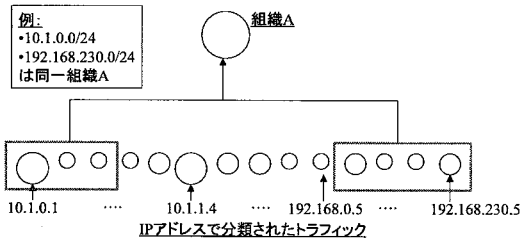

図 2: IPアドレスの所属組織情報を利用したプロファイル集約

で，より現実のインターネットの構成に基づいたトラフィックプロファイリングを行えると考えられる．一般的に各組織は，ネットワークプレフィックスの異なる複数のサブネットワークを運用していると考えられる．つまり，一般的には，ネットワークプレフィックスが近接していても全く別の組織である場合もあり，その逆も考えられる．そのため図 1 で示したプレフィックスツリーに基づくプロファイルの集約結果とは異なる結果が得られることが予想される．

本論文においてはこのことを確かめるため，現実のネットワークのトラフィックデータを用い，上述した 2 つの手法によるトラフィックプロファイリングとプロファイルの集約を実験した．本論文においてIPアドレスと対応する組織の情報として利用したのは，AS(Autonomous System) 情報である．ASは，同一の管理ポリシーで運用されるネットワークの集合体であり，その下に複数のサブネットワークを含んでいる．各ASには 16bit のAS番号が割り当てられ，AS 間

[1]KOIDE Kazuhide,SAITOH Takeo,Glenn Mansfield Keeni, SHIRATORI Norio
[2]Research Institute of Electrical Communication/ Graduate School of Information Sciences, Tohoku Univ.
[3]Cyber Solutions Inc.

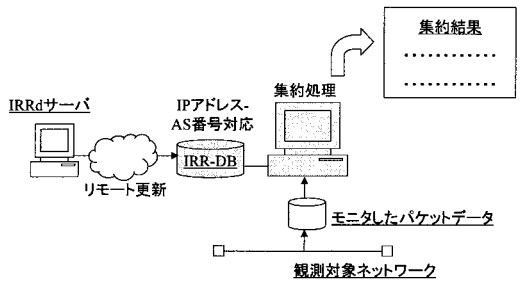

図 3: 実験環境

でのルーティングに用いられている．この AS の情報は IRR(Internet Routing Registry) と呼ばれるデータベースに登録され運用されており，IRR を利用することで IP アドレスから所属 AS 情報を得ることができる．今回はこの IRR を利用し，IP アドレスの所属 AS 情報を取得し，プロファイルの集約に用いる．

3 実トラフィック解析実験

使用したトラフィックデータは本研究室のネットワーク上で観測されたもので，メディアは Ethernet である．研究室内部から研究室外部へルータを越えて流れたトラフィックにのみ注目している．30 分間のトラフィックでパケット数は 9559,Destination IP アドレスの種類は 144 であった．このデータの Destination IP アドレスに注目して，2 種類の実験を行った．実験システムを図 3 に示す．

- 実験 1：Destination IP アドレスで分類後，プレフィックスツリーを用いて集約
- 実験 2：Destination IP アドレスで分類後，所属 AS 情報を用いて集約

実験 1 においては，プレフィックスツリーに基づいたトラフィックプロファイリングを行うツールである AGURI[2] の出力結果を用いた．AGURI は IP アドレスを用いてトラフィックをプロファイリングすることができ，トラフィック量の少ないプレフィックスはプレフィックスツリーを用いて集約する．今回の実験では，全体のトラフィック量に対して 1%未満のトラフィック量のプロファイルを集約することとした．実験 1 の結果を図 4，実験 2 の結果を図 5 に示す．

実験 1 では，最終的に 51 プロファイル (ネットワークプレフィックス形) にプロファイリングされた．一方実験 2 では，53 プロファイル (AS 番号) にプロファイリングされた．ここで注目すべきは，実験 1 で生成された 211.8.0.0/14 というプロファイルである．このネットワークプレフィックスで示されるサブネットワークは現実には存在せず，AS4725(ODN) 所属の 211.8.0.0/16，AS4713(NTT-OCN-AS) 所属の 211.11.0.0/16 等複数のサブネットワークを集約している．実験 2 ではこれらのサブネットワークはそれぞれの所属 AS に集約されている．一方，実験 1 で生成された 61.112.43.211 お

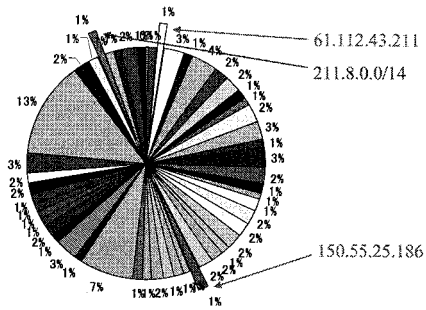

図 4: プレフィックスツリーによる集約結果

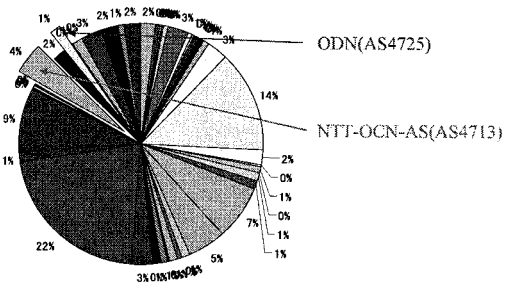

図 5: 所属 AS 情報に基づいた集約結果

よび 150.55.25.186 というプロファイルは，どちらも現実には AS4713(NTT-OCN-AS) という同一 AS 所属であり，実験 2 ではこの 2 つを AS4713(NTT-OCN-AS) という単一のプロファイルに集約している．

4 まとめと今後の展望

本論文では，トラフィックプロファイリングを行う際のプロファイルの集約について，プレフィックスツリーを用いる手法と，現実のネットワーク構成に基づいた手法とを実験し，結果の差異を比較し考察した．今後は，より長期間のデータを処理して得られる時系列的な集約結果について，両手法の結果にどのような差異が存在するのかを比較・考察する予定である．

参考文献

[1] K.C.Claffy,H.-W.Braun,G.C.Polyzos,“A parameterizable methodology for Internet traffic flow profiling,” IEEE Journal on Selected Areas in Communications,1995.

[2] 海崎　良, 長　健二郎,“トラフィックプロファイラ AGURI の設計と実装” WIT2001,2001 年 9 月 5 日．