

インターネット上の通信と経路情報の相関関係に関する研究

原田 義明[†] 岡村 耕二^{††} 千代延隆嗣^{†††}

[†]九州大学 システム情報科学府, 〒812-8581 福岡県東区箱崎 6-10-1

^{††}九州大学 情報基盤センター, 〒812-8581 福岡県東区箱崎 6-10-1

^{†††}九州大学工学部 電気情報工学科, 〒812-8581 福岡県東区箱崎 6-10-1

E-mail: †{harada,chiyonobu}@ale.csce.kyushu-u.ac.jp, ††oka@ec.kyushu-u.ac.jp

あらまし 本論文では九州大学と九州ギガポッププロジェクト(QGPOP)の経路情報及びフローデータを利用し、ASパス長の変化と実際にインターネット上に流れるフローデータとの相関関係についての解析、及びある組織における通信相手の分布の解析を行った。実際にインターネット上に流れる通信データと各種データとの相関関係を解析することで、今後必要になってくるであろう経路や改善すべき経路などを知ることができ、インターネット構築に有用なデータを得ることができる。QGPOPのフローデータに対して、パケット数、フロー数とASパス長の相関関係を解析したところ、年間を通して平均ASパス長には増加傾向が見られた。しかし、解析データを一般的に利用される通信であるウェブアクセスに限定したところ、フロー数から計算した平均ASパス長については減少傾向が見られた。九州大学の通信相手分布解析の結果、通信全体では最大で全ASの約半数との通信のみみられた。しかし、通信データをウェブアクセスに絞って解析を行うと、年間を通じて九州大学の通信の80%はインターネット全体の0.4%の組織を相手とする通信に偏っていることが示された。

キーワード フローデータ, BGP, AS, AS Path,

A Study on Distribution of flows Relationships with AS path length and flows

Yoshiaki HARADA[†], Koji OKAMURA^{††}, and Takashi CHIYONOBU^{†††}

[†] Graduate School of Information Science and Electrical Engineering(ISEE), Kyushu University, 6-10-1
Hakozaki, Higashi-ku, Fukuoka 812-8581 Japan

^{††} Computing and Communications Center,

Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

^{†††} Department of Electrical Engineering and Computer Science (EECS), Kyushu University, 6-10-1
Hakozaki, Higashi-ku, Fukuoka 812-8581 Japan

E-mail: †{harada,chiyonobu}@ale.csce.kyushu-u.ac.jp, ††oka@ec.kyushu-u.ac.jp

Abstract This paper analyzed the relationship between the change of AS-path length and Flow Data, and analyzed the distribution of subject that a certain organization is communicating. It is expected that the route should be improved and will be needed for future by obtaining BGP routing information. We used the collected BGP tables stored in Hard Disk and communication data passed through BGP tabled, and analyzed correlation between AS path length and flow QGPOP which has its own AS. We wanted to find the distribution of BGP neighbors of Kyushu university by analyzing collected Flow Data.

Key words Flow Data, BGP, AS, AS Path

1. ま え が き

近年、急速にインターネットの普及が進み、莫大な数のホストがインターネットに接続されるようになった。インターネットで

は、この莫大な数のホストを管理するためにAS(Autonomous System)という単位を利用している。ASとは、共通のポリシーや同じ管理下で利用されているルータやネットワークの集合のことである。インターネットにおける通信はAS内とAS

間で別々に管理されており、AS 単位に分割して管理することで、ネットワークの構築を容易にしている。現在、AS の数は 2 万を超え、さらに増加傾向にある。隣接していない AS にホスト間が通信を行う場合、他の AS がパケットを中継することで、目的の AS までの通信を行っている。このため、通信を行う AS の位置、つまり、通信経路によって通信速度は影響を受けることになり、ネットワーク距離は経路選択の一つの指標となっている。このネットワーク距離を示す、代表的な指標の一つとして AS パス長というものがある。現在インターネット上の AS 間の経路制御として利用されている BGP (Border Gateway Protocol) というプロトコルでは、AS パス長が短い経路を優先的に選択するようになっている。しかし、AS パス長の短い経路を選択するといっても、AS パス長の長い経路しか存在しなければ冗長な経路しか選択できず、効率のいい経路制御が行われているとは言えない。しかし、ネットワークが発達することで AS 間に新しい接続や、新規 AS の出現によって長い経路が短絡されれば、その経路を通過する通信の AS パス長は短くなる。本論文では、インターネット上の経路情報の変化が、実際にインターネット上に流れる通信に対してどのように影響を及ぼしているのかを解析する。これは今後のインターネット構築において有益な情報となり得る。また、インターネットに接続している全ての組織やネットワークが常に相互通信を行っているわけではない。例えば九州大学では、日本の大学や交流のある外国の大学とは多く通信がなされているだろう。つまり、主要な通信相手の分布には偏りがあると考えられる。そこで、本論文では、実際にインターネット上を流れているフローや経路情報を解析し、AS パス長との相関関係や通信の偏りを時系列解析することで、今後のインターネット構築に有用なデータを収集する。

本論文では、第 2 章において AS 間の距離を表す AS パス長、及びインターネット経路制御についての説明を述べる。第 3 章では、本研究の解析手法を述べ、どのように解析を進めていくのかを説明する。そして、第 4 章では、実際に得られた結果に基づき、AS パス長と他データとの相関関係についての考察を行う。第 5 章で本研究のまとめと今後の課題について述べる。

2. インターネットルーティング

2.1 AS と AS 間の経路制御

現在、インターネットには莫大な数のホストが接続されているが、これを管理するために AS [1] (Autonomous System) という単位が利用されている。AS とは、共通のポリシーや同じ管理下におかれているルータやネットワークの集合のことである。インターネットでは、AS 単位に分割して管理することで、ネットワークの管理を容易にしている。AS には、それぞれの AS を一意に識別するために AS 番号という 16bit の識別子が割り当てられている。AS や IP アドレスは RIR (Region Internet Registry) や NIR (National Internet Registry) などが管理しており、ユニークな AS 番号が保持、運用されている。また、それぞれの機関が管理している地域の AS や IP アドレスの情報については WHOIS データベースを参照することで細かい情報

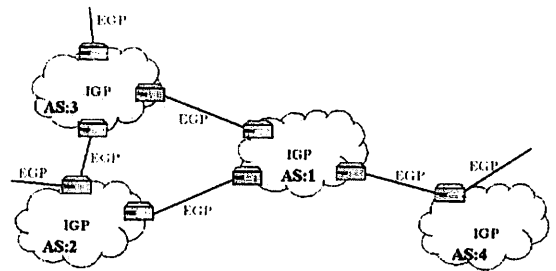


図 1 AS とインターネット上の経路制御
Fig. 1 AS and Internet routing

を得ることができる [2]~[6]。ここで、インターネットを一つの巨大な IP ネットワークと捉えた場合、インターネットは多数の AS の集合体と考えることができる (1)。そこで、AS 内と AS 間の経路制御を分け、階層化することで、ネットワークの耐障害性の向上や、管理の効率化を図っている。インターネットの経路制御は、AS 内における経路制御と AS 間における経路制御の 2 つで大別でき、AS 内の経路制御を行うプロトコルを IGP (Interior Gateway Protocol)、AS 間の経路制御を EGP (Exterior Gateway Protocol) が呼ぶ。IGP には、OSPF (Open Shortest Path First)、RIP (Routing Information Protocol)、Is-Is (Intermediate System-to-Intermediate System) などがある。一方、EGP には、BGP (Border Gateway Protocol) がある。現在では、EGP では、BGP [4] [7], [8] というプロトコルが標準となっている。

BGP は、AS 間の経路情報を交換するためのプロトコルである。BGP の経路情報には、次にパケットを転送すべきルータのアドレスや、宛先に到達するまでに経由した AS 番号のリスト (AS パス) が含まれており、通常 AS パス長の短いものを最短ルートとして使用する。また、BGP にはパス属性と呼ばれるパスに関する情報が含まれており、パス属性を変更することで、より柔軟な経路制御を行うことができる。経路情報には、AS ごとに AS パスが割り当てられているのではなく、プレフィックスという単位ごとに AS パスが割り当てられている。プレフィックスとは、ネットワークアドレスとネットマスクから成り、一群のネットワークアドレスを指す。経路選択の際には、宛先ネットワークアドレスを含むプレフィックスを検索することで、それに対応する AS パスを知ることができる。また、BGP では到達可能な AS への経路情報を隣接ルータで交換しあうことで、到達可能な全てのプレフィックスへの経路情報を作成する。ここで、経路情報での AS パス長とは、宛先に到達するまでに経由した AS 番号の数であり、AS 単位での距離と考えることができる。実際、AS パス長を視覚化する研究など、インターネット上の通信距離に関する多くの研究がなされている [9]~[11]。

3. 解析手法

3.1 フローデータの経路情報の収集

本研究では、フローデータと経路情報を収集し、これらの

srcIP	dstIP	router_sc	prot	srcPort	dstPort	octets	packets
155.230.104.29/16	133.43.121.50/16	0.0.0.0	6	20	26586	130720	85
141.223.1.31/16	150.100.2.3/16	0.0.0.0	17	33424	53	85	1
133.5.9.34/16	140.113.38.243/16	0.0.0.0	1	0	0	56	1
133.243.3.35/16	133.5.207.232/16	0.0.0.0	6	80	1636	40	1
133.5.9.35/16	163.19.173.100/16	0.0.0.0	1	0	0	56	1
207.46.4.38/16	133.5.20.172/16	0.0.0.0	6	1863	50089	48	1

図 2 フローデータの一例

Fig. 2 example of Flow Data

データから AS パス長の情報や必要のデータを抽出し解析を行う。本節では、フローデータと経路情報の収集方法について述べる。

今回の解析では経路情報を QGPOP [12](Kyushu Giga POP Project) の AS である AS2523 のルータから抽出し、フローデータを AS2523 と AS2507 (九州大学) から抽出した。QGPOP とは、日本の様々な研究機関が参加する学術研究用ネットワークである SINET や韓国の学術ネットワークである KOREN と BGP ピアリングを持ち、それぞれの AS と直接通信するとともに、SINET と KOREN のトランジットの役割も果たしている。また、IIJ (Internet Initiative Japan) や APAN-JP (Asia Pacific Advanced Network - Japan) などとの BGP ピアリングを持っている。

フローデータに関しては、QGPOP、九州大学ともに蓄積されているフローデータを使用する。また、全てのフローを収集するのではなく、サンプリングされたデータを使用する。サンプリングレートは 10 分の 1 である。フローデータからは、送信元アドレス、送信先アドレス、プロトコル番号、使用ポート番号やパケット数などが収集できる (図 2)。フローデータは、インターネット通信の遅延を解析するなど多くの研究で利用されている [13]~[15]。

ポート番号とは、インターネットでの TCP/IP による通信において、相手先を特定するために用いられる番号である。IP アドレスから通信相手のホストを特定し、ポート番号からそのホストの利用しているプログラムを特定する。ポート番号は通信の種類毎にそれぞれ 0 から 65535 (16bit) で割り当てられる。本研究では、通信の用途を限定した解析を行う際にこのポート番号を用いる。

次に、経路情報の収集について述べる。本論分では BGP の経路情報を取得する。経路情報からは、到達可能な全てのホストに対応するプレフィックスとそれに対応する AS パスを取得することができる。AS パスは、宛先に到達するまでに通過する AS 番号を順に一つずつ並べられた数字の列で示される (図 3)。この AS パスの列を数えることで AS パス長を得ることができる。また、終点の AS から、通信先 (元) AS を知ることもできる。ここで、AS パス長を計算するに際し、経路情報には同一 AS パスが連続で記述されている場合がある。これは AS-path prepend と言い、ルータが自分のパスにダミーのパスを追加し、自ルータを通る AS パス長が長いと見せかけるものである。BGP では、AS パス長が短い経路を優先的に選択していくので、自ルータに流れるフローを減らすことができる。AS-path prepend は、あまりフローを流したくない場合などに利用されている。AS-path prepend を含む AS パス長

Network	Next Hop	Metric	LocPrf	Weight	Path
*13.0.0.0	210.138.15.145	300	0	0	2497 2497 701 703 80 1
*14.0.0.0	210.138.15.145	300	0	0	2497 2497 3356 1
*14.23.112.0/22	210.138.15.145	300	0	0	2497 2497 174 21889 1
*14.23.180.0/24	210.138.15.145	300	0	0	2497 2497 701 6123 30576 1
*14.36.116.0/23	210.138.15.145	300	0	0	2497 2497 174 21889 1
*14.36.118.0/24	210.138.15.145	300	0	0	2497 2497 174 21889 1

図 3 BGP 経路情報の一例

Fig. 3 example of BGP table

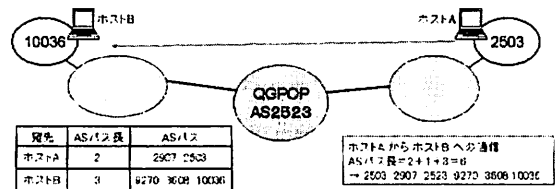


図 4 AS パスの計算方法

Fig. 4 method of calculating AS path length

は長く見えるが、実際に通過している AS を考えると、連続する AS を一つの AS と数えなくてはならない。また、本研究では、IIJ (AS2497) の AS-path prepend を減らして解析を行う。IIJ とは QGPOP に隣接している AS で、2005 年 8 月 17 日に AS-path prepend を追加している。IIJ は QGPOP との通信頻度も高く、QGPOP の経路情報上の AS パスへの影響も大きい。この影響を無視して解析を行うと年間を通した AS パス長の解析が有用でなくなってしまうので、今回の解析では IIJ の AS-path prepend を短縮し解析を行う。

3.2 フローデータの解析手法

3.2.1 AS パス長とフローデータの相関関係の解析

QGPOP から取得した経路情報とフローデータを解析することで、AS パス長と他データとの相関関係に関するデータを解析する。QGPOP の蓄積されているフローデータを、2004 年 10 月から 2005 年 12 月まで、4 週ごとの水曜日のデータを解析する。これは、曜日によってデータの傾向に差がでる可能性があるため、特定の曜日のデータを取ることが望ましいと考えられるためである。また、一日分のデータについては、それぞれ毎時 0 分~5 分のデータを 24 個取得する。

まず、AS パス長の計算方法について述べる。本研究では QGPOP の経路情報を利用するため、各フローに関しては、QGPOP から宛先までの AS パスしか取得できない。フローの往路と復路の AS パスが同じとは言い切れないが、基本的には最短 AS パスを経路として選択するため、AS パス長の誤差は少ないと考えられる。そこで、QGPOP からの AS パスの復路と往路が同じであると仮定して QGPOP から送信元アドレスへの AS パスを逆にたどり、次に QGPOP の AS 番号 2523 をはさみ、QGPOP から送信先アドレスまでの AS パスを足し、これを送信元から送信先への AS パスと見なし、解析を行う。AS パス長の計算の一例を図 4 に示す。

このようにして AS パス長を計算し、フロー、パケットと AS パス長との相関関係を調査する。また、全てのフローデータを解析した際に、不正アクセスなどによるフローが大量に流れたとすると、本来の目的で使われているフローの解析という

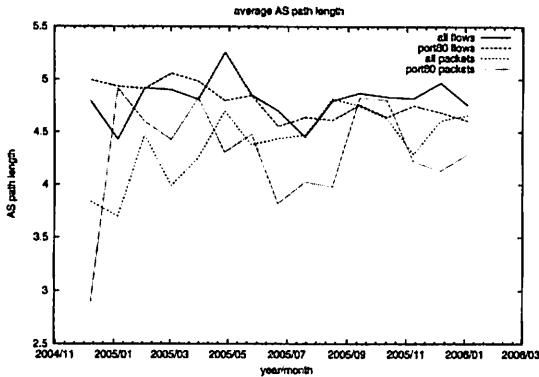


図5 年間を通した平均 AS パス長の時間変化
Fig.5 Time change of average AS path length

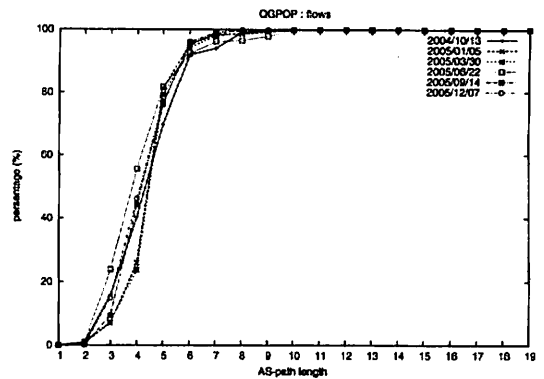


図6 AS パス長に基づいた、フローの占有率の分布
Fig.6 Distribution of port 80 flows and AS path length

意味にはならない。そこで、本研究では、全フローについての解析と同時に、日常的に使用されるプロトコルであるポート 80(HTTP)に照準を絞り、データの傾向及び AS パス長との相関関係の解析を行う。

3.2.2 通信相手の分布の解析

九州大学の蓄積フローデータと QGPOP の蓄積 BGP テーブルを用いて、フロー、パケット、通信 AS 数について通信相手分布の解析を行う。九州大学のフローデータは、毎週水曜日のデータを取得する。また、九州大学の BGP テーブルはある理由により取得できないため、QGPOP の BGP テーブルを用いて、通信相手の AS を参照する。

また、QGPOP のフローデータ解析と同様に、一般的に利用される通信として通信をポート 80 に限定しての解析も行う。

4. 解析結果

4.1 フローデータと AS パス長の相関関係

全通信とポート 80 の通信においてパケット数・フロー数から平均 AS パス長を計算し、時間変化を表したものが図 5 である。全体的にフローよりもパケットの平均 AS パス長の方が小さいことが多く、トラフィックを少なくするように経路制御されていることがわかる。また、全フローに対しての平均 AS パス長には増加傾向が見られ、ポート 80 のフローについては平均 AS パス長に減少傾向が見られる。パケットに関しては日よりの増減が大きく、全通信・ポート 80 の通信のどちらにおいても際立った傾向も見られなかったため、AS パス長と時間変化との相関関係は得られなかった。

次に、平均 AS パス長に減少傾向が見られたポート 80 のフローについて、AS パス長ごとの占有率を示したグラフを図 6 に表す。図 5 では、フローの平均 AS パス長が年間を通して減少していた。それを示すように、新しい月の AS パス長の分布のグラフは、左側にあることが多い。これは、月日が経つにつれ短い AS パス長を持つ通信の数が増え、グラフの立ち上がりが早くなった結果だと言える。また、AS パス長 5 で 70%、AS パス長 6 で 90% ほどのフローが流れており、短い AS パス長を持つ AS 間で多くの通信がなされていることがわかる。

ポート 80 のフローにおいて、AS パス長は年間を通して減少傾向にあり、時間変化とともに短い AS パス長を持つフローが増えていることがわかった。そこで、年間を通して通信を行っている AS の組やフローの多かった AS の組について細かく調査し、AS パス長の平均が減少した理由についての考察を行った。その結果、AS パス長の平均が減少した要因と思われる通信をいくつか発見した。その一つが Microsoft の AS である。Microsoft には AS8068 から AS8075 の AS を自由に使えるように割り当てられていて、年間を通してこれらの AS からは非常に多くの通信がされている。AS12076 (hotmail) もまた Microsoft の AS である。この Microsoft の AS、AS12076 と AS8072 についての AS パスの変化を見てみたところ、5月に AS12076 の一部と AS8072 が AS8075 に統合されていた。AS12076 と AS8072 は、AS8075 を経由して通信されていたため、これらが一つの AS にまとまったことで、AS パス長が 1 減少する結果となった。Microsoft の通信を見てみると、AS2508 (九州大学) との通信が多く、AS2508 (九州大学) は AS2523 (QGPOP) とは AS 間が直接つながっているため、総 AS パス長も短くなる。AS12076 と AS8072 が AS8075 に統合された場合、AS パス長が 4 と他の通信よりも短くなり、平均 AS パス長が減少したのだと考えられる。もう一つ AS パス長の平均が減少した理由として挙げられるのが、AS18088 (QIC 株式会社キューデンインフォコム) の通の増加である。2004 年 10 月から存在はしているのだが、通信自体は少なかった。しかし、徐々に通信が増え、2005 年 6 月頃からは、最も多く通信を行う AS の一つとなっている。AS18088 (QIC) はほとんどが AS2508 (九州大学) との通信で、AS パス長が 3 と短い。短い AS パス長を持つ通信が増加したことで、平均 AS パス長も減少したと考えられる。

4.2 ある組織の通信相手分布

九州大学のフローデータについてフロー数、パケット数、通信相手 AS の個数の年間の変化を図 7 に示す。

図 7 から、相手 AS の個数が特に少ないときはフロー数やパケット数も少ない傾向にあることが分かる。また、休日である 2004 年 11 月 3 日、2005 年 5 月 4 日、2005 年 11 月 23 日は前

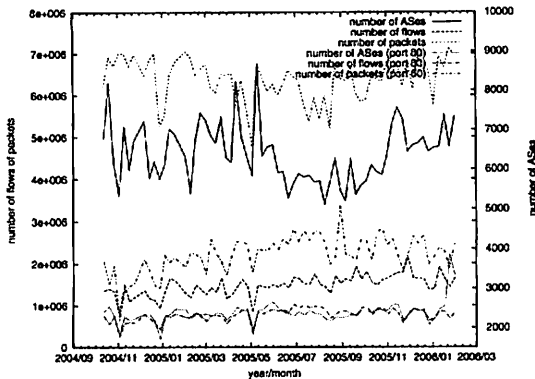


図7 フロー数、パケット数、通信相手 AS 数の時間変化

Fig.7 Time change of flows, packets, and communication target

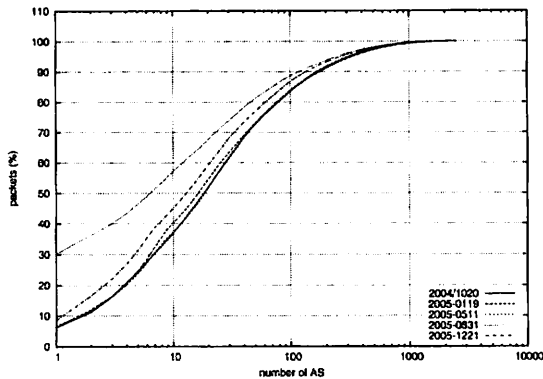


図8 通信相手によるパケット数の偏り (ポート 80)
Fig. 8 Distribution of communication (log scale)

後と比較してフロー数やパケット数が少ない傾向にある。その他に 2004 年 12 月 29 日や 2005 年 12 月 28 日、2006 年 1 月 4 日についても同様の傾向が見られる。また、夏期休業のためか 7 月から 9 月の間は通信相手 AS が少ない傾向にある。現在 AS は約 20,000 個存在しており、多いときでは全体の約半分と通信を行っていることになる。しかし、ポート 80 の通信においては、常に 2,200 個から 2,500 個程度の相手 AS と通信が行われており、通信分布にはある程度の傾向があると考えられる。また、ポート 80 のフロー数・パケット数は時間経過とともに増加している。これは、九州大学においてインターネット接続の機会が多くなっていると考えられる。

ポート 80 に関する AS の頻度分布のグラフを図 8 に示す。縦軸は通信パケット数が多い順に AS をソートし、上位 AS からパケット数を足していったものをパーセントで示しており、横軸が AS 数を示している。2005 年 8 月 31 日のグラフにおいて、他の月日と別の傾向が見られる。ひとつの AS との通信で約 30%パケットが使われているが、これは AS2504 の京都大学のネットワークであった。このグラフの影響を除外すると、どの日のデータも通信量の多い 20 個程度の AS が全通信量の 50%を

	2004		2005	
	のみ	重複	のみ	重複
flow	23,519	548,140	58,776	875,022
flow(%)	4.1	95.9	6.3	93.7
packet	61,509	1,276,678	226,662	2,377,060
packet(%)	4.6	95.4	8.7	91.3

表 1 2004 年 10 月と 2005 年 10 月の通信相手の詳細

Table 1 Detail of communication target compared between 2004/10 and 2005/10

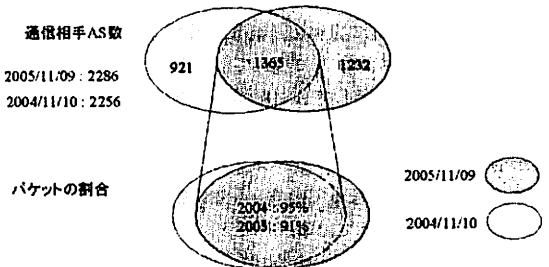


図9 通信相手によるパケット数の偏りと通信相手 AS 数の変化
Fig. 9 Changes of communication targets in port 80 flow data

占めている。上位 80 個 (インターネット全体の 0.4%) の AS で、全通信量の 80%を占めており、通信傾向に偏りがあることがわかる。通信の内約を見ても、常に通信パケット数の上位にある AS として AS23816 (Yahoo! Japan AS) や AS9370 (さくらインターネット), AS20940 (Akamai) などの商用 ISP が多く見られた。

次に、通信を行っている AS の内訳について詳細な調査を行う。例えば新規に多くの通信をするようになった AS があり、そのトラフィックが大きいのであれば、ネットワーク管理の際に考慮すべき問題となり得るだろう。そこで、1 年間隔でみたときに相手 AS の分布が変化しているのかを調査する。2004 年 11 月 10 日と 2005 年 11 月 9 日を対象とし、両者の通信相手 AS の内訳を調査した。2004 年 11 月と 2005 年 11 月に共通して通信のあった相手 AS は 1,365 個、2004 年のみ通信のあった AS が 921 個、2005 年のみ通信のあった AS が 1,232 個という結果であり、半数程度の AS が入れ替わっている。また、これらの AS 群の集合それぞれの通信量を表 1 に記す。

図 9 をみると、重複している AS 群の通信量が多いことが分かる。2004 年と 2005 年ともに通信のある AS 群が通信の 90%以上を占めている。全体としては、常に通信のある AS 群が多くの通信をしているといえる。これは予想を裏付ける結果となっている。2004 年のみ通信のあった AS 群は 2004 年全パケットの 4.6%を占めるにとどまっているが、2005 年新たに通信するようになった AS 群の全パケットに占める割合は 8.7%と大きくなっている。新たに通信するようになった AS 群は、通信量に占める割合が大きくなってきているといえる。

次に、BGP 経路表を参照し、AS が存在するかどうかを調査した。2004 年のみ通信のあった AS 群 921 個のうち、2005 年 11 月 9 日の経路表から消えているものは 78 個であった。こ

れらは、AS 自体が無くなっていると考えられる。残りの 843 個は、経路表には残っているので、AS 自体が無くなったのではなく通信が無くなったものと考えられる。同様に 2005 年のみ通信のあった AS 群 1,232 個に関して 2004 年 11 月 10 日の経路表を参照すると、経路表に無いものが 211 個あった。これらはこの 1 年の間に新規にできた AS である。残りの 1,021 個は経路表自体には存在していたので、もともと通信が無かった AS 群であると考えられる。以上から、実際に消えたり新規にできた AS の個数は相対的に少なく、多くはもともと存在する AS であるといえる。

さらに、各 AS 群に関する詳細な解析を行う。まず 2005 年のみに通信があった AS 群について解析を行う。実際に AS 自体が無くなったものとして AS10007 (I P Revolution, inc) があげられる。AS10007 とはほとんどいつも 10,000 パケット以上の通信があったが、ソフトバンク BB に統合され AS が消えた 7 月で通信が無くなっている。AS8072 (Microsoft AS) も 4 月に通信が無くなっている。

次に、2005 年のみに通信があった AS 群について詳細な解析を行う。通信量が増えた AS として、AS8075 (Microsoft AS) の増加が目立っている。そこで BGP 経路表を参照したところ、2004 年 11 月に AS8072 に属していたプレフィックスの多くが、AS8075 に変更されていることが分かった。また、AS5572 (BOTIC AS) や AS30968 (Info Box) はともにロシアの組織だが、2005 年 11 月に通信が偏っている。一時的にはあるが、ロシアとの通信が増えているので、二つの AS の増加には何らかの関連性があると考えられる。

続いて、2004 年と 2005 年に重複して通信のあった AS 群に関する解析を行う。ポート番号 80 番全体でみて常にパケット数の上位にある AS として、AS23816 (Yahoo! Japan AS) や AS9370 (さくらインターネット)、AS20940 (Akamai) などの商用 ISP が多く見られた。また、これらの AS は年間を通して常に上位 20 個の中に含まれていた。ポート番号 80 番では、商用 ISP との通信が特に多いといえる。また、この他に上位に挙がっている AS としては、AS2504 (京都大学) がある。

5. おわりに

本研究では、実際にインターネット上に流れているフローデータの解析を行った。QGPOP を流れている全フローに関して、フロー、パケット数と AS パス長との相関関係を調査したところ、どちらのデータに関しても AS パス長の短い通信が多いという結果が得られたが、年間を通して AS パス長との相関関係を解析すると、平均 AS パス長には増加傾向があるという結果が得られた。しかし、不正アクセスなどの影響が少なく、一般的に使用される通信であるポート 80 のフローのみに対して解析を行ったところ、フロー数に関して AS パス長に減少傾向が見られた。また、AS パス長が減少した原因を解析するために細かい AS パスの打ち分けを調査してみたところ、AS パス長の短い通信のフロー数が増加し、またフロー数の多い通信の AS パス長が減少することで、全体的に AS パス長の短い通信が増えていた。インターネットの発達により AS 間の経路が

増え、また AS が統合されることで、AS パス長が減っていた。これより、実用的な通信に関しては AS パス長は減少傾向にあるという結果が得られた。九州大学のフローデータ解析の結果、通信全体では最大で全 AS の約半数との通信がみられた。しかし、不正な通信の影響を除外し、一般的なインターネット利用であるウェブアクセスなどに絞った解析を行うと、通信相手 AS に偏りがあることが分かった。年間を通じて九州大学の通信の 80% はインターネット全体の 0.4% の組織を相手とする通信に偏っていることが示された。通信相手 AS の内訳には年間で流動があったが、通信量の多い AS はいつも共通しており、商用 ISP や京都大学などとの通信が多く、通信量の 50% を常に上位 20 個の AS が占めていることが分かった。一方、ロシアなどの新たに通信が増加している AS もみられた。

本研究では、ポート 80 の解析において、フローと AS パス長との相関関係は解析できたものの、パケットと AS パス長との相関関係を見出すことはできなかった。より短い間隔でフローデータの解析を行い、パケット数についても相関関係を見出す。通信傾向の解析では通信相手として AS 単位で分布をまとめていったが、これを国や地域ごとで整理すると、それぞれの通信分布に新たな傾向がみられると考えられる。フローデータと AS パス長の相関関係を見出すことは容易なことではないが、今後のインターネット構築に有益な情報を得るために更に研究を深める必要があるだろう。

文 献

- [1] RFC 1772 - BGP-4 Application, <http://www.ietf.org/>.
- [2] American Registry for Internet Numbers (ARIN), <http://www.arin.net/>.
- [3] Resource IP Europeans Network Coordination Centre (RIPE-NCC), <http://www.ripe.net/>.
- [4] Asia Pacific Network Information Centre (APNIC), <http://www.apnic.net/>.
- [5] Japan Network Information Center (JPNIC), <http://www.nic.ad.jp/>.
- [6] Korea Network Information Center (KRNIC), <http://whois.nic.or.kr/>.
- [7] RFC 1771 - A Border Gateway Protocol 4, <http://www.ietf.org/>.
- [8] Stephen A. Thomas : "IP Switching and Routing Essentials", WILEY, pp.181-219, December.2001.
- [9] Huffaker Bradley, Marina Fomenkov, Daniel J. Plummer, David Moore and K claffy, "Distance Metric in the Internet", IEEE International Telecommunications Symposium (ITS), 2002.
- [10] L. Amini and H. Schulzrinne, "Observations from router-level internet trace", in DIMACS Workshop on Internet and WWW Measurement, mapping and Modeling, 2002.
- [11] Toshiyuki KAWASAKI and Koji OKAMURA, "Evaluation on Scalability of Conference System using Request-Routing", 18th International Conference on Advanced Information Networking And Applications, p533-534, 2004.
- [12] Kyushu GigaPOP Project (QGPOP), <http://www.qgpop.net/>.
- [13] C. Estan and G. Varghese, New Directions in traffic measurement and accounting, ACM SIGCOMM '02, 2002.
- [14] Y. Zhang et al. On the characteristics and origins of Internet flow rates, ACM SIGCOMM '02, 2002.
- [15] Connie Logg, Les Cottrell, "Passive Performance Monitoring and Traffic Characteristics on the SLAC Internet Border", PAM2001 A workshop on Passive and Active Measurements, 2001