

固有値解析を用いた AS 間接続関係の数値評価手法の提案

浅井 大史[†] 江崎 浩[†]

[†] 東京大学大学院情報理工学系研究科 〒 113-8656 東京都文京区本郷 7-3-1

あらまし インターネットは、インターネットサービスプロバイダ (ISP) や企業、大学等の異なる管理ドメインによって運用されている数万の自律システム (AS: Autonomous System) が相互に接続することで構築されている。隣接する各 AS 間の接続関係は、1) トランジット、2) ピアリング、3) シブリングの 3 種類に分類される。近年、ピア・ツー・ピア (P2P) 型のコンテンツ配信ネットワーク (CDN) システムにおけるネットワーク資源の効率的利用を目的として、これらの AS 間接続関係をオーバーレイネットワークの経路制御メトリックとして利用した研究が行われている。しかし、AS 間接続関係は一般には公開されていない情報であるため、これらの研究において、実ネットワークへの応用方法が課題となっており、公開可能な情報から AS 間接続関係を数値評価することが必要とされている。本稿では、AS 間の隣接関係行列に対して固有値解析を利用することで AS 間接続関係を数値評価する手法を提案する。

キーワード AS 間接続関係の数値評価、固有値解析

AS Relationships Estimation by Use of Eigenvalue Analysis

Hirochika ASAI[†] and Hiroshi ESAKI[†]

[†] Graduate School of Information Science and Technology, The University of Tokyo Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656 Japan

Abstract The Internet consists of thousands of Autonomous Systems (ASs) operated by different administrative domains such as Internet service providers (ISPs), companies and universities. The relationships among ASs can be categorized into three groups; 1) *transit*, 2) *peering* and 3) *sibling relationship*. In recent studies, these relationships are used as routing metric of peer-to-peer content delivery network systems in order to achieve efficient resource utilization. However, the relationships are generally non-disclosure for most of commercial ISPs. It is necessary to estimate the relationships for these studies. In this report, we propose an analytic way to estimate the AS relationships by use of eigenvalue analysis.

Key words AS relationships estimation, eigenvalue analysis

1. はじめに

インターネットは、インターネットサービスプロバイダ (ISP) や企業、大学等の異なる管理ドメインによって運用されている数万の自律システム (AS: Autonomous System) が相互に接続することで構築されており、各 AS 間の接続関係は、1) トランジット、2) ピアリング、3) シブリングの 3 種類に分類される。近年、ピア・ツー・ピア (P2P) 型のコンテンツ配信ネットワーク (CDN) システムにおけるネットワーク資源の効率的利用を目的として、AS 間接続関係をオーバーレイネットワークの経路制御メトリックとして利用した研究が行われている [1], [2]。我々は、AS 間接続関係を経路制御メトリックとして用いた経路制御アーキテクチャを提案し、ネットワークドメイン間の高コストなリンクを通過するトラフィックを削減できることを示した [2]。しかし、AS 間接続関係は一般には公開されていない情

報であるため、これらの研究において、実ネットワークへの応用方法が課題となっており、公開可能な情報から AS 間接続関係を数値評価することが必要とされている。

各 AS 間接続関係における経路交換ポリシーの非対称性より、インターネット上の経路が Valley-Free 構造 [3] となることを利用して、複数の AS における経路表 (AS PATH) から探索的に AS 間接続関係を推定する手法 [4], [5] は数多く研究されてきた。しかし、実際に運用されているインターネットにおける AS 間接続関係は、接続料や経路交換ポリシーの点で均一ではなく、経路交換の非対称性だけから AS 間接続関係を数値評価することは困難であった。

AS 間接続関係を数値評価する手法として、AS の規模を表す隣接 AS の数の差を用いるが手法 [6] は広く利用されているが、この手法により評価された数値にはピアリングとトランジットの違いが反映されていない。

本稿では、経路表を用いずにトランジットおよびピアリングを区別しながらインターネット全体における2AS間の規模の差を表現することを目的として、AS間の隣接関係行列に対して固有値解析を利用することでAS間接続関係を数値評価する手法を提案する。

2. ドメイン間ルーティング

インターネットは数万のASから構成されている。それぞれのASはISPや企業、大学等の異なった管理ドメインにより運用されている。AS間の経路情報は、Border Gateway Protocol (BGP) に代表されるドメイン間ルーティングプロトコルにより交換・共有・構築される [3].

2.1 AS間接続関係

AS間接続関係は、以下の3つに分類でき [4], [7], それぞれの接続関係ごとに経路交換ポリシーが存在する。この経路交換ポリシーにより、通常、インターネット上の経路はValley-Free構造となる [3].

(1) トランジット (Transit): 大規模なネットワークを運用するネットワークプロバイダ (プロバイダAS) とそれより小規模なネットワークを運用するネットワークプロバイダ (カスタマAS) の接続形態 (provider-to-customer/p-to-c 関係, customer-to-provider/c-to-p 関係)。カスタマASは、プロバイダASからグローバルなインターネットへの接続性を利用帯域に応じて購入する。一般的に、プロバイダASは全経路を広告し、カスタマASは自分のカスタマASの経路のみを広告する。

(2) ピアリング (Peering): 同規模のネットワークを運用するネットワークプロバイダ同士の対等な関係での接続形態 (peer-to-peer/p-to-p 関係)。一般的に、ピアASは、カスタマASの経路のみを広告し、プロバイダASや他のピアASの経路は広告しない。ピアリングはトラフィックの交換量で対等な立場となるため、これらの間ではトラフィックの交換を自由に行うことができる (無償での相互接続)。しかし、交換されるトラフィック量が不均衡になった場合、一方のASが他方のASに接続料を要求する場合もある (paid peer)。

(3) シブリング (Sibling): 同一組織に属するネットワークプロバイダによって運用されるネットワーク同士の接続形態 (sibling-to-sibling/s-to-s 関係)。一般的に、シブリングASは、カスタマASの経路だけでなく、プロバイダASやピアASの経路も広告する。

2.2 経路解析に基づくAS間接続関係の推定

AS間接続関係は、AS間の規模を表す隣接AS数の差を解析することにより推定できる [4]。各AS間接続関係の推定手法は、以下の3つに分類される。

(1) 隣接AS数による推定: 隣接AS数はASの規模を表しており、一般に、トランジット関係においては、カスタマASの隣接AS数よりもプロバイダASの隣接AS数の方が大きくなる。このことより、トランジット関係におけるプロバイダ・カスタマの関係性を推定することができる [5], [6].

(2) 経路解析による推定: インターネット上の経路が Valley-

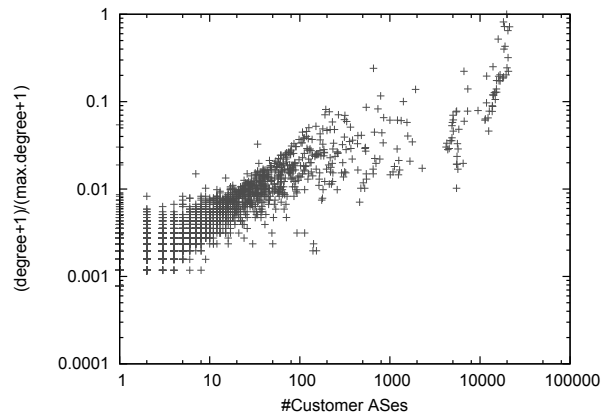


図1 カスタマの数と隣接AS数の相関関係

Free構造となることから、複数のASにおいて取得した経路情報 (AS PATH) を解析することでピアリング関係およびトランジット関係を推定することができる [4], [8].

(3) 組織情報からの推定: シブリング関係については、IRR (Internet Routing Registry) データベース等の公開されている登録情報から類推することができる [9].

しかし、これらの手法において、AS間接続関係の数値評価は行われていない。次節にて、AS間接続関係の数値評価手法を提案する。

3. AS間接続関係の数値評価

AS間接続関係の数値評価手法として、従来、トランジットにおけるプロバイダ・カスタマ関係を推定するために用いられてきた隣接AS数による数値評価手法と本稿の提案手法である固有方程式の解ベクトルによる数値評価手法について説明を行う。

3.1 隣接AS数による数値評価

AS間接続関係は、隣接AS数から推定することができる [5], [6]. “The CAIDA AS Relationships Dataset (09/06/2008)” [10] を用いて計算した、カスタマ (カスタマのカスタマを含む) の数と隣接AS数の相関関係を図1に示す。これより、カスタマの数と隣接AS数には正の相関があり (相関係数: 0.78), 隣接AS数がASの規模を示していることがわかる。つまり、2AS間の隣接AS数の差は、ASの規模の差を数値的に表現していると言える。

式 (1) により、正規化した隣接AS数ベクトルを $rank^{(deg)}$ と定義する。 \mathbf{d} は隣接AS数ベクトル (要素 d_i はAS i における隣接AS数) を表す。また、隣接するAS $i-j$ 間の $rank^{(deg)}$ の差 $\Delta rank_{ij}^{(deg)}$ を式 (2) で定義する。

$$rank^{(deg)} := \log_{10} \left(\frac{\mathbf{d} + 1}{\max(\mathbf{d}) + 1} \right) \quad (1)$$

$$\Delta rank_{ij}^{(deg)} := rank_i^{(deg)} - rank_j^{(deg)} \quad (2)$$

3.2 固有方程式の解ベクトルによる数値評価

隣接AS数による数値評価手法では、隣接ASとの直接的な関係から数値評価を行った。しかし、隣接AS数のような局所的な関係性だけでは、ASの規模を適切に表すことはできてお

らず、ピアリングの判別が困難であると考えられる。本稿では、固有方程式の解ベクトルを AS の規模として表すことで、隣接 AS との直接的な関係性だけでなく、インターネット全体における AS の規模を考慮した数値評価手法を提案する。

AS i から隣接 AS j へのトラヒックが、AS j の規模に比例すると仮定し、トラヒックの分布を求めることにより、インターネット全体から見た AS の規模を再帰的に求める。初期状態の AS の規模として隣接 AS 数を用いる、

まず、式 (3) により、AS 間隣接関係行列 A を定義する。

$$A = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} \quad (3)$$

式 (3) における AS 間隣接関係行列 A の各要素は次のように定義する。

(i) $n = 1$ のとき

$$a_{ij} = \begin{cases} j \text{ の次数} : (i, j) \text{ に接続関係がある場合} \\ 0 : (i, j) \text{ に接続関係がない場合} \end{cases}$$

(ii) $n > 1 (n \in N)$ のとき

$rank^{(eig(n-1))}$ から $rank^{(eig(n))}$ を再帰的に定義する。

$$a_{ij} = \begin{cases} 10^{rank_j^{(eig(n-1))}} : (i, j) \text{ に接続関係がある場合} \\ 0 : (i, j) \text{ に接続関係がない場合} \end{cases}$$

式 (4) により、AS 間隣接関係行列 A を各行で正規化した行列 B を定義する。

$$B = \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \quad (4)$$

$$b_{ij} = \frac{a_{ij}}{\sum_k a_{ik}}$$

B の最大固有値に対応する左固有ベクトルを求め、式 (5) により、AS 規模ベクトル $rank^{(eig(n))}$ を定義する、

$$\begin{aligned} \mathbf{x}B &= \lambda \mathbf{x} \\ rank^{(eig(n))} &:= \log_{10}(\mathbf{x}) \\ (\mathbf{x} : \text{corresponding to the maximum } \lambda) \end{aligned} \quad (5)$$

$rank^{(eig(n))}$ は、AS i から AS j へのトラヒックを b_{ij} としたときの各 AS のトラヒックの流量分布の定常状態を表している。

隣接する AS i - j 間の $rank^{(eig(n))}$ の差 $\Delta rank_{ij}^{(eig(n))}$ を式 (6) で定義する。

$$\Delta rank_{ij}^{(eig(n))} := rank_i^{(eig(n))} - rank_j^{(eig(n))} \quad (6)$$

4. 評価

本稿では、CAIDA [11] により配布されている “The CAIDA AS Relationships Dataset (09/06/2008)” [10] を用いて提案手法の評価を行った。

図 2 に各接続関係における $\Delta rank^{(deg)}$ および $\Delta rank^{(eig(n))} (n \in \{1, 2\})$ の分布を示す。なお、図 2 は、各 AS 間接続関係ごとに分布図の面積が 1 になるように正規化した。

図 2(a) より、隣接 AS 数を用いた AS 間接続関係の数値評価手法を用いることで、トランジットにおけるプロバイダ・カスタマ (p-to-c) 関係とカスタマ・プロバイダ (c-to-p) 関係の違いを数値表現できていることがわかる。しかし、ピアリングの $\Delta rank^{(deg)}$ が広い範囲に分布しており、隣接 AS 数を AS 間の規模の差として用いた場合、ピアリングの判別が困難であることがわかる。

図 2(b) より、固有方程式の解ベクトルを用いた AS 間接続関係の数値評価手法を用いることで、 $|\Delta rank^{(eig(n))}|$ はトランジットでは大きくなっている一方、ピアリングでは小さくなっていることがわかる。この特性により、トランジットとピアリングの区別を行いながら AS 間接続関係の数値評価を行うことが可能であることが示される。

4.1 ピアリングの判別

隣接 AS 数を用いた AS 間接続関係の数値評価手法および提案手法である固有方程式の解ベクトルを用いた AS 間接続関係の数値評価手法における、ピアリングの判別についての評価を行う。

トランジットは規模の異なる AS 間の接続であり、ピアリングは同程度の規模の AS 間の接続であるため、AS の規模の差である $\Delta rank$ を評価することでトランジットとピアリングを判別することが可能であると考えられる。 $\Delta rank^{(deg)}$ および $\Delta rank^{(eig(n))} (n \in \{1, 2\})$ により、式 (7)、式 (8) の閾値を変化させ、各 AS 間接続関係の判別を行い、評価する。

$$\begin{cases} \Delta rank^{(deg)} > th_{deg} : \text{p-to-c} \\ \Delta rank^{(deg)} < -th_{deg} : \text{c-to-p} \\ -th_{deg} \leq \Delta rank^{(deg)} \leq th_{deg} : \text{p-to-p} \end{cases} \quad (7)$$

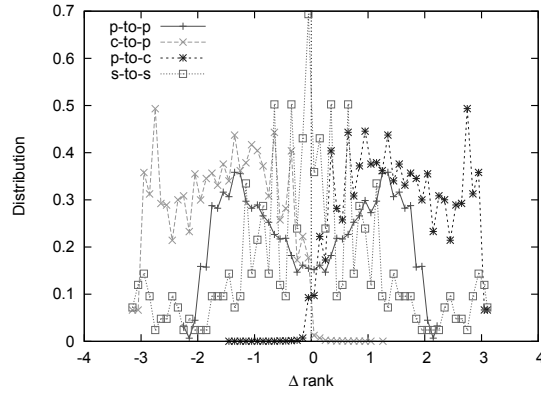
(th_{deg} : 閾値)

$$\begin{cases} \Delta rank^{(eig(n))} > th_{eig(n)} : \text{p-to-c} \\ \Delta rank^{(eig(n))} < -th_{eig(n)} : \text{c-to-p} \\ -th_{eig(n)} \leq \Delta rank^{(eig(n))} \leq th_{eig(n)} : \text{p-to-p} \end{cases} \quad (8)$$

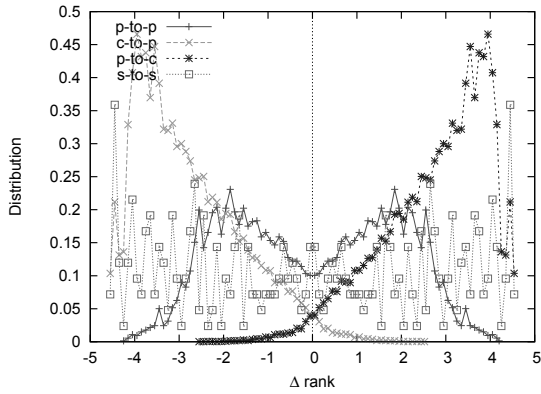
($th_{eig(n)}$: 閾値)

表 1 に各 AS 間接続関係判別時における ROC^(註1) 曲線下面積を示す。また、図 3 にピアリング判別時における ROC 曲線を示す。ROC 曲線における *false positive rate* (*fpr*) および *true*

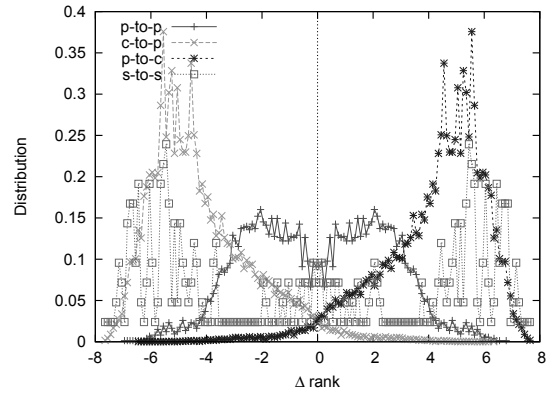
(註1) : receiver operating characteristic (受信者動作特性)



(a) 隣接 AS 数による数値評価 ($\Delta rank^{(deg)}$)



(b) 固有方程式の解ベクトルによる数値評価 ($n = 1, \Delta rank^{(eig(1))}$)



(c) 固有方程式の解ベクトルによる数値評価 ($n = 2, \Delta rank^{(eig(2))}$)

図 2 各 AS 間接続関係における隣接 2AS 間の rank の差 ($\Delta rank$) の分布

Fig. 2 Distribution of $\Delta rank$ for each relationship

表 2 評価結果: 最適閾値と false positive rate および true positive rate

Table 2 Result: optimal threshold and false positive rate / true positive rate

| Threshold | fpr, tpr (p-to-p) | fpr, tpr (c-to-p/p-to-c) |
|----------------------|-------------------|--------------------------|
| $th_{deg} = 1.86$ | 0.628, 0.945 | 0.00392, 0.373 |
| $th_{eig(1)} = 2.70$ | 0.402, 0.879 | 0.00897, 0.599 |
| $th_{eig(2)} = 3.62$ | 0.304, 0.883 | 0.0115, 0.694 |

表 1 評価結果: ROC 曲線下面積

Table 1 Result: area under ROC curve

| | p-to-p | c-to-p/p-to-c | s-to-s |
|-------------------|--------|---------------|--------|
| $rank^{(deg)}$ | 0.653 | 0.980 | 0.680 |
| $rank^{(eig(1))}$ | 0.783 | 0.983 | 0.541 |
| $rank^{(eig(2))}$ | 0.834 | 0.981 | 0.420 |

positive rate (tpr) は、それぞれ式 (9), 式 (10) で表される。ここで, fp , fn は、それぞれ false positive, false negative の総数を示している。また, ni , pi は、それぞれ negative instance, positive instance の総数を示している。

$$fpr = \frac{fp}{ni} \quad (9)$$

$$tpr = \frac{tp}{pi} \quad (10)$$

ROC 曲線下面積は、各手法の判別精度を示している。表 1 よ

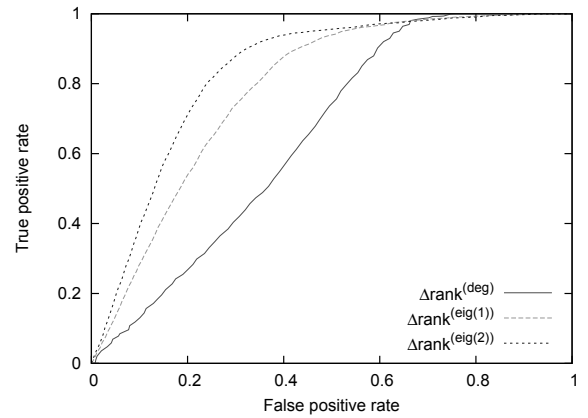


図 3 評価結果: ピアリング判別における ROC 曲線

Fig. 3 Result: ROC curve on inferring peering relationship

り、隣接 AS 数を用いた AS 間接続関係の数値評価 ($\Delta rank^{(deg)}$) による判別手法と比べ、最大固有値に対する固有ベクトルを用いた AS 間接続関係の数値評価 ($rank^{(eig(n))}$ ($n \in \{1, 2\}$)) による判別手法の方が、ピアリングおよびトランジットの高い判別精度を得られることが示された。

一般に、ROC 曲線上の点と $y = x$ の距離が最大となるときの閾値が最適とされている。表 2 に、ピアリング判別時の ROC 曲線 (図 3) に対する最適閾値 (th_{deg} , $th_{eig(1)}$, $th_{eig(2)}$) における *false positive rate* および *true positive rate* を示す。 $\Delta rank^{(deg)}$ によるピアリング判別では、高い *true positive rate* を得ることができている一方、*false positive rate* が高くなってしまっている。つまり、ピアリングの判別精度が良くないということがわかる。 $\Delta rank^{(eig(1))}$ および $\Delta rank^{(eig(2))}$ によるピアリング判別では、 $\Delta rank^{(deg)}$ と比べ、*false positive rate* が低く抑えられていることがわかる。さらに、 $\Delta rank^{(eig(2))}$ によるピアリング判別における最適閾値によるトランジット判別においても、高い *true positive rate* を得ることができていることが示された。つまり、最大固有値に対する固有ベクトルを用いた AS 間接続関係の数値評価では、トランジットとピアリングの区別を行うことができることが示された。

5. 結 論

本稿では、固有値解析を AS 間接続関係の数値評価に用いることで、トランジットとピアリングを区別しながら、隣接 AS 間の規模の差を数値表現する手法を提案した。本提案手法により、従来の経路表 (AS PATH) を解析する手法とは異なり、隣接情報のみを用いて AS 間接続関係を推定し、数値評価することが可能であることを示した。また、従来の数値的解析手法である隣接 AS 数による数値評価手法と比較して、ピアリングおよびトランジットの判別において高い精度を実現できることを示した。

今後は、本提案手法を Routing Information Service [12], Route Views Project [13] 等から得られる実経路データに適用し、AS 間接続関係に対するモデル理論からの詳細な評価を行う予定である。また、本提案手法の AS 間接続関係判別時における閾値付近の AS 間接続関係について評価を行い、paid peer 等の経路表から推定することができない AS 間接続関係についての検出可能性を確認することを今後の課題とする。

謝辞 本稿の執筆にあたり固有値解析についてのご助言を頂きました。国立情報学研究所 福田健介氏および東京大学 肥村洋輔氏に感謝します。

文 献

- [1] H. Xie, Y.R. Yang, A. Krishnamurthy, Y.G. Liu, and A. Silberschatz, "P4p: provider portal for applications," SIGCOMM '08: Proceedings of the ACM SIGCOMM 2008 conference on Data communication, pp.351–362, New York, NY, USA, ACM, 2008.
- [2] 浅井大史, 江崎浩, "マルチドメイン P2P 型 CDN における経路制御アーキテクチャの提案と評価," インターネットコンファレンス 2008, pp.78–84, Oct. 2008.
- [3] T. Erlebach, A. Hall, A. Panconesi, and D. Vukadinovi, "Cuts and disjoint paths in the valley-free path model of

internet BGP routing," 2005.

- [4] L. Gao, "On inferring autonomous system relationships in the internet," Networking, IEEE/ACM Transactions on, vol.9, no.6, pp.733–745, Dec. 2001.
- [5] G.D. Battista, T. Erlebach, A. Hall, M. Patrignani, M. Pizzonia, and T. Schank, "Computing the types of the relationships between autonomous systems," IEEE/ACM Trans. Netw., vol.15, no.2, pp.267–280, 2007.
- [6] B. Huber, S. Leinen, R. O'Dell, and R. Wattenhofer, "Inferring AS relationships beyond counting edges," Technical Report Nr.446, D-INFK, 2004.
- [7] S. Shakkottai and R. Srikant, "Economics of network pricing with multiple isps," IEEE/ACM Trans. Netw., vol.14, no.6, pp.1233–1245, 2006.
- [8] X. Dimitropoulos, D. Krioukov, B. Huffaker, kcclaffy, and G. Riley, "Inferring AS relationships: Dead end or lively beginning?," Experimental and Efficient Algorithms, vol.3503/2005, pp.113–125, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005.
- [9] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, kcclaffy, and G. Riley, "As relationships: inference and validation," SIGCOMM Comput. Commun. Rev., vol.37, no.1, pp.29–40, 2007.
- [10] "The CAIDA AS Relationships Dataset (09/06/2008)," <http://www.caida.org/data/active/as-relationships/>.
- [11] "Cooperative Association for Internet Data Analysis (CAIDA)," <http://www.caida.org/>.
- [12] "Routing Information Service (RIS)," <http://www.ripe.net/projects/ris/>.
- [13] "University of Oregon Route Views Project," <http://www.routeviews.org/>.

