ECONOMIZING ISP INTERCONNECTIONS AT INTERNET EXCHANGE POINTS

Jörn Altmann
TEMAP, Department of Industrial Engineering, School of Engineering, Seoul National University, South-Korea
School of Information Technology, International University, Bruchsal, Germany

Deepak Goel
School of Information Technology, International University, Bruchsal, Germany

The Internet service provider market is very competitive. Small and medium-size Internet service providers (ISPs) are competing for customers, while, at the same time, they are under price pressure from upstream providers. Therefore, these ISPs have to reduce their overall cost of interconnection. In order to address this issue, Internet exchange points (IXPs) have been built up in recent years, which allow small and medium-size ISPs to go into public or private peering with other ISPs. However, those ISPs do not have sufficient information to select the optimal set of ISPs, with which they should go into private peering agreements. In this paper, we describe an approach, which provides ISPs with the information about the most economical interconnections to other ISPs. This approach helps small and medium-size Internet service providers to reduce their interconnection costs for upstream connectivity and to improve network performance for their customers. To achieve that, our approach uses Internet topology information in close neighborhood of the ISP (which is determined by the set of ISPs connected to the IXP), measurement information about the number of bytes transmitted, and traffic pricing schemes. Based on real data, our analysis results demonstrate that our approach provides the necessary information to ISPs for locally optimizing their interconnection agreements (e.g. peering, sibling, transit agreements).

Keywords: ISP interconnection, peering agreements, transit agreements, traffic measurements, business intelligence, traffic analysis, routing inefficiencies, pricing, autonomous systems.

1. Introduction

The work presented in this paper will provide business intelligence support for small and medium-size Internet service providers (ISPs). Small and medium-size ISPs offer network services to their customers (e.g. end-users, enterprises, or small ISPs) by connecting them to upstream service providers (tier-1 or tier-2 service providers). These small and medium-size ISPs have to compete for customers and, at the same time, are under high price pressure from their upstream ISPs. For example, currently, ATT and BellSouth, which control a large portion of the USA backbone, are planning to differentiate their network services by offering two different network quality levels to their customers, such as small and medium-size ISPs and data center providers (Paczkowski, 2006). This plan is an attempt to extract the surplus of their customers. Therefore, this situation challenges tier-2 ISPs to find the optimal price-quality offerings for their upstream connectivity.

The use of Internet exchange points (IXP) is one way to address this issue. IXPs provide an infrastructure to ISPs to lower the setup cost for interconnections. It also allows small ISPs to go into public peering agreements. However, it is still difficult for medium-size ISPs to find the most appropriate interconnection partner among a set of ISPs that are connected to the IXP. Lack of this knowledge leads to peering/transit agreements that are uneconomical for these ISPs.

This paper shows how a medium-size ISP that is connected to an IXP can determine the optimal set of upstream service providers. This information would lower the cost and improve the network performance. Our approach does not try to determine the topology of the entire Internet and provide a global optimum, as it is done in previous research (Vazquez, et al., 2002)(Chang, et al., 2001)(Mao, et al., 2003)(Ubramanian, et al., 2002). Instead, we only determine all ASs and the AS topology in close neighborhood of a single ISP. This information together with the amount of traffic exchanged between those ASs is used to find a set of interconnection agreements that is locally optimal.

The current ISP interconnection agreements (contracts) are established based on trust relationships, the reputation of providers, the price, and the network performance. Aspects such as the routes of packets and the amount of bytes exchanged between ASs have not been considered yet.

To determine the AS-level forwarding paths as seen by a single ISP, we use a well-known approach. First, *Traceroute* is executed to the destination IP addresses of traffic of the analyzing ISP. Then, the resulting IP addresses are mapped to AS numbers.

For the analysis, our approach measures the number of bytes exchanged between the analyzing ISP and all destination ISPs. Then, these numbers of bytes are added up for each AS that is involved in the traffic forwarding. The result is a weighted tree structure with ASs as nodes and number of bytes exchanged as link weights. Using this approach, the analyzing ISP can visualize the flow pattern of its traffic. This enables small and medium-size ISPs to economize their interconnection agreements based on the price, the route, the market situation, and the byte volume exchanged.

The rest of the paper is organized as follows: In the next section, we give an overview about the different types of ISP interconnection agreements, the pricing schemes between ISPs, and related work. The third section explains the methodology for establishing economical interconnections. The software architecture of our approach and how it can be used in making economical decisions for interconnection agreements is described in the fourth section. The fifth section illustrates our measurement environment and presents some measurement results, which demonstrate that our approach reduces the cost of upstream connectivity and improves network performance.

2. State-of-the-Art

2.1. Internet Topology

Today's Internet is hierarchically structured. Some ISPs, which are called tier-1 ISPs, act as the Internet backbone and provide global connectivity. Others provide regional connectivity, which are called tier-2 or tier-3 ISPs. In order to provide end-to-end network service, an ISP establishes peering and/or transit agreements with other ISPs. In peering agreements, two equal-sized ISPs agree on exchanging their traffic with each other (Norton, 1999). That means, traffic, which originates in one ISP, is terminated in the network of the other ISP and vice versa. It is a mutual agreement without any flow of money. In case of transit agreement, a customer-provider relationship exists. The provider ISP allows transit of traffic of the customer ISP, which pays for that service. Therefore, in transit agreements, a flow of money from the customer to the provider exists. In addition to these categories of interconnection agreements, two ISPs can also undergo a sibling interconnection agreement. A sibling relationship is defined for a pair of ISPs, which have a common upstream provider. It is a mutual transit agreement in which both siblings provide connectivity to the rest of the Internet for each other. This type of interconnection usually exists between two small ISPs (or for example universities), which are located close to each other and which cannot afford to upgrade their upstream Internet connectivity.

In order to provide a platform for these agreements, various Internet exchange points (IXPs) have come into existence throughout the world. IXPs are physical locations where ISPs can connect to each other via a shared infrastructure. Each participating ISP has to simply buy a connection to the exchange point and has to co-locate a router at this location. An IXP can be a commercial company or a non-profit organization.

Figure 1 illustrates the different interconnection agreements between ISPs as well as the current Internet topology. Tier-1 ISPs are located at the top of the hierarchy, providing global connectivity. Down the hierarchy, we find regional ISPs, which business is purely based on revenues from end-user or enterprises. Peering agreements exist between ISPs of the same level (e.g. between tier-3 ISPs), and transit agreements exist between ISPs that belong to different levels (e.g. between a tier-2 ISP and a tier-3 ISP).

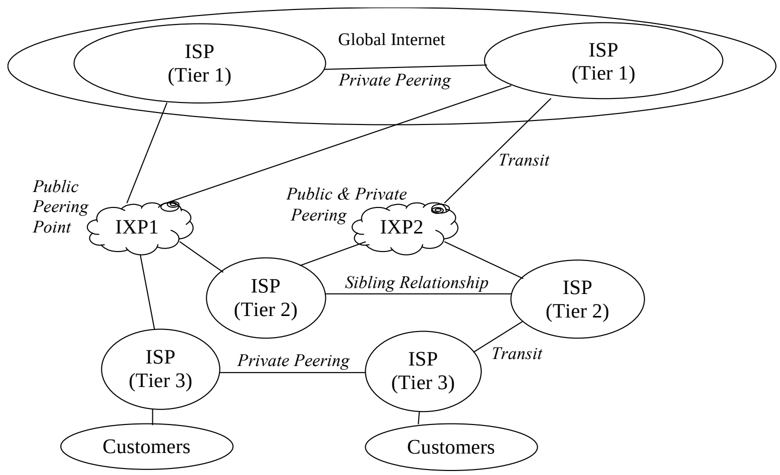

Fig. 1 Example of interconnected ISPs

However, despite the existence of infrastructure such as IXPs, there are no criteria for an ISP to choose an appropriate upstream ISP or an appropriate peer. Currently, an ISP bases its decisions only on its trust relationship to other ISPs, the size of their networks, the reputation of the other ISPs, the price, the market situation, and the network performance.

2.2. Pricing of ISP Interconnections

Peering agreements are business relationship whereby ISPs reciprocally provide access to each other's customers. Under this agreement, both ISPs do not charge each other. Peering agreements can be established in two ways. Most common is "Public Peering", under which ISPs openly announce their routes to all others. The other type of peering is "Private Peering". It is defined as peering between two parties that have a private agreement to terminate each other's traffic. In practice, the interconnection is established by using the network infrastructure (LAN) of the IXP or by having a direct point-to-point interconnection. Public peering is common at Internet exchange points.

Since there is no money transfer if ISPs are in peering agreements or sibling agreements, they only have to cover the cost for interconnecting the two networks. This process is quite inexpensive if both ISPs are already connected to an IXP. The monthly costs for connecting to a non-profit IXP range from \$200 to \$3000. It depends on the type of connection and the one time installation cost (DE–CIX, 2005).

Transit relationships are based on business arrangements, whereby an upstream ISP sells access to the global Internet. After a transit agreement is in place, the upstream service provider is responsible for providing access to the global Internet for their ISP customers (Hwang, et al., 2002). The cost of transit varies, depending on the size of the two ISPs involved in the agreement. Typically, the traffic is metered and charged, using traffic-sampling techniques. A small network service provider might have one to four contracts with upstream providers.

The current pricing scheme for interconnections to upstream ISPs is based on the capacity of the line. Within the transit agreement, the two-network service providers agree on the price for a minimum transmission rate that the downstream Internet service provider will consume (flat rate pricing scheme component) and on the price if the transmission rate is exceeded (usage-based pricing component) (Altmann and Chu, 2001). Table 1 shows a sample of a tiered flat rate pricing scheme for transit traffic.

Table 1 Example of a tiered flat rate pricing scheme for transit traffic

Mbps/month	Flat Rate
1-15	$425
16-30	$395
31-45	$365
46-60	$325

Regarding the usage-based pricing component, there are two kinds of pricing, which are offered by ISPs: 95%ile pricing and average-usage pricing.

In case of the 95%ile pricing scheme, the contract might state that the downstream network service provider has to pay a certain price per Mbps, if he exceeds the minimum transmission rate. The bandwidth consumed is measured in 5-minute (or 10-minute) time intervals. For example, the price of 50 $/Mbps has to be paid for each Mbps exceeding the threshold of 60 Mbps within the charging time interval (which is usually a week or month). The Mbps exceeding the threshold are calculated using the 95%ile pricing. Under 95%ile pricing, the 5 (or 10) minutes measuring intervals are sorted according to the Mbps rate consumed. In order to determine the Mbps value that has to be paid, the 5% of the highest Mbps rates are discarded. This means that the customer only has to pay for 95% of this usage. This pricing scheme charges the customer for the amount of bytes transmitted but also for the quality of the connection, namely the available capacity on the network.

The average-usage pricing scheme considers only the average Mbps rate that has been used over the charging time interval. This pricing is used for large customers that generate constantly the same traffic load.

2.3. Related Work

Recent research focused on inferring the topology of the Internet. The outcome can be classified into tools for discovering the topology at router-level and tools for discovering the topology at AS-level.

An early attempt started with discovering the router-level topology by executing *Traceroute* to a list of 5,000 destinations from a single network node (Pansiot and Grad, 1998). Govindan and Tangmunarunkit improved this basic tool by developing a heuristic for generating an Internet topology map (Govindan, Tangmunarunkit, 2000). Their tool, which is called *Mercator* and runs on a single host, uses informed random address probing to find destinations instead of using a list of hosts. However, their map of the Internet only shows the interconnection between routers. It does not classify the routers to an AS. The map provides a time-averaged topology, not an instantaneous view of the topology. In (Spring, et al, 2002), Spring et al. presented an Internet mapping technique, *Rocketfuel*, to directly discover router-level ISP topologies. Instead of determining the topology for the entire Internet, they focused on individual ISP networks to improve the completeness of the map within an ISP network.

Chang et al. use BGP tables to infer AS-level Internet topology from router-level path traces (Chang, et al., 2001). After generating a router-level topology from *Traceroute* results, they transferred the router-level topology into an AS-level path, using BGP tables. From those AS paths, they generated an AS-level Internet topology. Mao et al. improved this method further by creating route maps that have extremely few invalid BGP paths (Mao, et al., 2004). Once the AS-level forwarding path is determined, it is important for some analysis to determine business relationship between adjacent ASs, which can be done by analyzing BGP announcement (Gao, 2000)(Battista, 2003).

Apart from above-mentioned tools, there are various network-monitoring tools for determining network performances like delay, jitter, bandwidth, response time, and retransmission rate. CAIDA has provided a number of tools in this area. *Skitter*, a topology discovery project at CAIDA, uses BGP tables and a database of Web servers to find destination prefixes (Claffy and McRobb, 2005). It is a tool that measures the forward path and the round trip time (RTT) to a set of destination hosts by sending probe packets through the network. It probes the Internet from about 20 different locations worldwide. Another tool is the *InMon Traffic Server* (InMon, 2006). This workload tool is a web-based sFlow and NetFlow analyzer that provides access to real-time and historical traffic information.

Another workload tool of interest is *ASPATH* (NCNE, 2005). It is a suit of tools, which work in conjunction with the trace files produced by *Coral*. It can determine the total number of bytes transferred across an AS path fragment in a specified time period. It takes raw network trace information and distills it into an IP matrix. This matrix includes source and destination IP addresses, ports, protocol used, packets, bytes and a time stamp for the monitored flow. Using a local BGP table, a lookup is then performed on each destination IP to determine the AS level path and calculate the number of bytes transmitted.

Our work differs from the above approaches in how AS topology information is used. Instead of determining the full Internet topology either at router level or AS level, we are interested in the AS-level forwarding paths as seen by an ISP in its close network neighborhood. While *ASPATH* is used to determine DoS attacks, patterns of network abuse in the form of suspect FTP sites, high traffic web sites, or peer-to-peer clients, our approach uses the information about the AS-level topology along with network load and pricing schemes to determine the most economical connections for an ISP. Our approach provides guidance to an ISP for its interconnection agreements.

3. Method for Economizing Interconnection Agreements between ISPs

The inter-domain traffic engineering approach that we propose analyzes the traffic of its customers with respect to the routes of the packets, the actual destinations of packets, and the number of bytes sent. The analysis of the path of the packet reveals the distinct ASs that are involved in the transport of the packet. Combining this information with the number of bytes that has been sent, the ISP gets a tree-shaped AS topology graph with weights representing the number of bytes sent across the AS and the number of bytes which are terminated within the AS. Data can be collected for any time period. An example of this tree structure is given in Figure 2.

In detail, Figure 2 illustrates the topology of the Internet as seen from an analyzing ISP, the total average transmission rate to an AS, the average transmission rate that terminates at an AS, as well as the transmission rate z over a certain link. Note, instead of the average transmission rate, the 95%ile transmission rate could also be used. ISP F, which is the upstream provider for ISP E and ISP C, receives data at an average transmission rate of x_F. Local data arrives at a transmission rate of y_F.

Beside this information, our method for determining the most economical interconnections needs two additional kinds of information: First, the charging information, namely the transmission rate x_j in Mbps (based on average usage pricing or 95%ile pricing) that is sent to ISP j and the flat-rated threshold value t_j; second, the list of partners that are available for interconnection at an IXP. A cost C is associated with each of the possible interconnections. The cost is composed of a fixed component and a usage-based component. The fixed component c_j takes into consideration the monthly location charge of the IXP, the initial fixed cost for establishing the physical line to the IXP (or the upstream ISP), and the cost for managing the network.

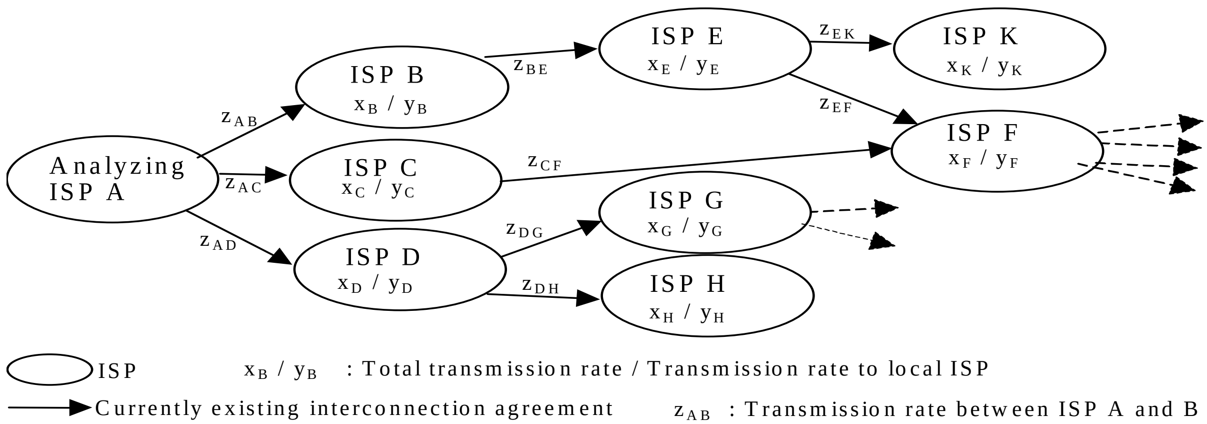

Fig.2 Internet topology and average transmission rate from the perspective of the analyzing ISP

For calculating the usage-based component, the ISP needs to know the prices p_j of the available upstream providers and the prices for the flat-rated component f_j. Based on these parameters, the network service provider can calculate the cost for each possible interconnection. Note, only those ISPs j of the tree-shaped AS topology graph are considered which belong to the list J of ISPs connected to the IXP. Therefore, the charge C_j that an ISP has to pay to an upstream provider is calculated as:

$$C_j = \begin{cases} ((x_j - t_j)\, p_j + f_j + c_j) & \text{, if } t_j < x_j \\ f_j + c_j & \text{, if } t_j >= x_j \end{cases} \qquad (1)$$

If there is no connection between the analyzing ISP and the ISP $j \in J$, the cost $C_j = 0$. If a peering agreement is possible at the Internet exchange point, the cost is calculated to:

$$C_j = c_j \qquad (2)$$

The optimization that an ISP has to perform is to calculate the cost S' for each possible combination of interconnection agreements K. The combination with the lowest cost S' is S, as shown in the following formula:

$$S = \min_{K} \sum_{j \in K} C_j \qquad (3)$$

Note, the brute force method of trying all possible combination will work. However, the classification of all ISPs according to being potential peers, potential upstream service providers, or potential customers could reduce the calculation time. Moreover, this method of classifying the neighboring ISPs will give an operator some means to influence the optimal solution. This might be desirable because of certain business conditions.

Once the optimal solution has been found, the analyzing ISP can initiate establishing peering/transit agreements with those ISPs.

For illustration, we consider the example of Figure 2, assuming that most of the traffic of analyzing ISP A is forwarded to ISP F via the upstream providers (i.e. ISP B and ISP C). Although it can also be assumed that all of these ISPs are connected to a common Internet exchange point, there is no direct connectivity between the traffic-analyzing ISP and ISP F (Figure 3). Since the cost per Mbps goes down with a higher transmission rate (progressive pricing), a high-bandwidth connection could be less expensive than two low-bandwidth connections. The network performance could simply improve through a single connection, since all packets would have to traverse less number of hops. However, currently, there is no way for the analysing ISPs to find out, due to the lack of technology for combining information about topology, pricing, and the amount of traffic exchanged.

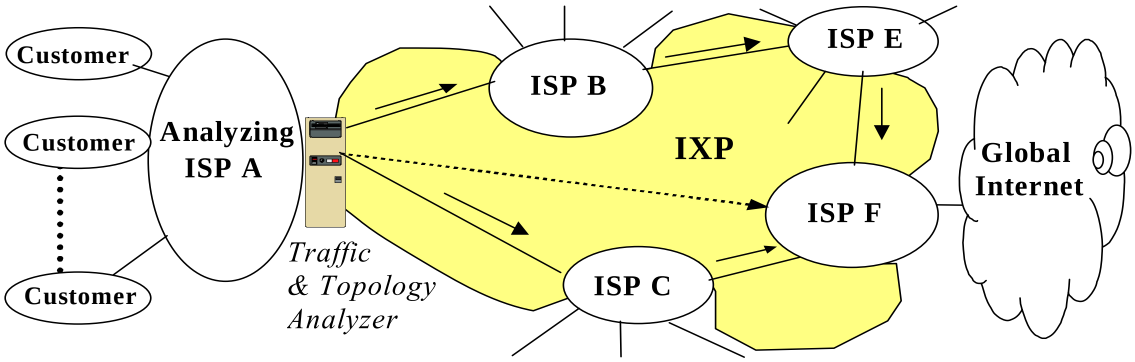

Fig.3 Example of neighboring ISPs at an interconnection at IXP

In the following, we give two examples that demonstrate how inefficiencies that can be detected and resolved with this analysis. The first example assumes that the analyzing ISP has a transit connection of 32 Mbps for $1185 with ISP B and another transit connection of 10 Mbps for $425 with ISP C (Table 1). In total, it is spending 1610 $/month for a 42 Mbps connection. More economical solutions would be to consolidate the connections into one, either with ISP B, ISP C, or ISP F (if most of the traffic is going to ISP F anyway, see Figure 3). This could result in a saving of 425$/month according to Table 1. Since information about the traffic flow to ISP F is available now, the analyzing ISP can negotiate with ISP F and with both upstream providers (ISP B and ISP C) for better peering or transit agreements.

The second example is based on the fact that a capacity upgrade for a connection is very expensive. An upgrade requires the purchase of an expensive, high-bandwidth network card. Considering this, we can assume a situation in which the analyzing ISP is connected to ISP F initially. Over time, however, its traffic grew such that its connection had to be upgraded. Knowing the pricing and traffic information, the analyzing ISP went for sibling agreement with another ISP, the most economical solution, and not for upgrading the network card. So, a purchase of a low-capacity network card was sufficient, resulting in a substantial saving.

The above two examples were based on economizing transit connections. If an ISP is connected to an Internet exchange point, it has the chance to go for peering agreements. A peering agreement is usually more economical than a transit agreement, since it only requires the purchase of a network card. However, due to business policies and lack of knowledge about the actual traffic flow, ISPs usually cannot find easily potential partners for private peering agreements. Consequently, ISPs have to undergo transit agreements to provide their customer with connectivity to potential private peering partners. Knowing the actual flow of traffic and its traffic flow pattern could help an ISP to choose its peers, saving costs from a lower load on the transit connection, and increasing the QoS of the network (lower number of routers that a packet has to travel).

Therefore, we can state that, using this method of analyzing the transmission rate to other ISPs connected to its IXP, an ISP can discover potential peering partners and reduce its cost for upstream connectivity.

4. Measurement Environment

Our approach, *iGuide*, is implemented on top of *NeTraMet* (Figure 4). *NeTraMet* is a metering and accounting tool that runs on Windows and UNIX (NeTraMet, 2005). It builds up packet and byte counts per traffic flows. The kinds of traffic flows, which have to be processed by NeTraMet, are specified in a user-defined rule set. The resulting flow data can be accessed through SNMP.

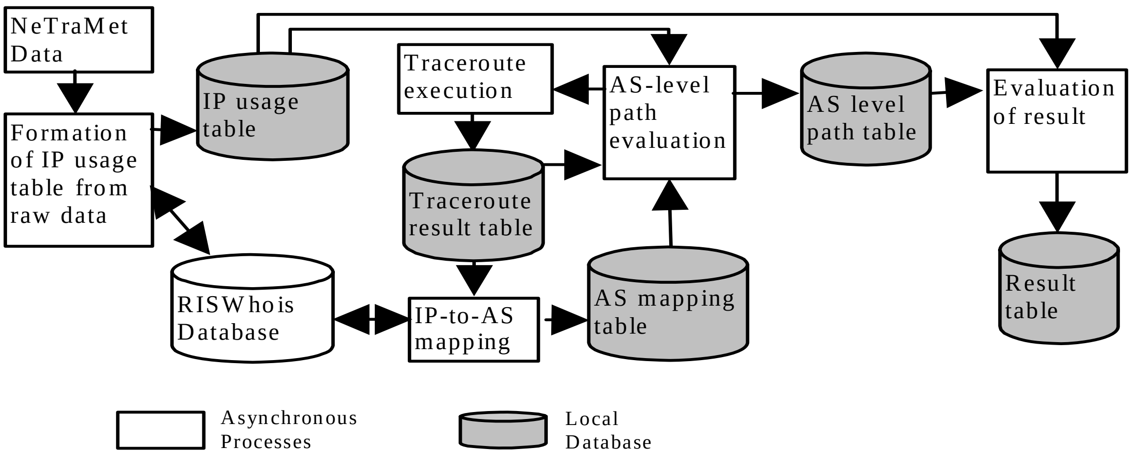

Fig.4 *iGuide* data processing architecture

An *IP usage table* is distilled from the raw NeTraMet data. This table includes source IP address, destination IP address, destination IP prefix, average Mbps sent, average Mbps received, number of bytes send, and number of bytes received (Table 2). The information about destination IP prefixes is taken from *RISWhois* database (RISwhois, 2005). The IP usage table, which is maintained in a local database, is updated every 5 minutes.

Table 2 Example of the IP usage table

source_ip_prefix	destination_ip_prefix	avg_Mbps_send	avg_Mbps_received	no_of_bytes_send	no_of_bytes_received
172.16.0.0/16	213.219.64.0/18	0.0530724	1.86856	8597732636	302706868754
172.16.0.0/16	130.117.0.0/16	0.0368002	1.54703	5961639049	250618420389
172.16.0.0/16	64.236.0.0/16	0.0282483	1.01833	4576218919	164969546838
172.16.0.0/16	208.53.136.0/22	0.0257665	0.99721	4174180760	161547980568
172.16.0.0/16	213.251.128.0/18	0.0245118	0.853635	3970918313	138288886576
172.16.0.0/16	217.75.128.0/24	0.111271	0.751198	18025975596	121694012038
172.16.0.0/16	193.178.175.0/24	1.05028	0.672096	170145898476	108879550781
172.16.0.0/16	64.224.0.0/16	0.111127	0.420338	18002535541	68094814610
172.16.0.0/16	64.235.246.0/24	0.216132	0.417091	35013395990	67568743119
172.16.0.0/16	81.23.224.0/19	0.00865807	0.291466	1402607459	47217432824
172.16.0.0/16	81.250.128.0/17	0.355418	0.170633	57577770546	27642550022
172.16.0.0/16	81.53.0.0/16	0.412287	0.136817	66790474414	22164307589
172.16.0.0/16	62.26.0.0/15	0.00518544	0.106829	840041639	17306314308
172.16.0.0/16	61.174.0.0/16	0.00477352	0.0984833	773309550	15954296271
172.16.0.0/16	69.31.98.0/23	0.00268478	0.0938211	434934031	15199018498
172.16.0.0/16	212.227.0.0/16	0.0584161	0.0904382	9463415341	14650983179
172.16.0.0/16	81.169.160.0/19	0.0267456	0.0805147	4332781057	13043377955
172.16.0.0/16	64.12.0.0/16	0.00994997	0.0729986	1611895515	11825773420
172.16.0.0/16	193.108.94.0/24	0.00548129	0.0673201	887969338	10905850523
172.16.0.0/16	64.94.224.0/20	0.033926	0.0671531	5496008831	10878799373

The local database provides both persistent storage of measurement results and a substrate for inter-process communication between asynchronously running processes (Figure 4). Each new destination IP address is checked asynchronously on whether the corresponding IP level path already exists in the *Traceroute result table*. If it exists, *Traceroute* has already been executed. If not, Traceroute will be executed to a randomly chosen IP address of the corresponding destination IP prefix. The IP address and the resulting IP-level path, obtained from Traceroute, are stored in the table (Traceroute result table). This technique is sufficient for determining the AS-level path and it reduces the number of Traceroute executions significantly. In order to cope with the high rate of new destination IP addresses that has to be added to the database, Traceroute executions are performed in parallel.

The period of time for updating the IP-level paths depends on the frequency of path fluctuations. However, stored IP-level path information can be updated after the expiration of a timeout, which can be set. Currently, the timeout is set to 15 days.

To infer the AS-level connectivity, we determine the AS of each router hop of the IP-level path and extract AS adjacency information from the resulting sequence of ASs associated with the IP-level path. The *IP-to-AS mapping* is performed with the help of the *BGP Table* from the remote *RISwhois* database. The mapping information is also stored in the local database (*AS-mapping table*) so that future queries to the same IP address are not made. The *RISwhois* database, which maintains BGP sessions of 14 BGP routers on the Internet, provides updated version of its BGP routing table every 8 hours (RISwhois, 2005). This guarantees that the information is up-to-date. Unlike conventional BGP routing tables, it retains not only the "best" AS path but also each advertised AS route. Therefore, it is essentially a collection of BGP routing tables of all its peers. Sometimes, a single prefix is mapped to more than one AS from the BGP routing tables because of hardware failure or configuration errors of some ISPs. In such cases, we remove the paths containing those prefixes from our database. An example of an *AS-level path table* is shown in Table 3.

Table 3 Example of an AS-level path table

source_ip_prefix	destination_ip_prefix	asn_1	asn_2	asn_3	asn_4	asn_5	asn_6	asn_7	asn_8
172.16.0.0/16	82.179.170.0/24	private	AS5409	AS286	AS1239	AS3549	AS4436	AS26627	AS1273
172.16.0.0/16	212.23.176.0/22	private	AS5409	AS286	AS8763	AS6453	AS15557	AS12626	AS12566
172.16.0.0/16	130.240.0.0/16	private	AS5409	AS286	AS8763	AS3356	AS2603	AS1653	AS2831
172.16.0.0/16	203.207.128.0/17	private	AS5409	AS286	AS209	AS4134	AS4847	AS4808	AS7549
172.16.0.0/16	82.179.166.0/24	private	AS5409	AS286	AS1239	AS3549	AS4436	AS26627	AS1273
172.16.0.0/16	202.108.128.0/17	private	AS5409	AS286	AS209	AS1239	AS4837	AS4808	AS9305
172.16.0.0/16	212.23.160.0/20	private	AS5409	AS286	AS8763	AS6453	AS15557	AS12626	AS12566
172.16.0.0/16	59.151.0.0/18	private	AS5409	AS5669	AS3320	AS4837	AS4808	AS9802	AS9308
172.16.0.0/16	163.118.0.0/16	private	AS5409	AS286	AS3356	AS19151	AS19962	AS10701	AS31967
172.16.0.0/16	211.98.96.0/19	AS5409	AS286	AS209	AS4134	AS4812	AS4808	AS17964	AS9394
172.16.0.0/16	202.37.240.0/23	private	AS5409	AS286	AS5459	AS4637	AS4763	AS9901	AS4768

After the AS-level forwarding path is available, the number of bytes sent and the number of bytes received by each AS are calculated based on the AS-level forwarding path information and the IP destination usage information. Based on this information, the transmission rates (i.e. average number of Mbps sent / received) of each AS are calculated (Table 4).

Table 4 Example of the result table

asn	avg_Mbps_send	avg_Mbps_received	no_of_bytes_send	no_of_bytes_received
AS5409	3.84968	12.2718	623657697557	1988057562588
AS286	2.04377	7.08641	331096287645	1148006859824
AS8763	1.8419	5.71874	298388314522	926441705236
AS1299	0.228539	2.77494	37023290297	449540057900
AS174	0.109094	2.74282	17673295750	444336448617
AS3249	0.0537766	1.87236	8711815695	303321610890
AS1668	0.0829531	1.28055	13438397471	207449536494
AS209	0.337365	1.25639	54653222539	203536144510
AS30058	0.0263986	1.00653	4276566992	163058653378
AS16276	0.0246362	0.856066	3991063237	138682724218
AS12301	0.112618	0.752999	18244053779	121985796162
AS9070	0.112546	0.75232	18232474428	121875925300
AS12867	0.111271	0.751198	18025977001	121694015891
AS8220	1.053	0.687521	170586835947	111378400552
AS6939	0.217659	0.456782	35260795627	73998626854
AS11305	0.111176	0.420943	18010521686	68192797890
AS1239	0.780319	0.418046	126411804250	67723436821
AS3356	0.0994153	0.337945	16105297918	54747112103
AS6762	0.0200372	0.301889	3246020816	48906096898
AS24730	0.0091527	0.294665	1482737436	47735664951

In addition to this, our approach can also provide information about the average transmission rate (i.e. number of Mbps exchanged) along a certain path.

5. Measurement Results

The results that are presented here are based on the entire traffic passing through the external gateway of our university over the course of 15 days. Our university is a customer of the ISP with the autonomous system number AS5409. Although the measurement experiments have been conducted at the gateway of our university, the measurement results still represent how ISP AS5409 and its upstream ISP (AS286) route their packets into the Internet (Table 5). The traffic accounting and traffic analysis have been performed using our approach. The ISP AS5409 is a tier-3 ISP and is connected to DE-CIX (DE–CIX, 2005). DE-CIX is an Internet exchange point in Frankfurt, Germany. 145 ISPs are connected to DE-CIX for private peering, public peering, or transit agreements. Out of those 145, 53 are not connected to any other IXP. We classified these 53 ISPs as small ISPs. Every external packet originating from AS5409 is routed to the location of DE-CIX, where it gets forwarded to upstream service providers or peers.

Table 5 Specific AS-level-forwarding paths observed

Destination IP Prefixes	First AS	Second AS	Third AS	Fourth AS	Fifth AS	Sixth AS
198.146.0.0/16	AS5409	AS286	AS209	AS3356	AS6389	AS7892
137.78.0.0/16	AS5409	AS286	AS209	AS2914	AS226	AS127
213.42.32.0/19	AS5409	AS286	AS209	AS6453	AS8961	AS5394
82.150.134.0/23	AS5409	AS286	AS5417	AS29686	AS29502	
81.22.32.0/24	AS5409	AS286	AS5417	AS3491		
69.31.10.0/21	AS5409	AS286	AS1239	AS3549	AS4436	AS26627
132.70.216.0/21	AS5409	AS286	AS1299	AS13264	AS8584	
81.214.236.0/22	AS5409	AS286	AS8763	AS6762	AS9121	
194.246.114.0/23	AS5409	AS286	AS8763	AS3356	AS24587	
212.66.96.0/20	AS5409	AS286	AS8763	AS6762	AS9035	

AS connected to DE-CIX (green) AS External to DE-CIX (white)

Table 5 presents some specific AS-level forwarding paths, starting from AS5409 via AS286. These routes are found by checking each route for two criteria: First, existence of an AS that does not belong to the IXP but its neighbors in the AS forwarding path do; second, existence of an AS forwarding path with more than two ASs belonging to the IXP. Looking at Table 5, it becomes immediately clear, that this constitutes a routing inefficiency. The routes shown in Table 5 are a very good example for situations, in which a physical interconnection between two ASs (e.g. AS286 and AS3356) exists but no business relationship. AS286 and AS3356 are both connected to DE-CIX but do not have an interconnection agreement. The light-gray highlighted cells of Table 5 indicate an Autonomous System that is connected to DE-CIX (i.e. is a member of DE-CIX). Otherwise, they are not.

For our analysis, we are only interested in the first 4 ASs that a packet crosses after leaving our university gateway. Therefore, we specified within *iGuide* to aggregate all AS-routes that we obtained (see Table 3) with respect to the first 4 ASs. Note, there can be a high number of AS-forwarding paths beginning with the same 4 ASs. While aggregating traffic data, the average transmission rate (Mbps) is calculated for incoming and outgoing traffic of all those aggregates on the fly. The results obtained are shown in Table 6.

Table 6 Transmission rates for some specific routes

First AS	Second AS	Third AS	Fourth AS	Average Mbps send	Average Mbps received
AS5409	AS286	AS209	AS3356	0.041	0.246
AS5409	AS286	AS209	AS2914	0.028	0.036
AS5409	AS286	AS209	AS6453	0.001	0.006
AS5409	AS286	AS1239	AS3549	0.002	0.012

While Table 6 provides a detailed view on different paths through the close neighborhood of the analyzing ISP, Figure 5 provides a broader view on the traffic pattern between those ISPs.

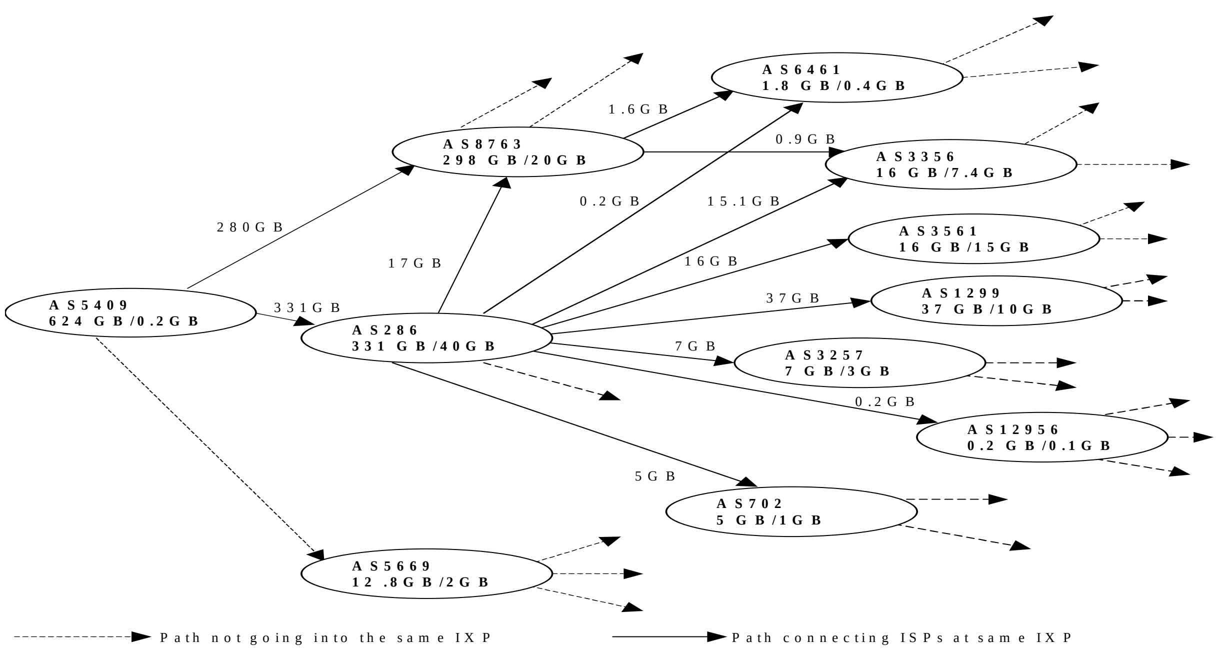

Fig. 5 Internet topology from the perspective of AS5409

Figure 5 presents a graph that represents information about the network topology (i.e. its business relationship) and the number of bytes sent to an AS. Nodes of the graph represent different ASs. Links between nodes denote an interconnection agreement between two ASs. The three weights are defined as described before: The two node weights represent the number of bytes that are sent to the AS and the number of bytes that do not get forwarded. The link weight indicates the total number of bytes sent across that link. ASs that are connected to DE-CIX are shown by solid arrows. Dotted arrows represent external links.

Based on the information in Table 6, Figure 5, and the pricing scheme from Table 1, our approach can provide suggestions about economical routes. It helps to find the most appropriate set of upstream providers, peers, or siblings. Note, since our experiments are carried out on our university gateway (and we did not have access to other networks), we extrapolate the amount of data (respectively the transmission rate) that the provider (AS5409) of our university and its upstream provider (AS286) deal with. Without loss of generality, we assume that AS5409 has 100 customers of the size of our university and AS286, which is bigger than AS5409, has 1000 such customers. In the following, four scenarios are presented, demonstrating how ISPs can optimize their interconnection agreements by using our approach.

5.1. Scenario 1

Figure 5 visualizes the traffic flow from the perspective of AS5409. Information about the amount of bytes transferred on a particular link and the amount of bytes terminated at a particular AS is shown. AS5409 can use this information to optimize its interconnection agreements. Figure 5 shows that 1.8 GByte of data is forwarded from AS5409 to AS6461 via 2 ASs (AS286 and AS8763). AS5409 could reduce its interconnection cost by establishing a peering agreement with AS6461. Since AS6461 is already connected to the same IXP, the interconnection cost would be low. This peering agreement would also lower the cost for the transit agreement with AS286 and AS8763. Besides, the agreement would improve the network performance for data directed to AS6461, because of the fewer number of hosts traversed.

5.2. Scenario 2

Assuming the case in which traffic of AS286 has gone up to such a level that it needs to upgrade its network card with AS209. In the absence of our analysis results, AS286 would have considered upgrading its network card as the economical solution (which is expensive). With the help of our approach, AS286 can analyze its data traffic. Using the information given in Table 6, AS286 recognizes that, in average, 0.041 Mbps are sent to AS3356 and 0.028 Mbps are sent to AS2914; respectively, after extrapolating, 41 Mbps are sent to AS3356 and 28 Mbps are being sent to AS2914. Since both of them are connected to the same IXP as AS286, AS286 can go into sibling agreements with either of them. If AS286 goes for this solution, it will not only save the upgrading cost but also the usage-based charges for not sending the traffic to AS3356 and AS2914 through its upstream provider (AS209). This could sum up to a saving of about 2000 $/month according to the transit cost in Table 1. Although it may still be possible that agreements with AS3365 and AS2914 cannot be reached due to business policies, it enabled AS286 to start those negotiations. This has not been possible before because of the simple lack of this kind of information.

5.3. Scenario 3

Although peering is free, many ISPs feel reluctant to go for peering. This is due to lack of information about whom to peer with and the effort involved in maintaining a peering connection. Our approach provides the required information to make informed decisions. Assuming that AS5409 has a usage-based transit agreement with AS286, AS5409 finds out from Table 6 that it is sending 0.106

Mbps per customer to AS8763 through this interconnection. According to our assumption that AS5409 has 100 customers, 10.6 Mbps (average transmission rate) is being sent to AS8763 in total. This would result in a saving of about $400 per month (Table 1), if AS5409 entered a peering agreement with AS8763.

5.4. Scenario 4

Our approach can also be used to find the most appropriate upstream providers. After analyzing its traffic flow pattern in a graph similar to Figure 5, AS286 found that a large amount of packets goes to a certain destination. Therefore, AS286 checks whether it is possible to improve the performance (with respect to the number of ASs that AS286's packets have to cross) to this certain destination for its customers. This improvement could be achieved by selecting a more appropriate upstream provider. If such an interconnection could be found but increases the cost, then AS286 has to decide based on its policies whether it wants to improve its customer service or go for a low-cost solution. We are unable to produce concrete data for this scenario, as we could only take measurements at our university, which is not multi-homed. Measurement would have to be done at larger ASs, such as AS286.

6. Conclusions

Our approach (*iGuide*) that we presented helps small and medium-size ISPs to find appropriate peering partners, siblings, and upstream providers for optimizing their network interconnectivity. The approach guarantees that the best possible set of transit and peering partners are found, which consequently results in low cost for upstream connectivity at desired quality.

In order to optimize the interconnectivity of a small or medium-size ISP, the approach needs AS-level topology information, measurement information about the number of bytes transmitted (respectively, the average transmission rate), and pricing scheme information. We demonstrated the effectiveness of our approach by presenting analysis results of real measurement data. Besides this, the analysis of 4 different scenarios showed how actual measurement information could be used to select the optimal set of interconnection agreements between ISPs. In all four scenarios, small or medium-size ISPs could lower their cost for interconnection or improve the network performance.

Previous work on analyzing the data traffic to a certain AS has focused on detecting DoS attacks or patterns of network abuse of certain applications. Our work is novel in the sense that data traffic (i.e. the amount of bytes transmitted) is analyzed with respect to the business interest of small or medium-size ISPs.

7. References

Altmann, J. and Chu K., 2001, "A proposal for a flexible service plan that is attractive to users and Internet service providers," IEEE INFOCOM, Conference on Computer Communications, Proceedings, Anchorage, Alaska, USA.

Baker, F., 1995, "RFC 1812: Requrements for IP version 4 routers," http://rfc.net/rfc1812.html (cited October 15, 2006).

Burch, H. and Cheswick, B., 1999, "Mapping the Internet," IEEE Computer, Vol. 32, No. 4.

Chang, H., Jamin S., and Willinger, W., 2001, "Inferring AS-level Internet topology from router-level path traces," SPIE ITCOM, Workshop on Scalability and Traffic Control in IP Networks, Proceedings, Denver, USA.

DE-CIX, Deutscher Commercial Internet Exchange, 2006, http://www.de-cix.net/ (cited October 15, 2006).

Di Battista, G., Patrignani, M., Pizzonia, M., 2003, "Computing the types of the relationships between autonomous systems," IEEE INFOCOM Conference on Computer Communications, Proceedings, San Francisco, USA.

EuroRings, 2006, "International service provider," http://www.eurorings.kpn.com (cited October 15, 2006).

Gao, L., 2001, "On inferring autonomous system relationships in the Internet," IEEE/ACM Transactions on Networking, Vol. 9, No. 6, pp 733-745.

Govindan, R. and Tangmunarunkit, H., 2000, "Heuristics for Internet map discovery," IEEE INFOCOM, Conference on Computer Communications, Proceedings, Tel-Aviv, Israel.

Huffaker, B., Fomenkov, M., Moore, D., and Claffy, K.C., 2001, "Macroscopic analyses of the infrastructure: measurement and visualization of Internet connectivity and performance," Usenix LISA.

Hwang, J., Altmann, J., Oliver H., and Sánchez-Beato Suárez, A., 2002, "Enabling market-managed QoS interconnection in the next generation Internet by a modified BGP mechanism," IEEE ICC2002, International Conference on Communications, Proceedings, New York, USA.

InMon, 2006, "trafficserver," http://www.inmon.com/trafficserver.htm. (cited October 15, 2006).

Mao, Z.M., Johnson, D., Rexford J., Wang J., and Katz R., 2004, "Scalable and accurate identification of AS-level," IEEE INFOCOM, Conference on Computer Communications, Proceedings, Hong Kong.

Mao, Z.M., Rexford J., Wang, J., and Katz, R.H., 2003, "Towards an accurate AS-level Traceroute," ACM SIGCOMM, Proceedings, Karlsruhe, Germany.

NCNE, 2005, "aspath-qa," http://www.ncne.org/software/tools/ASPath-QA.html (cited October 15, 2006).

NeTraMet & NeMaC, 1999, Release Notes, http://www2.auckland.ac.nz/net/Accounting/ntmref.pdf (cites October 15, 2006).

Norton, W.B., 1999, "Interconnection strategies for ISPs," NANOG – The North American Network Operators' Group, http://www.nanog.org/mtg-9905/norton.html (cited October 15, 2006).

Norton, W.B., 2002, "A business case for ISP peering," http://www.equinix.com/pdf/whitepapers/Business_case.pdf (cited October 15, 2006).

Paczkowski, J., 2006, "Prioritize some of these services right out of business," http://blogs.siliconvalley.com/gmsv/2006/01/net_neutrality_.html (cited October 15, 2006).

Pansiot, J. and Grad, D., 1998, "On routes and multicast trees in the Internet," ACM SIGCOMM CCR, Computer Communication Review, Vol. 28, No. 1.

RISwhois, 2006, "IPv4/IPv6 address to origin mapping," http://www.ripe.net/ris/riswhois.html (cited October 15, 2006).

Spring, N., Mahajan R., and Wetherall, D., 2002, "Measuring ISP topologies with rocketfuel," ACM SIGCOMM, Pittsburgh, USA.

Ubramanian L., Agarwal S., Rexford J., and Katz R.H., 2002, "Charaterizing the Internet hierarchy from multiple vantage points," IEEE INFOCOM, Conference on Computer Commuinications, Proceedings, New York, USA.

Vazquez, A., Pastor-Satorras, R., and Vespignani A., 2002, "Internet topology at the router and autonomous system level," http://www.arXiv.org, cond-mat, cond-mat/0206084 (cited October 1, 2006).