A Value Based Framework for Provider Selection of Regional ISPs

Miao Li, Hui Wang, Jiahai Yang, Chenxi Li
1. Network Research Center, Tsinghua University
2. Tsinghua National Laboratory for Information Science and Technology(TNList)
Beijing , R.P. China
mileslmw87@126.com, {hwang,yang}@cernet.edu.cn, li-cx13@mails.tsinghua.edu.com

Abstract—**To access to the Internet, regional Internet Service Providers (ISPs) have to buy transit service from global ISPs. Provider selection strategies are related closely to ISPs` economic interests. With the growing number of potential transit provider and the flattening topology of the Internet, it's getting harder for ISPs to select upstream provider empirically as before. In this paper, we propose a concept of bargaining power as an important decision-making criterion of ISPs during their provider selection process, and design a value based framework to help ISPs' provider selection based on it. The bargaining power of each involved ISP is computed by applying the Shapley Value based transit value distribution mechanism to each involved traffic flow, taking into consideration the cooperative possibility among ISPs and the market roles these ISPs play, i.e., potential providers or potential competitors. It reflects not only the cost and link level transit performance, geographical constraints, but also includes the influence of interconnection impacts, demand/supply relationships by analyzing the traffic content and commercial relationships among ISPs. We then design a quantitative provider selection framework and instantiate our framework using the operation data of a real-world network, CERNET, a national ISP in China. In addition, we evaluate our provider selection results for CERNET and the experimental results show the effectiveness and practicability of our solution in this paper.**

Keywords- network management; network economics; provider selection strategy; Shapley value

I. INTRODUCTION

The Internet is a worldwide network of interconnected networks operated by different ISPs. Most ISPs, especially the regional ISPs, have to pay other ISPs to transit their traffic to the world. Each ISP has to optimize its provider selection periodically to increase its profit and/or provide good performance for its customers.

ISPs are used to select providers empirically and provider selection is often considered as an "art" rather than science. In most cases, ISPs make decisions based on scale, customer cone size of potential providers or rough traffic analysis. On one side, the number of potential providers was limited due to the limited infrastructure of ISPs before. On the other side, for a long time, the Internet can be viewed as a hierarchical network [1] and transit relationships between ISPs are stable. 10-15 ASes peer with each other, forming a clique and occupying the highest layer in the core (or say Tier-1). Some more ISPs (Tier-2), such as national ISPs, large regional ISPs and large content providers, buy transit service from tier-1 ISPs while providing transit service to lots of small regional ISPs and small corporations in the edge of Internet (Tier-3). ISPs can make decisions according to the layer attribute empirically [2], i.e. a smaller regional ISP pays tier2 ISPs for the traffic transited in both directions of the link; ISPs in the same layer tend to free peer with each other. We can see to some extent ISPs was able to make provider selection decisions intuitively at that time.

However, recent 10 years witnessed enormous changes of the Internet and provider selection for regional ISPs gets harder and harder. It can be attributed to the following reasons. First, the potential provider number increases dramatically in these years. It is due to the increasing number of ISPs worldwide and the high speed development of ISP infrastructure. This is more obvious with the emergence of Internet eXchange Point (IXP) which decreases the cost of interconnections among ISPs a lot. Because of these developments, ISPs can have a lot of choices when they make provider selection decisions. Secondly, lots of recent works validate the flattening trend of the Internet from different points of view [3-5] and the provider selection is turning to be more difficult to this background. Therefore, intuitive or empirical provider selection may not work well and a quantitative scientific framework of provider selection for regional ISPs is valuable and necessary.

To make an intelligent provider selection decision, a regional ISP must consider many factors comprehensively. The monetary cost is one important metric which is widely considered by other works. The performance is also an important metric for the provider selection, but the definition and evaluation method varies a lot from each other. Most works evaluate the performance of one ISP based on the average performance to all possible destinations through the provider under study. However, monetary cost and simple average performance are not enough. For example, performance evaluation should also consider the importance of traffic content or user demands. Good performance to a destination with a large amount of traffic demand should induces a bigger plus than good performance to a destination with small amount of traffic demand. Besides, we believe the interconnection impact is a quite important aspect in the provider selection which is neglected by other works. Considering a provider ISP which holds large valuable or hot resources. Obviously, its price can be set to a high value since other ISPs would be eager to be connected with it. It is the same truth for customer ISP. An access ISP with a lot of end-users can be valuable for transit ISPs. Potential providers may agree a low price to get this access ISP to be their own customer, since other of their ISPs would like to buy their service due to these end-users and then the loss can be compensated from other ISPs. In short, ISPs can be viewed with different values, according to their infrastructure, customers, contents etc. We define it as their bargaining power. In detail, to provider ISPs, the demand/supply of transit services between certain ISPs impacts the prices directly. The ISPs that monopolizes transits to some destinations

978-1-4799-0913-1/14/$31.00 ©2014 IEEE

would have a strong bargaining power; to access ISPs, ISPs with more important end-users would have a strong bargaining power. A regional ISP with less competitors has more bargaining power in the negotiation because it is hard for transit ISPs to find other ways to access these end-users, which is negative to their performance of service to other customers.

Therefore, we argue that ISP provider selection should follow these principles:

1. The cost for the connection must be low enough, both the cost for setting up the connection and for maintaining the connection.
2. The provider that provides better performance for most traffic of the customer ISP would be preferred.
3. We should consider the competitions and cooperation among all involved ISPs. ISPs should consider the bargaining power of potential ISPs and select the one with highest ratio of bargaining power to price.

In this paper, following these principles, we take cost, performance, and bargaining power into consideration and propose our value based framework using game theory to solve the provider selection problem for regional ISPs.

Let us denote the regional ISP which needs to select provider ISPs intelligently as ISP_m. We calculate the bargaining power of potential providers based on analysis of every traffic flow of ISP_m. We consider that each flow is with some amount of value. For one flow related to ISP_m, we classify all ISPs in the Internet into three groups, i.e., potential providers, competitors and non-directly-related ISPs. The potential providers are ISPs which can transmit the flow to its destination for ISP_m, while the competitors are ISPs which can provide transit service for the users in ISP_m to transmit this flow, i.e., it is possible that ISP_m may lose customer due to these competitors. Obviously competitors can cooperate with potential providers to provide transit to ISP_m's users bypassing ISP_m. We take the idea of cooperative game theory and distribute the value of this flow to all directly related ISPs using Shapley Value. We can see that the value assigned to these ISPs can represent their bargaining power on this flow. Taking all traffic flows of ISP_m into consideration, we will obtain the overall bargaining power of each ISP as the metric for its provider selection.

The bargaining power in fact is the contribution of one ISP to ISP_m to transmit ISP_m's traffic flows. If we choose the metrics for potential provider and competitors carefully, we can also include the first two principles into the concept of bargaining power. For example, the ISP must be able to access ISP_m' customers with a reasonable performance without the help of ISP_m to be a competitor.

We demonstrate the practicability and effectiveness of our method by applying our framework to a real world network environment. Based on the real world operation data of China Education and Research Network (CERNET), which includes the flow data of all international traffic and routing tables, we try to optimize its provider selection for CERNET at its international point of presence (PoP) in Hong Kong. We also evaluate our provider selection results in different ways. The evaluation results demonstrate the effectiveness of our solution.

The main contributions of our work can be summarized as follows:

- We propose a value based quantitative framework for regional ISPs to select their upstream providers when facing many transit choices.
- We are the first to incorporate the transit supply/demand situation description -- the bargaining power of potential ISPs in the framework so that the decision can be made more intelligently.
- We apply our framework to CERNET and show the practicability and effectiveness of our method.

The rest of the paper is organized as follows. In Section Ⅱ, we discuss related work. In Section Ⅲ, we present our value based decision model for ISP provider selection problem. In Section Ⅳ, we discuss the Shapley Value distribution method to distribute the value of flows among competitor ISPs and potential provider. Then in Section Ⅴ, we instantiate our model using the real world data from CERNET and evaluate our provider selection results. Finally, in section Ⅵ we conclude our work.

II. RELATED WORK

Peer selection strategy is drawing more and more attention. In recent ten years, William Norton published many valuable reports on various issues in ISP's peering from practical point of view. Based on interviews with several hundreds of ISP peering coordinators, both the ISP peering decision-making process and tactics adopted by peering coordinators are introduced in [2, 11, 12]. These works provide a good summary of empirical methods which play important roles in industry. However, these methods have been less powerful when ISPs are facing more complicated situations now and future.

In academia, a lot of previous research efforts focus on free peering settlements [9] and pricing mechanisms [10]. However, since the Internet is showing a flattening trend, recent works begin to focus more on how to select providers properly. A major part of these works model ISPs as players that try to find tradeoff between cost and performance to optimize their own revenue, without considering the peering impacts among several ISPs. [6] is the earliest work as far as we know. The author assumes ISPs' optimization objective is to minimize the AS distances. In [7], ISPs' goal is to optimize the ISP gaining taking the price factor into optimization. Based on a widely used pricing scheme, the authors solve the optimization problem using dynamic programming in polynomial time. In [8], the author takes both cost and performance into consideration and formulate it as a multi-objective optimization problem. Its objects include cost, link performance, routing diversity and average AS path.

Some recent works tend to consider the cooperation and competition among ISPs using cooperative game theory. In [9], the author defines peering value in free peering relationships and propose a quantitative framework for ISPs to set up peering agreements, both free peering and paid peering. In [13], the author formulate a cooperative game and propose team buying to reduce the transit cost for customer ISPs. Cooperative game theory is also applied in revenue distribution among ISPs to support network neutrality. Though the scenario is different, Ma's work [14,15] is suggestive for our scenario, both the network model and the revenue distribution mechanism.

III. Decision-Making Criterionin in Provider Selection

In this section, we would try to analyze the considerations of regional ISPs, summarize the key problems during their provider selection, and propose important criterions in their decision making process.

A. Considerations in Provider Selection

When an ISP considers whether to build customer-provider relationships with a potential provider, its considerations might be very complex: the customer cone size of the provider, the brand of the provider, the cost to build and maintain the relationships, the customer experience after building relationships, and even its own market strategies on cooperation and competition. Generally speaking, all the considerations could be attributed into three categories: cost, performance and bargaining power of potential ISPs.

Cost -- the cost is definitely an important consideration during network planning. The cost includes two aspects: the cost for connection establishment and the cost for maintaining the relationships. Obviously, ISP_m would prefer to potential providers that are close to its PoPs geographically since the connection establishment cost would be low. After the connection is set up, the customer ISP should pay its provider for the traffic transited between them. How the provider charges the customer, i.e., price and cost function, is one of the most important criteria during provider selection.

Performance – Intuitively, ISP_m prefers providers which can provide high quality service to its end users. But what is "high quality" and how to evaluate the performance of a potential provider? There may be various ways and metrics. In this paper, we evaluate a potential provider as follows:

- Single-flow Performance: the performance between the potential provider and the destination of a flow generated by ISP_m. It measures the performance of this flow if ISP_m delivers the flow to the potential provider to transmit to its destination. The performance can be evaluated using different metrics based on user requirement for this flow. For example, we can use RTT for flows of high-reliability applications and use throughput for flows of high-bandwidth applications. It can be simplified by assuming all flows to a same destination have a same performance.
- Overall Performance: the overall performance of a potential provider is a weighted result of performance of all flows. The weighting factor can be traffic volume of a flow, which means the transit service quality to hot destinations deserves more attention in the evaluation of potential providers.

Bargaining Power – In a market, the demand and supply condition determines the price of a product or service. In the Internet, ISPs with more bargaining power would deserve a higher price. The bargaining power reflects the significance of one potential provider in transmitting flows of ISP_m. The potential providers with different transit cost functions can be either irreplaceable, useless, or replaceable. Consider a particular flow, a related potential provider has its potential competitors and cooperators. For ISP_m, less competitors and more potential cooperators will lead to the decreasing of each potential provider's (cooperator's) bargaining power in the trading.

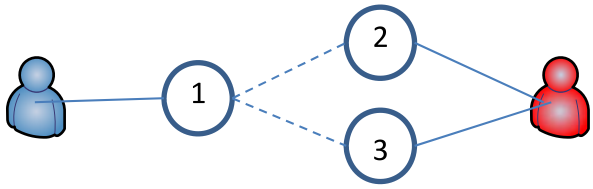

Figure 1. Bargaining Power Analysis in the Trading

We further explain the concept of "competitor", "cooperator", and "bargaining power" by the example shown in Figure 1. In this example, for the flow from the left person to the right person, ISP1 has more bargaining power than ISP2 and ISP3, because ISP1 is necessary to connect these two persons while ISP2 and ISP3 can be replaceable. To enable the communications between two persons, ISP1 needs a cooperator to provide transit service. ISP1 can negotiate with either ISP2 or ISP3. However, ISP2 and ISP3 must cooperate with ISP1 to finish the transmission. Since the communication is also a demand from customers of ISP2 and ISP3, these two ISPs must connect with ISP1. For ISP1, ISP2 and ISP3 are both cooperators, and it has no competitor. For ISP2 or ISP3, its competitor is the other ISP and each of them has only one cooperator.

Obviously, for this single flow, ISP1 have a stronger bargaining power than ISP2 and ISP3. Bargaining power represents the significance of the ISP in the interconnection market and should be taken into consideration in the provider selection problem.

This explains the analysis of a single flow. We should analyze all flows generated by (i.e., to or from) ISP_m, and then summarize them to get the overall bargaining power of each involved ISP. Therefore, the calculation of bargaining power of an involved ISP can be divided into two stages: bargaining power analysis of each flow, and overall bargaining power calculation.

To obtain bargaining power of one ISP in transmitting a flow, we should identify the potential competitors and providers of ISP_m. Figure 2 illustrates an example.

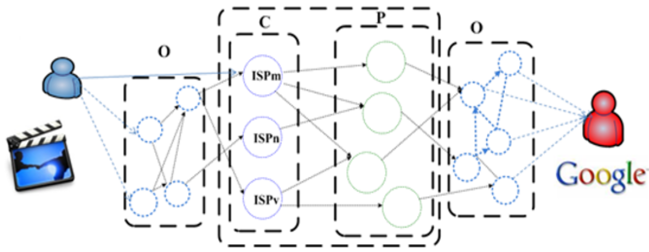

Figure 2. Bargaining Power Analysis

In Figure 2, in the middle box, *C* represents the set of potential competitors of ISP_m and ISP_m itself, while *P* represents the set of potential providers of all competitors and ISP_m that can reach the destination of the flow with acceptable performance. The left person is a customer of ISP_m, either directly connected to ISP_m or through a customer ISP of ISP_m. The right person is Google, which is only an example of destination of flows from the left person. We can see all flows from the person to Google must traverse on a path (person, *, ISP_m or a competitor, a cooperator, *, Google).

This graph for single flow analysis is constructed as follows. First, we need to identify potential competitors of ISP_m in terms of this flow. An ISP is considered to be a competitor if it can accept the flow from the person with a reasonable performance, i.e., the performance between the person and the competitor is acceptable. Second, we also need to identify potential cooperators, i.e. providers of all competitors. An ISP is considered to be a potential cooperator if it can reach the destination of the flow with a reasonable performance. In the following section, in order to simplify the analysis, we will deploy more rules to reduce the number of potential cooperators.

After we identify potential competitors and cooperators of ISP_m, we would connect ISPs in *P* and *C* according to their business relationship. If one ISP in P and one ISP in C is connected directly or it is possible for them to be connected directly, we will draw a line in the graph, which means this is a possible path to transmit the flow. The information of ISPs in the Internet is publicly available on CAIDA's website [17,19].

Based on this graph, we can analyze the bargaining power of each ISP in P and C as the example shown in Figure 1. We assume each flow is with a value, and distribute the value to each ISP according to a distribution mechanism in cooperative game theory, i.e. Shapley Value.

Finally, we can summarize the result for all flows and get the overall bargaining power of each ISP.

B. *Identifying Potential Providers and Bargaining Power Calculation*

In this part, we will describe formally how to calculate the bargaining power of one ISP. For convenience, we first define some notations. Let N denote the set of all existing ISPs in the Internet. The set of traffic flows from or to ISP_m is denoted by T and each element T_i defines a single flow. A flow is identified by a six-dimensional vector:

$T_i = \langle Src_i, Dst_i, Type_i, StartTime_i, EndTime_i, Volume_i \rangle$,

Wherein *Src*, *Dst* and *Volume* indicate source ISP, destination ISP and the volume of this flow respectively. *Type* records the type of application which generates this flow, which is related to performance demands and then measurement metrics, e.g. low latency, low loss rate.

First, let us define following notations. For now, we just assume ISP_m has a set of potential providers, and we denote one potential provider as ISP_j. Take ISP_j as an example, we define the following notations and describe how to calculate the significance of ISP_j to ISP_m.

- PoP_j – PoP locations of ISP_j.
- P_j – the price ISP_j charges ISP_m as provider.
- $Q_{j,k}$ – the performance of ISP_j to transit the traffic flow T_k from ISP_m to its destination.
- $O_{i,j}$ – the performance for the source of the flow T_i to reach ISP_j.
- V_j – the bargaining power of ISP_j, or say the value of provider ISP_j to ISP_m. It is determined by the traffic flow set T, the potential relationships map M and its potential competitors.
- R_j – the value price ratio of ISP_j to ISP_m .

$$R_{j,m} = V_j / P_j \quad (1)$$

Based on these notations, we will describe how to find potential provider for ISP_m, which is the first task for ISP_m to select providers properly. Only when ISP_j satisfies following two conditions, ISP_m would consider ISP_j as a potential provider.

1) The distance between ISP_j and ISP_m should be less than a threshold: $\min\{dist(s,t) \mid s \in PoP_j, t \in PoP_m\} < D_{max}$
2) Performance of ISP_j should be higher than a threshold: $(Q_{j,i}) > Q_{i,min}$. $Q_{i,min}$ is the performance threshold of T_i. This condition can be slightly relaxed by allowing a small part of flows to exceed the threshold.

After identifying all potential providers, we will calculate the bargaining power of each potential provider, and select the provider ISP_j with maximum R_j or V_j.

As we mentioned above, for each traffic flow T_i, from the view of ISP_m, we can categorize ISPs as three types: potential providers P_i, competitors C_i and others.

- $C_i = \{C_{i,1}, C_{i,2}, \ldots, C_{i,|Ci|}\}$ is the set of competitor ISPs of ISP_m in terms of the flow T_i. As we defined before, those ISPs who could reach the source of T_i (also end users of ISP_m) with good performance are viewed as competitor ISPs, i.e., ISP_c is a competitor if

$$O_{i,c} > O_{min} \quad (2)$$

  Since competitor ISP_c can cover end users of ISP_m related to the flow T_i, other ISPs can cooperate with ISP_c with a proper price instead of ISP_m. The more competitors, the less bargaining power ISP_m. Note that it is possible for the potential provider ISP_j to cooperate with ISP_m.

- $P_i = \{P_{i,1}, P_{i,2}, \ldots, P_{i,|Pi|}\}$ denotes the set of potential providers of ISP_m and all of its competitors that can reach the destination of T_i. In other words, ISP_p can appear in this set if 1) it is potential provider of a competitor or ISP_m; and 2) $Q_{p,i} > Q_{min}$. Note that we have to find out potential providers for each ISP in C_i using the two conditions stated above. And in fact, ISPs in the set P_i are competitors for each other.

Other ISPs which are not in these two sets are not directly related to the provider selection problem in terms of the flow T_i. It is possible for them to be on the transmission path of this flow. But they are not competitors and not potential providers of these competitors, which means they do not affect the bargaining power distribution of the value of T_i.

Again, let us take the scenario in Figure 2 as an example. In terms of the flow from the left person to Google, ISP_m has two competitors ISP_n and ISP_v which can cover the left person with acceptable performance. There are four ISPs in the set of potential providers/cooperators. ISP_m has three potential providers, and each of ISP_n and ISP_v has one. If there are potential peering cooperation relationships (i.e. one ISP satisfies the two conditions to be potential provider of the other), there will be a link between ISPs in C_i and P_i.

From this graph, we can further distribute the value of this flow T_i and compute $V_{j,i}$ for each potential provider ISP_j, which will be discussed in details in the next section. Then, we will have the total value of provider ISP_j to customer ISP_m.

$$V_j = \sum_{i=1}^{|T|} V_{j,i} \tag{3}$$

IV. VALUE BASED PROVIDER SELECTION FRAMEWORK

We have introduced ISPs' considerations on provider selection and describe the basic idea to solve the problem. In this section, we will describe how to implement these ideas and introduce the development of our value-based provider selection framework according to the consideration of decision-making criteria above. As we introduced before, we need to implement three algorithms to solve three key problems to help ISPs to select provider properly, i.e. identifying competitors, identifying cooperators, and value distribution mechanism.

A. Potential Competitors

To leverage the bargaining power analysis of transit ISPs, it is necessary to identify the potential competitors and providers of ISP_m.

As mentioned before, we regard an ISP as potential competitor if it can provide transit service with good performance to the end users of ISP_m. In other words, if the performance is good over the link between an ISP and the end users, we can recognize it as a competitor. Hence, we can directly find the competitors from the view of users straightly.

Considering performance evaluation of an ISP varies from different users,based on their IP prefixes, we divide all the users into *K* groups according to their locations geographically. Then, we launch measurment probes from servers in the *K* groups to all the transit ISPs in the world. In addition, to achieve a better performance evaluation, our method takes the content type of flow and user demand into consideration. As an example, we classify the application types into two categories: high-bandwith applications and high-reliability applications and related metrics are bandwith, loss rate correspondingly.

Good performance is recognized by comparison between measurement result and a threshold, which can be adjusted approximately.

B. Potential Providers

Different to the analysis of competitor selection, the potential providers should provide good transit service to the specific destinations of the end users and comply to the geographial constraints at the same time. Hence, after exempting ISPs "far from" ISP_m and its competitors, we need to assess the performance between left ISPs and the destination ISP of each flow. We can combine the results of measurements and estimation to leverage the performance assessment.

First, we can use Looking Glass tools to do the measurements. Looking Glass tools are deployed on border route servers of ISPs to help troubleshoot Internet-wide routing problems. The measurement, trough Looking Glass tool, also uses delay and loss rate as our metrics.

However, as mentioned above, not every ISP provides Looking Glass tool and we consider to apply other measurements or estimations to make the performance results more convinced. We use CAIDA AS rank [19] to estimate the performance, assuming larger ISP coming with better performance. The rank metrics includs its customer cone size, prefixes announced and so on. We then order the ISPs according to their rank accendingly.

Another estimation approach is based on the topology distances. Based on the Internet topology in both AS level [17] and router level [18], we obtain AS hops and router hops between each ISP pair. Then, we can estimate the performance of the communication by AS hops and router hops, assumming less hops leads to a better performance which is well accepted by other researches[3,4].

Finally, to each flow, we can consider the ISPs as potential providers which perform well in all the above three aspects.

C. Value Distribution of Certain Traffic Flows

With the information of potential competitors, providers and the cooperation possibility among them in geographical view, how to evaluate the contribution of each related ISP to certain flows is a key problem and a fair value distribution mechanism is necessary. In this section, we explore the desirable properties and derive the value distribution mechanism.

We consider *N* is a set consisted of ISP_m and its potential providers, competitors as Figure 2 illustrated. The mechanism is to analyze the contribution of each ISP to this end to end communication assuming that they have the knowledge all the possible connections. We treat this problem as a cooperative game and follow the notations of coalitional game to do formulization. Any non-empty subset S is defined as a coalition if ISPs in S form a subnetwork that can enable the communications between users. v(S) denotes the contributions of coalition S. The distribution mechanism is a vector denoted as ϕ ($\tilde{N}$, v) = { $\phi_1, \phi_2, ..., \phi_N$ } and ϕ_i ($\tilde{N}$, v) indicates the assigned value of ISP_i.

A fair distribution system should conform to the rules below:

Efficiency: The total value distributed should be equal to the value of the flow and no ISP should receive extra value.

$$\sum_{i \in \mathrm{N}} \phi_{\mathrm{i}}(\mathrm{N}, v) = v(N) \tag{4}$$

Symmetry: If contributions of two ISPs are the same, their value should be same.

$$\mathrm{v}(\mathrm{S} \cup \{\mathrm{i}\}) = \mathrm{v}(\mathrm{S} \cup \{j\}) \text{ for all S}$$
$$\longrightarrow \phi_i(N, v) = \phi_j(N, v) \tag{5}$$

Fairness: For any pair of ISPs, ISP_i and ISP_j, they should have same mutual contribution. It ensures that if the contributions of

ISP_i to all the coalition are larger than ISP_j, value of ISP_i should be higher than ISP_j.

$$\phi_i(N, v) - \phi_i(N\backslash\{j\}, v) = \phi_j(N, v) - \phi_j(N\backslash\{i\}, v) \quad (6)$$

Shapley Value is a famous model in cooperative game theory. It conforms to all the properties above. The Shapley Value distribution mechanism can be described as below.

$$\phi_i = \sum_{S \subset N\backslash\{i\}} \frac{|S|!(n-|S|-1)!}{|N|!}(v(S \cup \{i\}) - v(S)) \quad (7)$$

Based on the Shapley Value distribution mechanism, we explore our traffic value distribution mechanism. First, we should define the total value V of a certain flow which is related to many factors of the flow, i.e., destination, volume, content. We use volume as the flow value for general case analysis. Second, we design and implement the value distribution mechanism based on Shapley Value. We consider all the orders of $\tilde{N}$, denoted as *A(i)*. $A_l(i)$ indicates the *l*th element of the permutation *A(i)*. To each flow, as demonstrated in Figure 3, if $\{A_{l'}(i) \mid l'=1,2,...,l-1\} \neq S$ and $A_l(i) \cup \{A_{l'}(i) \mid l'=1,2,...,l-1\} = S$, we recognize $A_l(i)$ as the key to this coalition *S* and add 1/|N|! to the value of $A_l(i)$. Hence, considering all of the flows, the value of ISP_j to ISP_m is

$$V(j, m) = \sum_{i=1}^{|T|} \sum_{k=1}^{|S_j|} \frac{1}{|N|!} * T_i(Volume) \quad (8)$$

As line 6 in Figure 3, we addup all the value after the keys appearing instead of calculating all the permutations to reduce the complexity to O(2^n). For our analysis, as we can see in the experiments, the number of competitors under analysis in this stage is less than 10 in general cases. The complecity of this algorithm will not be a big concern.

Input: Volume of flow T_i, a random order A of participant ISPs in set N and their potential connectivity
Output: Distributed value of each participant ISP to flow T_i

Initialization: sortedNum = 0

```
1.   For k=sortedNum, …, len(A)
2.   {
3.       A[sortedNum-1]←→ A[sortedNum+k-1];
4.       if({A[l] |l=0,1,…, sortedNum-1}=S)
5.       {
6.           V(j,m,i)+=factorial(len(A)- sortedNum)*Volume;
7.       }
8.       else
9.       {
10.          Recursively calculation with sortedNum increment by 1
11.      }
12.      A[sortedNum+k-1] ←→A[sortedNum-1];
13.  }
```

Figure 3. Value Distribution Method for Each Flow

D. Provider Selection Framework

According to the analysis above, we can explore our value based provider selection framework. Our framework is built on the operational data of the ISP_m which mainly includes the routing tables and traffic data. Published data, such as Internet topology data and ISP geographical data is also necessary.

As Figure 4 illustrates, in our framework ISP provider selection process can be divided into several steps: First is the performance measurements and estimations. Second, it will begin the value analysis to each flow, including competitor selection, provider selection and value distribution and then obtain the transit value for each related ISP based on their value distributed in each flow. Further, considering the transit price, it will sort the potential providers according to the descending order of R(j, m) and evaluate their performance again one by one. If performance of ISP_j reach the standards or say Evaluation($Q_{d, j, m}$) > $Q_{d,max}$, ISP_j will be an recommended provider in our model.

```
// Data Preparation for Competitor Selection
1. foreach user group k of the understudy ISP_m
2.      foreach transit ISP_j
3.           measure the performance q^s_{j,m,b,k} and q^s_{j,m,r,k}
// Data Preparation for Potential Provider Selection
4. foreach ISP pair<ISPa, ISPb>
5.      measure or estimate the performance q^d_{a,b}
// Bargaining Power Analysis for Related ISPs
6. foreach traffic flow T_i with content type t
     // Competitor Selection
7.      foreach transit ISP_a
8.           k = the source user group T_i
9.           if(q^s_{j,m,t,k}is better than threshold)    recognize ISP_a as competitor
10.     Gather the geographical info. of competitors as constraint C1
11.     // Potential Provider Selection
11.     foreach transit ISP_a
11.          if(it complies to C1 && Evaluation(q^d_{a,desi}) ranks top level)
12.                recognize ISP_a as potential provider
     // Value Distribution
13.     Value Distribution(Volume_i) to each related ISP
14.     foreach related ISPa       V(a, m) += V(a, m ,i)
15.End for
// Provider Selection Result based on bargaining power analysis
16.Rank transit ISP_a according to the descending order of V(a, m)/P(a, m)
17.     if(Evaluation(Q_{d, j, m}) > Q_{d,max})      Recommend ISP_a to ISP_m
```

Figure 4. ISP Transit Provider Selection Framework

If all the potential providers could not meet the performance standard, we should do a partition to our traffic and adopt multihoming mechanism. For example, we can simply consider providers for domestic traffic T_D and international traffic T_I respectively.

V. Experiment on Provider Selection of CERNET

In this section, we implement our decision model and take CERNET as an example to demonstrate the executions and practical effects of our method.

CERNET is a nationwide ISP in China. It has 38 domestic PoPs. More than 100 universities and other research entities are connected to it. It is an important eyeball ISP which serves over 25 million users and becomes a major part of the Chinese Internet community. Recent years, China is very positive with a opening telecommunication policy. CERNET would like to connect with more ISPs to optimize its business at its international PoP in Hong Kong. But how to choose its providers intelligently is an important issue, and CERNET would like to see reasonable data analysis to make persuasive decisions.

A. Data Sets

Routing Table and traffic: We obtain two datasets from CERNET: CERNET international netflow traffic data and CERNET routing tables from Nov. 1 to Nov 7, 2012. To speed up our selection, we convert the IP addresses of each flow into their corresponding ASs according to CERNET routing table, which means that we have a traffic matrix at AS level. So one flow is in fact all traffic between an AS pair.

Geographical Data of ISPs: We use the data published by Internet eXchange Point (IXP) to achive geographical data of all ISPs. We can find out the locations of one ISP by summarizing all locations of the IXPs which the ISP attaches to. The data published on [16] includes 374 public IXPs and 1134 private facilities and we mark locations of ISPs based on it.

Internet AS Level Topology and the Relationships among ISPs: CAIDA infers the AS relationships among ISPs in the Internet, i.e. provider-customer and peer-peer. We use the data [17] published by CAIDA at the end of 2012 and consider the non-stub ISPs as transit ISPs.

Internet Router Level Topology and the Relationships among ISPs: CAIDA deploys a Internet router level topology discovery platform, named Skitter to plot Internet topology. We use its public topology snapshot [18] of Dec. 2012.

B. Performance Measurement

Competitor Selection

We classify the IP addresses in CERNET into 38 groups based on the prefixes of the 38 domestic PoPs of CERNET. Then, we launch our probes from the 38 servers at these PoPs to all transit ASs which appear in the Netflow data. We conduct measurement both on the link bandwith and loss rate.

The ISP QoS analysis results of traffic from Beijing to Seoul are presented in Figure 5. We take the best two ISPs, i.e., China Telecomm and China Unicomm as competitors of CERNET.

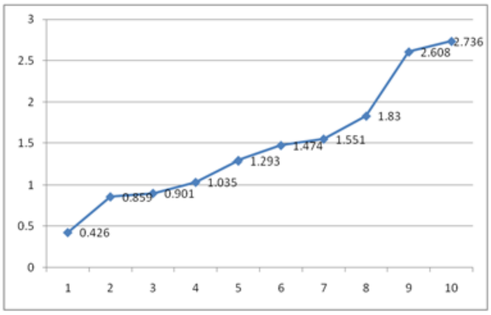

Figure 5. Performance Measurement Results to Find CERNET Competitors

Potential Provider Selection

We maintain a list of Looking Glass sites and query Looking Glass servers through a web-based interface using a script. We choose top 20 ISPs for further analysis and all ISPs without any available looking glass servers.

For ISPs without looking glass servers, we try to estimate their performance. We apply Dijkstra algorithm to the AS level topology graph and router level topology graph of the Internet to estimate the distance between an AS pair. We sort ISPs by AS hop number and router hop number assendingly and consider the top 20 ISPs as ISPs with good performance.

At last, for each flow, we select 10 ISPs with the highest AS rank as potential providers and add them in both List1 and List2.

C. Provider Selection and Evaluation

After settling down the potential providers and the competitors, we begin to distribute the value of each flow and calculate the final results as we introduced above. The values of top 10 recommanded ISPs to CERNET are demostrated in Figure 6.

ISP	AS#	Value(Normalized)
Pacnet	10026	0.243
Flag Telecom	15412	0.172
Deutsche Telecom	3320	0.162
TransTelecom	20485	0.14
Hutchison	9304	0.092
MTS OJSC	8359	0.082
INIT7	13030	0.076
GlobalCrossing	3549	0.025
Swisscom Ltd	3303	0.008

Figure 6. Recomanded Providers for CERNET

We can see Pacnet ranking the first followed by Flag Telecom. Pacnet, the current provider of CERNET, formed from the operational merger of Asia Netcom and Pacific Internet, owns PoPs in 19 cities worldwide, including HK, LA, NY and so on. The average AS path length between Pacnet and destination ISPs of CERNET traffic is just 2.12 while the average AS path length of other ISPs is 6.2; Traffic routed to Pacnet or its close neighbour ISPs (within 1 AS hop, we say "cover" below) accounts for 50.6% of the total traffic of CERNET. Flag telecom is another potetial transit provider for CERNET. As current provider of China Telecom, which is a competitor of CERNET, Flag telecom owns PoPs in a lot of cities, i.e., HK, NY, TYO. The average path length of it is 2.76 and it covers 45.5% traffic of CERNET.

Further optimization could be performed by dividing CERNET traffic as we introduce in Section 3. We divide the traffic by their destinations, i.e. North America Area, Asis & Pacific Area, Europe Area. We apply our model to the divided traffic data respectively and the results are shown in Figure 7.

Group#	North America	Europe	Asia &Pacific	Africa
Rank1	Pacnet	Deutsche Telecom	Pacnet	Trans Telecom
Value	128617712	92478196	357175775	2036255
Rank2	Flag Telecom	Trans Telecom	Hutchison	Flag Telecom
Value	79566776	71206369	266852740	1522615

Figure 7. Provider recomended for CERNET Devided Traffic

We can see the recomended providers are different for different flows. Pacnet contributes a lot to CERENT for the traffic to Asia & pacific and North America. Its average AS path length to the resouces in these areas is 2.01 and 1.92. However, for the traffic to Europe, European ISPs present their advantages. Deusache, with AS distance 2.26 to European resources, ranks the top among all ISPs.

From the analysis, we can see that the results are reasonable which can be supported from both theory and practice. We do not present more details on other ISPs in the recommended list due to commercial considerations.

VI. Conclusion

The provider selection problem is closely related to ISPs' economic interests. However, regional ISPs do not have a clear idea about which networks they should connect to and the price they should accept. Previously, they select providers empirically based on the Internet hirarchical architecture. With the development of Internet, it becomes a challenging task for ISPs and cannot be solved intuitively as before.

In this paper, we try to solve this problem for regional ISPs and help them find providers intelligently. We first explore the decision making criteria during ISPs' provider selection and then propose our decision model based on game thoery. In our model, we analyze each potential provider's bargaining power which we believe is a key factor in the evaluation of a transit provider. We distribute the transit value of each flow to related ISPs using Shapley Value distribution mechanism and obtain the bargaining value for each ISP by summing up their value of all flows. This value-based framework includes the influence of cost, performance, peering supply/demand relationships and so on.

We further instantiate our model and apply it to a national ISP CERNET. Based on the real world operational data, we try to find proper ISPs for CERNET and find out whether the result is consistent with our experience. The experimental results show the pracibility and effectiveness from the pratical point of view.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under Grant No. 61202356, Tsinghua University Initiative Scientific Research Program under Grant No. 2012Z02151, the National Natural Science Foundation of China under Grant No. 61170211, and Specialized Research Fund for the Doctoral Program of Higher Education under Grant No. 20110002110056.

References

[1] Subramanian, L.; Agarwal, S.; Rexford, J.; Katz, R.H.; , "Characterizing the Internet hierarchy from multiple vantage points," *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE* , vol.2, no., pp. 618- 627 vol.2, 2002.

[2] Norton W B. The art of peering: The peering playbook[J]. 2002

[3] Miao Li, Hui Wang, Jiahai Yang. Flattening and Preferential Attachment in the Internet Evolution. Network Operations and Management Symposium (APNOMS), 2012 14th Asia-Pacific, vol., no., pp.1,8, 25-27 Sept. 2012.

[4] Xiang Y, Yin X, Wang Z, et al. Internet flattening: Monitoring and analysis of inter-domain routing[C]//Communications (ICC), 2011 IEEE International Conference on. IEEE, 2011: 1-6.

[5] Labovitz C, McPherson D, Iekel-Johnson S, et al. Internet observatory report[J]. NANGO: NANGO47, 2009.

[6] Orda A, Rom R. Multihoming in computer networks: A topology-design approach[J]. Computer Networks and ISDN Systems, 1990, 18(2): 133-141.

[7] Wang H, Xie H, Qiu L, et al. Optimal ISP subscription for Internet multihoming: Algorithm design and implication analysis[C]//INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE. IEEE, 2005, 4: 2360-2371.

[8] Dhamdhere A. Provider and peer selection in the evolving internet ecosystem[M]. ProQuest, 2009.

[9] Bono C. Sharing the Cost of Backbone Networks[J]. 2012.

[10] Dhamdhere A, Dovrolis C, Francois P. A value-based framework for internet peering agreements[C]//Teletraffic Congress (ITC), 2010 22nd International. IEEE, 2010: 1-8.

[11] Norton W B. Internet service providers and peering[C]//Proceedings of NANOG. 2001, 19: 1-17.

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

[13] Stanojevic R, Castro I, Gorinsky S. CIPT: using tuangou to reduce IP transit costs[C]//Proceedings of the Seventh COnference on emerging Networking EXperiments and Technologies. ACM, 2011: 17.

[14] Ma R T B, Chiu D, Lui J, et al. Interconnecting eyeballs to content: a shapley value perspective on isp peering and settlement[C]//Proceedings of the 3rd international workshop on Economics of networked systems. ACM, 2008: 61-66.

[15] Ma R T B, Chiu D M, Lui J C S, et al. On cooperative settlement between content, transit, and eyeball internet service providers[J]. Networking, IEEE/ACM Transactions on, 2011, 19(3): 802-815.

[16] http://www.peeringdb.com

[17] The CAIDA AS Relationships Dataset, <2004-2012>, http://www.caida.org/data/active/as-relationships/

[18] The CAIDA Skitter Dataset, <2004-2012>, http://www.caida.org/data/active/Skitter/

[19] The CAIDA AS Rank Dataset, 2012, http://www.caida.org/data/active/as-rank/.