Analysis on Ideal Network Structures to Improve Reliability by Multi-path and Multi-homing BGP Routing in the Internet

Hiroshi Fujinoki

Department of Computer Science
Southern Illinois University Edwardsville
Edwardsville, IL 62026-1656
E-mail: hfujino@siue.edu

Abstract

We analyzed how reliability, defined as the probability of continuing transmissions on link failures, will be affected by inter-domain multi-path and multi-homing routing when the structure in the future Internet changes. The goal of this project is to find properties for the ideal network structure that maximizes the advantage of multi-path and multi-home routing. We focused on how each end-to-end path is built, how many multi-paths exist and how each multi-path consists of no multi-path and multi-path segments. The results of the analyses showed that multi-path and multi-home routing can improve the reliability by 10 to 30% in absolute probability of survivals on link failures, which were two to six times better than the existing no multi-path and no multi-home routing. After the analyses, several interesting properties are identified. It is important to keep the path length short to maximize the benefit from multi-path routing but a large number of multi-paths nor multi-homing connections are not necessary. The multi-homing configurations of degree three improved the reliability up to 50% for link failure rate up to 50% compared to degree of two in our analyses. It is found that single-path edge sections in a path should be short for multi-path routing. The results of this analysis can be a guide in structuring the future Internet.

Key Words: inter domain routing, network topology, reliable networks.

1. Introduction

Today's Internet handles huge volume of network traffic every single second. Its routing mechanism is responsible for efficiently delivering network traffic, which has significant impact to transmission rate, end-to-end delay, and network resource utilization. It is responsible also for reliability in a sense that ongoing transmissions will be sustained by detouring them through alternative routes if some of the links are down.

Border Gateway Protocol (BGP) has been the default routing protocol for routing network traffic in the Internet. As a path-vector routing protocol, BGP does not recognize whole path information for each route [1]. Each BGP speaker knows only the next hop AS to a specific destination. Although this property in BGP has been known to cause routing loops, there are two other significant problems. The first problem is incapability of multi-path routing. Even if more than one possible path exists between two end hosts, only one, no matter how many exist, will be selected for actual payload transmissions. Another path can be used only when the currently used path is down, but not at the same time. This prohibits multi-path routing, which allows network traffic to flow through more than one parallel path to a destination at the same time [2].
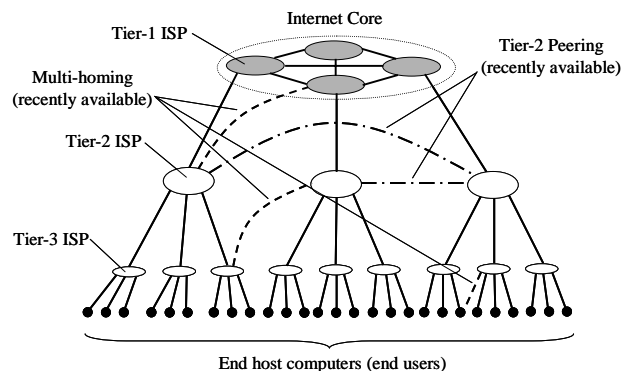

Figure 1. Simplified internal structure in the Internet

Lack of support for multi-path routing has not been a serious problem, since most of the Internet is structured as a tree and it is not structured to take the advantage of multi-path routing. Figure 1 shows the internal structure in the Internet. The Internet is known to have three layers of tier-1, 2, and 3 networks. Tier-1 layer mainly consists of world-wide long distance carriers and those that belong

to this group are called "tier-1 ISP's". Teir-2 ISP's are typically continent-wide or even country wide network carriers. Tier-3 ISP's are typically regional carriers. It used to be that tier-2 ISP's were connected only to a tier-1 ISP, but not to any other tier-2 ISP's. The important point here is that only the tier-1 layer has a network structure, while others form "stub" to the network of tier-1 ISPs.

The model shown in the figure has recently changed. Many tier-2 and 3 ISP's recently added connections to other ISP's in the same layer, mainly to reduce end-to-end delay to access information stored in content delivery networks. This type of connections is called "peering" (shown as dotted links in Figure 1). Peering is obviously invalidating the simple tree structure that has been used as an excuse for negating need for multi-path routing in the inter-domain routing.

The other problem is lack of BGP's support to multiple-path transmissions to multi-homed networks [3]. Multi-homed networks are tier-3 or 2 ISP's that are connected to more than one higher-layer network (e.g., a tier-3 provider connected to multiple tier-2 ISP's and a tier-2 provider connected to multiple tier-1 providers). When a network transmits network traffic to a multi-homed destination, existing BGP does not allow multi-path routing to a multi-homed network by the same reason as described for multi-path routing.

Support for multi-path routing becomes more important than ever for the following reasons. First, ability to sustain transmission on link failures (we defined this ability "reliability") becomes critical since many of the Internet users are now business users. Even a few minutes network down can cause tremendous financial losses for such users. The problem is that BGP does not support seamless route switching on a link failure. Although BGP is capable of detecting and switching to alternative paths on a link failure, on-going TCP connections will be dropped due to BGP's long convergence delay, which necessitates human end users to restart the transmissions [4, 5].

Second, the traffic load has continued to increase and the network resource is never enough to handle such huge volume of network traffic. Use of multi-path routing can maximize network hardware resource utilization by offloading excess network traffic from over-utilized paths to under-utilized ones. If BGP can dynamically divert burden on some already over-utilized resources to less utilized resources, it will realize Internet-wide load-balancing.

Third, tier-1 ISP's used to be the only networks that formed a network. Recent peering connects tier-2 and 3 ISP's as a network. This change implies that multi-path routing can be applied not only to the network of tier-1 ISP's but possibly to the entire Internet, warranting the benefits of multi-path routing to be performed as a part of BGP routing.

This paper focuses on studying network properties, such as the path length, number of multiple paths, degree of multi-homing and composition of multi-paths, that are expected to have impact to the performance of future multi-path routing in the Internet. The goal in this work is to find the properties that allow us take the most benefits of future multi-path routing, hoping that this work will be a guide to how to organize the structure of the Internet in the near future.

The rest of this paper is organized as follows. Section 2 discusses the existing related work. Section 3 describes how our analyses were designed and implemented. Section 4 presents observed results and analyses on the results. Section 5 summarizes the conclusions, followed by a list of the selected references.

2. Related work

The structure of the Internet, from a view point of interconnections of AS's, has been actively studied in the last decade. The common goal of the existing work has been to model the Internet structure by applying concepts in graph theory, such as degree of connectivity, its distribution and path length. For example, Faloutsos argued that the degree of AS connectivity follows power laws [6]. Zhang developed a technique to improve accuracy of modeling the AS-topology using BGP routing tables obtained from multiple different sources [7]. Subramanian [8] and Gao [9] made a step further by developing methods to infer different types of routing relationships, which adds more insight regarding the structure in the Internet. The concept of multi-path routing [2] and load balancing based on multi-homing [10, 11] have been introduced, but, by the best knowledge of us, there has not been a study that analyzed the impact of the network topology to multi-path and multi-homing, especially from a view point of enhancing reliability.

3. Descriptions of Analyses

In quantifying expected improvement for multi-path routing, we focused on reliability in data transmissions. We defined reliability in data transmissions to be the probability of continuing transmissions on link failures. Each data transmission route is assumed to consist of multiple links. It is also assumed that some parts of a route can have multiple paths.

The model we made for our analysis is shown in Figure 2. Node S indicates a source end host computer, while D indicates a destination end host. S is subscribed to an ISP, connected to the ISP's router, R_{15}. D is connected to multiple ISP's, connected to R_1 in its first

ISP and to R_7 in its second ISP. Each route from S to D consists of multiple links, each of which is terminated by different routers (not all such routers are shown in Figure 2).

Each route is assumed to consist of three logical sections. The first section corresponds to the source side's edge links, where there is no multiple path exists (the section is called "source edge section" hereafter). The second section is the one where multiple paths are available ("multi-path section"). The third section corresponds to the destination side's edge links ("destination edge section"). Destination edge section is same as the source edge section.
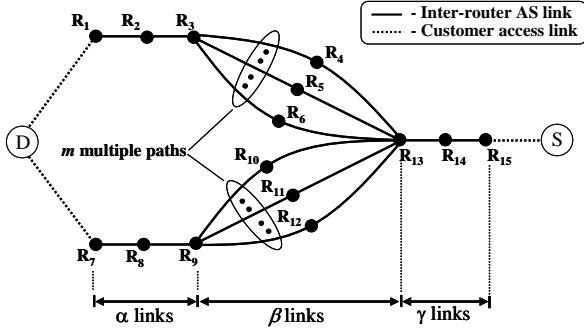

Figure 2. Model for reliability analysis using multi-path routing to a multi-homed destination

Based on the above model, the following control parameters are identified for our analysis:

- m: number of multiple paths available in the multi-path section
- n: degree of multi-homing
- α, β, γ: number of links for the source edge, multi-path, and destination edge section
- p: link failure rate for each link ($0 < p < 1$)

Using the parameters above, the reliability in data transmissions from S to D, as the probability of continuous transmission from S to D when each link fails at a probability of p, will be estimated as (1):

$$(1-((1-(1-((1-((1-p)^{\beta}))^{m})) \times ((1-p)^{\alpha}))^{n})) \times ((1-p)^{\lambda}) \quad (1)$$

We estimated the reliability in data transmissions for different m, n, and p, while we changed values for α, β, and γ, as well as their relative ratio to analyze their impact to the reliability mainly for the purpose of finding properties of network structure that allow maximizing the benefits from multi-path routing.

Analysis 1 ("path length analysis"): The impact of end-to-end path length to the transmission reliability was analyzed for different link failure rates. In the analysis, the transmission reliability was calculated using equation (1) while a set of α, β, and γ was changed from (α, β, γ)

= (1, 1, 1) up to (6, 6, 6). To increase the path length without modifying the ratio of the three sections, the same value was applied to α, β, and γ. For other parameters, m = 4 was applied and the reliability was calculated for three different cases of n = 1, 2, and 3.

Analysis 2 ("ratio of multi-path section analysis"): The transmission reliability was calculated while the length of multi-path section (β) was changed from 1 to 8 to see the effect of different ratio of β to α and γ. While β was changed, α and γ remained unchanged (α = 2 and γ = 2). Same as Analysis 1, m = 4 was used and the reliability was calculated for n = 1, 2, and 3.

Analysis 3 ("degree of multi-path analysis"): This analysis estimates the impact of the number of multiple paths available in the multi-path section. The expected reliability was calculated for m = 1 (no multiple path available) to 8. For other parameters, (α, β, γ) = (2, 2, 2) was applied. Similar to the previous analysis, the reliability was estimated for n = 1, 2, and 3. For a multi-homed environment (i.e., $n > 1$), the same value of m was applied to each multi-home connection to the destination network.

Analysis 4 ("ratio of no multi-path section analysis"): The reliability was estimated for different ratio of (α, γ) to β. While β was fixed, the values of α and γ were changed from (α, γ) = (1, 1) to (8, 8). β = 2 and m = 4 were applied to all the estimations in this analysis. The reliability was estimated for n = 1, 2, and 3.

4. Observed Results and Analysis

Observations from "path length analysis": Figure 3 shows the ranges of the link failure probability (p) that yielded 10%+ improvement in the raw difference between the reliability calculated for non multi-path routing (existing BGP) and multi-path BGP routing for three different levels of multi-homing (n = 1, 2, and 3). At (α, β, γ) = (1, 1, 1), the range of p that yielded 10%+ improvement for n = 1 (no multi-homing) was p = 0.14 through 0.52. For n = 2 (multi-homing with degree of 2), it was p = 0.06 through 0.69. For n = 3, it was 0.06 through 0.74. The range rapidly shrunk when the path length was increased from (1, 1, 1) to (3, 3, 3), followed by gradual, but monotonical, decreases to (6, 6, 6). At (6, 6, 6), the ranges of 10%+ improvement were reduced to p = 0.03 through 0.11 for n = 1, p = 0.01 through 0.17 for n = 2, and 0.01 through 0.20 for n = 3.

Analysis for "path length analysis": The results of Analysis 1 suggest that the path length should be short to maximize the advantage of multi-path routing. The width of p that yielded 10%+ difference for n = 1, 2, and 3 at (α, β, γ) = (1, 1, 1) was 0.38 (0.52-0.14) for n = 1, it was 0.63 and 0.68 for n = 2 and 3. The ranges of p shrunk to 0.15, 0.30, and 0.34 at (3, 3, 3) and 0.08, 0.16, and 0.19 at

(6, 6, 6) for n = 1, 2, and 3 respectively. However, the ratio of the width of p for n = 1 decreased as the path length increased from (1, 1, 1) to (6, 6, 6). The ratio of p for n = 1, 2, and 3 at (1, 1, 1) was approximately 1.0:167:1.74, while the ratio was 1.0:2.0:2.26 at (3, 3, 3) and 1.0:2.0:2.4 at (6, 6, 6). This implies that multi-homing will yield more benefit when path length is longer, while multi-path routing does so when path length is shorter.
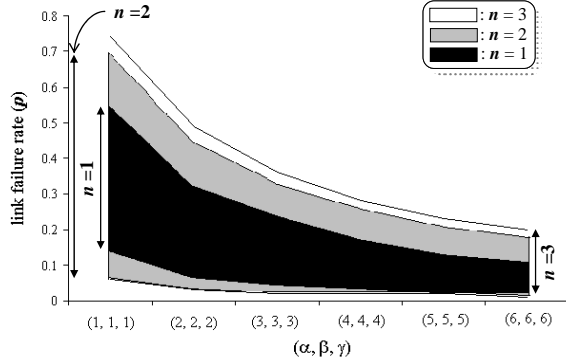

Figure 3. Ranges of link-failure rate (p) that yielded 10%+ improvement over non multi-path routing for Analysis 1

Observations from "ratio of multi-path section analysis": Figure 4 shows how the range of p that yielded 10%+ difference between no multi-path and multi-path routing changed when the ratio of the multi-path section was increased from (β: α, γ) = 1:2 through 8:2. When the ratio was 1:2, the ranges of 10%+ improvement was p = 0.04 through 0.44 for n = 2, and p = 0.04 through 0.50 for n = 3. For n = 1, the calculated transmission reliability was completely same as that for the no multi-path routing, which validated the equation (1). When the ratio was 8:2 (at β = 8), the ranges of 10%+ improvement were p = 0.02 through 0.26 for n = 1, 0.02 through 0.32 for n = 2, and 0.02 through 0.36 for n = 3. After β = 5, β did not significantly affect the ranges of 10%+ improvement.

Analysis for "ratio of multi-path section analysis": The width of p that yielded 10%+ improvement shrunk from 0.40 at (2, 1, 2) to 0.30 at (2, 8, 2) for n = 2 (that was only 25% difference from 0.40). For n = 3, the width shrunk by 26% from (2, 1, 2) to (2, 8, 2). For n = 1, there was no difference in the width from (2, 2, 2) to (2, 8, 2). The differences in the width of p were all minor compared to those observed in Analysis 1, where the differences were between 79 to 87%. For the effect to multi-homing, the ratio of the width of p for n = 1, 2, and 3 at (2, 2, 2) was approximately 1.0:1.67:1.92. The ratio was 1.0:1.25:1.42 at (2, 8, 2). Contrary to our prediction, these results suggest that the impact of the ratio of multi-path to the single-path sections will not significantly affect the

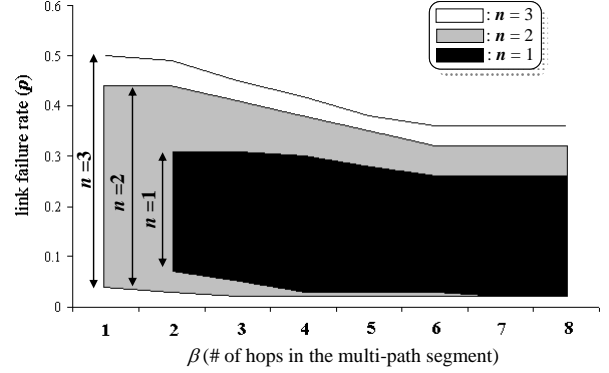performance of multi-path routing although multi-homing routing was more significantly affected.

Figure 4. Ranges of link-failure rate (p) that yielded 10%+ improvement over non multi-path routing for Analysis 2

Observations from "degree of multi-path analysis": Figure 5 shows how the range of p that yielded 10%+ difference changed when the degree of multi-path was increased from m = 1 to 8. When m = 1 (no multi-path available), the ranges were p = 0.04 through 0.28 for n = 2 and 0.02 through 0.36 for n = 3. The single-home configuration resulted in no improvement by the same reason for Analysis 2. At m = 5, the ranges for n = 1, 2, and 3 were 0.07 through 0.33, 0.03 through 0.45, and 0.03 through 0.50 respectively. After m = 5, increase in the range was minor (at most 2.4% increase from m = 5 to 6 for n = 2).
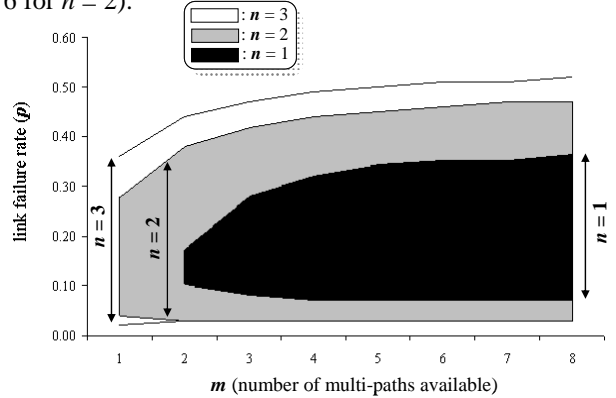

Figure 5. Ranges of link-failure rate (p) that yielded 10%+ improvement over non multi-path routing for Analysis 3

Analysis for "degree of multi-path analysis": The results imply that a relatively small number of multiple paths will significantly contribute to improvement of the reliability. However, the results also suggest that the reliability will not be improved proportionally to the number of multiple paths. At and above m = 4, the improvement slowed down. From m = 1 to 4, the range of

p increased for $n = 1$, 2, and 3 by 242.9% (from $m = 2$ to 4 for $n = 1$), 70.8% and 35.3%, while from $m = 4$ to 8, the increases were 16.7%, 7.3%, and 6.5%. These results suggest that a small number (three to four, especially three) of multiple paths will be most cost-effective.

Observations from "ratio of no multi-path section analysis": Figure 6 shows the ranges of p that yielded 10%+ difference when α and γ were changed from 1 to 8 (always the same value was assigned to α and γ) while the value of β was fixed to 2. At $(\alpha, \beta, \gamma) = (1, 2, 1)$, the ranges of 10%+ improvement were $p = 0.06$ through 0.51 for $n = 1$, 0.04 through 0.62 for $n = 2$, and 0.04 through 0.66 for $n = 3$. At $(\alpha, \beta, \gamma) = (3, 2, 3)$ and above, the single-home configuration did not improve the transmission reliability more than 10%. The transmission reliability for configurations $n = 2$ and 3 dropped rather rapidly (the ranges shrunk by 29.3, 26.8, 26.6, and 18.2% from $(1, 2, 1)$ to $(5, 2, 5)$ each time α and γ were increased for $n = 2$ and by 25.8, 26.1, 20.6, and 18.5% for $n = 3$). After $(\alpha, \beta, \gamma) = (5, 2, 5)$, the drop rate of the reliability slowed down though.
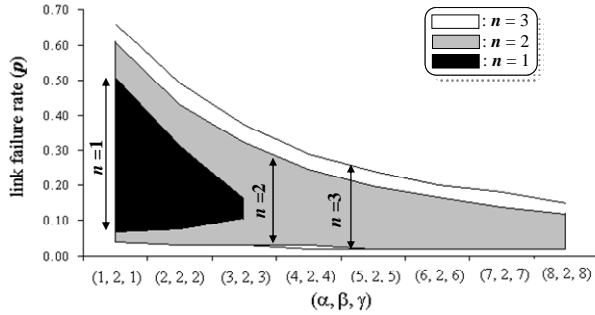

Figure 6. Ranges of link-failure rate (p) that yielded 10%+ improvement over non multi-path routing for Analysis 4

Analysis for "ratio of no multi-path section analysis": The results of this analysis indicate the difference between single-homed ($n = 1$) and multi-homed ($n = 2$ and 3) networks. As the length of the single-path sections between two end hosts (α and γ) increases, the advantage of multi-path routing rapidly decreased. When α and γ were four hops or more, the range of link-failure rate that produced 10%+ improvement disappeared for the single-homed configuration, while for the multi-homed configurations, the range slowly decreased.

Figure 7 shows the absolute and relative differences for the transmission reliability between the no multi-path routing and the three multi-path routing configurations ($n = 1$, 2, and 3) for $m = 4$ and $(\alpha, \beta, \gamma) = (2, 2, 2)$ for a range of p: $0.01 \leq p \leq 0.50$.

For the three multi-path routing configurations, the absolute differences were all positive (multi-path routing always resulted in a better reliability). The graph also shows that the relative differences monotonically increased for all the three multi-path routing configurations, although their absolute differences reached a peak while p was between 0.1 and 0.2. The same pattern was observed for all other configurations of m, α, β, and γ tested in Analysis of 1, 2, 3, and 4. The only differences observed were in the height of the peaks and the skews in the absolute difference.

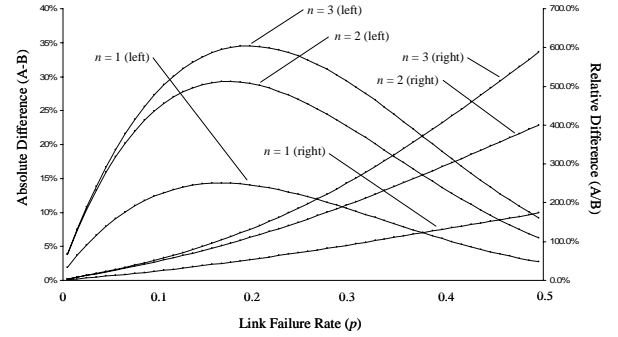**Figure 7**. Absolute and relative differences for the transmission reliability

5. Conclusions

After the four analyses, the following conclusions were drawn from the results of the analyses. To take the advantage of multiple-path routing, the end-to-end path length should be short (from Analysis 1). This implies that peering will be an effective solution to shorten path length by increasing the degree of AS-level connectivity.

The path length of the multi-path section (as notated "β") will not have a strong impact to the improvement of reliability (from Analysis 2 and 4). This means that, if the path length of the multi-path section increased especially when the degree of connectivity in tier-2 and 3 networks increases in the future, BGP multi-path routing will benefit mainly from increased degree of multiple paths without negatively impacted by increase in the path length (increased degree of connectivity will create paths that have longer path length).

A large number of multiple paths will not be necessary to maximize the advantage of multi-path routing (from Analysis 3). The results from Analysis 3 showed that three to four multiple paths will be most cost effective.

Increasing the degree of multi-homing for edge networks more than two will not be cost-effective. Although peering in tier-2 and 3 ISP's will be helpful because that will increase the degree of multiple paths, degree of multi-homing of two was more cost effective than three in our analyses. The multi-homing configurations of degree three improved the reliability up to 50% for link failure rate up to 50% compared to degree of two (Analysis 1, 2, 3, and 4).

One of the situations multi-homed connections benefit to multi-path routing will be when the single-path sections are relatively longer than the multi-path section (Analysis 4). This implies that multi-homing will be a good interim solution before degree of AS-level connectivity in the Internet becomes high enough for multi-path routing.

6. References

[1] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," *RFC-1771*, March 1995.

[2] M. Yannuzzi and X. Masip-Bruin, "Open Issues in Interdomain Routing: A Survey," *IEEE Network*, November/December 2005, pp. 49-56.

[3] Hiroshi Fujinoki, "Improving Reliability for Multi-Home Inbound Traffic: MHLB/I Packet-Level Inter-Domain Load-Balancing", *Proceedings of the International Conference on Availability, Reliability and Security*, March 2009, pp. 248-256.

[4] T. Griffin and B. Premore, "An Experimental Analysis of BGP Convergence Time," *Proceedings of the International Conference on Network Protocols*, 2001, pp. 53.

[5] D. Pei, X. Zhao, L. Wang, D. Massey, A. Mankin, S. Wu, and L. Zhang, "Improving BGP Convergence through Consistency Assertions," *Proceedings of IEEE INFOCOM*, vol. 2, April 2002, pp. 902-911.

[6] C. Faloutsos, P. Faloutsos, and M. Faloutsos, "On Power-Law Relationships of the Internet Topology," *Proceedings of ACM SIGCOMM*, September 1999, pp. 251-262.

[7] B. Zhang, E. Liu, D. D. Massey, and L. Zhang, "Collecting the Internet AS-level Topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, January 2005, pp. 53-61.

[8] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz, "Characterizing the Internet Hierarchy from Multiple Vantage Points," *Proceedings of IEEE INFOCOM*, vol. 2, June 2002, pp. 618-627.

[9] L. Gao, "On Inferring Autonomous System Relationships in the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, December 2001, pp. 733-745.

[10] F. Guo, J. Chen, W. Li, and T. Chiueh, "Experiences in Building A Multihoming Load Balancing System," *Proceedings of IEEE INFOCOM*, vol. 2, March 2004, pp. 1241-1251.

[11] R. Gummadi and R. Govindan, "Practical Routing-Layer Support for Scalable Multihoming", *Proceedings of IEEE INFOCOM*, vol. 1, March 2005, pp. 248-259.