A modeling method for router-level topologies based on network-cost optimization

Naoto Hidaka, Shin'ichi Arakawa, Masayuki Murata

Graduate School of Information Science and Technology, Osaka University, 1–5 Yamadaoka, Suita, Osaka 565–0871, Japan

Abstract

Measurement studies of Internet topologies show that the degree distribution of the topologies exhibits a power-law attribute. However, it is apparent that only degree distributions do not determine the structure of ISP topologies, where ISP designs router-level topologies based on their own design policies. Other structural properties than degree distribution are important to generate realistic Internet topologies. In this paper, we develop a modeling method for generating realistic ISP Internet topologies that obey a power-law degree distribution and have similar structural properties observed in the measurement studies. Our modeling method adds nodes one by one, and each node connects to optimal nodes to minimize overall network-cost. Then we answer the following question: what design factors are important to form realistic Internet topologies? Our results show that the answers are node locations and traffic demands.

Key words: power-law network, ISP router-level topology, topology generator, network-cost optimization, gravity model

1. Introduction

Measurement studies on Internet topologies have shown that the degree distribution of the topologies obeys a power-law for both Internet Service Provider (ISP) router-level and Autonomous System (AS)-level topologies [1, 2, 3]. Power-law degree distribution means that the probability that a node has k links is proportional to $k^{-\gamma}$, where γ is a constant value called *degree exponent*. In these topologies, a few nodes connect with many nodes, while most nodes connect with a few nodes.

Models of realistic Internet topologies are essential to study methods of network controls, such as routing control and congestion control, because topologies sometimes have a major impact on the performance of the network controls. A theoretical foundation to form power-law networks has been introduced [4], where Barabási and Albert presented their BA model in which the topology increases incrementally and links are placed based on the connectivities of the topologies.

However, even if the degree distributions of some topologies are the same, more detailed characteristics are often quite different. Li et al. [5] enumerated various topologies with identical degree distributions and showed the relation between their structural properties and performances. Li et al. pointed out that with the technology constraints imposed by routers, the degree of nodes limits the capacity of links that are connected to the node. Thus, to maximize the performance of router-level topologies, higher-degree nodes should be located at edges of the topology. They then demonstrate in an Abilene-based topology [5] that a power-law network can actually be constructed by maximizing the network throughput with the technology constraints imposed by routers.

Results of Ref. [5] indicated that topology performance does not only depend on the degree distribution. Actually, each ISP constructs its own router-level topology based on strate-gies such as minimizing the mileage of links, redundancies, and traffic demands, rather than focusing on the degree distribution. In a previous work [6], we showed that the structures of ISP topologies are quite different from those obtained by conventional modeling methods. More specifically, ISP topologies are highly clustered; a node connects two or more nodes that are also connected to each other. Thus, structural properties other than the degree distribution are important to generate realistic Internet topologies.

In Ref. [6], we developed a modeling method to generate ISP topologies. When a new node joins the network, it probably connects to the nearest nodes to minimize the mileage of links. In addition, we added new links based on node utilization in the topology that corresponds to enhance the performance of the network equipment in ISP networks. However, that paper does not discuss what design policies are crucial to form realistic ISP topologies.

Since ISPs have their own design policies based on budget constraints and customer satisfaction, it is important to unveil the essential design policies to form realistic ISP topologies. We therefore developed a modeling method where topologies are generated based on network-cost optimization. With our modeling method, important topology-related metrics such as clustering coefficients and the number of node pairs passing through links have almost the same as the actual ISP network by appropriate parameter settings, while still keeping the degree distribution of the topology to follow the power-law. We then investigated the structural properties of the generated topologies by changing three design factors: cost function for optimization, traffic demands, and node locations. The results show that all three factors are necessary for realistic clustering coefficients. We also show that the cost function and traffic demands differ by the distribution of the number of node pairs passing through links, while the node locations differs by the average

shortest path length.

This paper is organized as follows. In Section 2, we explain conventional modeling methods to generate topologies and discuss the issues of these methods when we use the generated topologies to evaluate methods of network control. In Section 3, we provide a new modeling method for ISP topologies based on network cost. In Section 4, we evaluate the topologies generated by our method and investigate the impact of changing cost function, traffic demands, and node locations on the structural properties of the generated topologies. Finally, we conclude this paper in Section 5.

2. Models of topologies

Many studies have focused on modeling methods for Internet topologies. In this section, we explain the conventional modeling methods of Internet topologies and discuss their problems when we apply the methods to evaluate methods of network control. We first introduce a Barabási-Albert (BA) model [4] in which the topology grows incrementally and links are placed based on the connectivities of the topologies to form power-law networks. We then introduce a Fabrikant-Koutsoupias-Papadimitriou (FKP) model [7] that incorporates geographical information to generate topologies.

2.1. Barabási-Albert (BA) model for power-law networks

Barabási and Albert proposed a BA model to generate topologies having a power-law degree distribution. The BA model is characterized by two features: *Incremental Growth* and *Preferential Attachment*. The model starts with a topology with a small number of nodes and works as follows:

Step 0 Make an initial topology that has m_0 nodes.

Step 1 *Incremental Growth*: Add a new node at each time step.

Step 2 *Preferential Attachment*: Connect the new node to m different nodes chosen with probability Π:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}, \qquad (1)$$

where k_i is the degree of node i.

2.2. Fabrikant-Koutsoupias-Papadimitriou (FKP) model

Fabrikant et al. [7] also presented a FKP model for generating topologies having a power-law degree distribution. The model also uses the incremental growth model, but the rules for link attachment are different from the BA model. The FKP model does not use preferential attachment to add links. Instead, it uses minimization-based link attachment. More specifically, the FKP model works as follows:

Step 0 Randomly place an initial node in a Euclidean space $[0, 1]^2$.

Step 1 Add new node i to the topology. Its location is also random in the Euclidean space $[0, 1]^2$.

Step 2 Calculate the following equation for each node j that already exists in the network:

$$D(j) = \alpha \cdot d_{ij} + h_j, \qquad (2)$$

where d_{ij} is the Euclidean distance (i.e., physical distance) between nodes i and j, h_j is the hop-count distance between node j and an initial node, α is a constant value.

Step 3 Select node j' that minimizes $D(j)$. Then connect nodes i and j', and go back to Step 1.

The FKP model introduces two distance-related metrics for attachment: the physical distance of nodes d_{ij} and the hop-count distance to the initial node h_j. The cost of attachment is the sum of these two metrics, but the physical distance is weighted by α. Depending on the value of α, the resulting topology has different characteristics. If α is a lower value, respective nodes seek to connect to higher-degree nodes. Especially when $\alpha = 0$, the resulting topology is star-like. If α is a higher value, the new node tries to connect to its geographically close nodes. In this case, the obtained topologies behave as an Erdös-Rényi (ER) topology that has a Poisson degree distribution [8]. A power-law attribute of degree distribution emerges when α is a medium value. The FKP model is further generalized in Ref. [9] so that AS-like topologies can be generated.

2.3. Structural properties by conventional models

In previous sections, we introduced two conventional models to generate power-law topologies. However, several problems exist when we use these models to generate Internet topologies. For example, topologies by the BA model are different from AS-level topologies in terms of the average shortest path length and clustering coefficient [10]. Topologies by the FKP model have many more nodes with single links than Internet topologies [11]. As a result, when we use the topologies generated by these models, we can not correctly evaluate the performance of network control methods like routing control [6]. These models focus on the degree or distance metrics. However, it is insufficient to consider only these metrics when we obtain realistic Internet topologies. Therefore, to correctly evaluate the performance of the network control methods in the Internet, we need realistic modeling methods for Internet topologies. Bu and Towsley [10] developed a variant of the BA model to generate topologies that have similar properties to the AS-level topology in terms of the average shortest path length and clustering coefficient. In this paper, we focus on ISP's router-level topology instead of AS-level topology.

3. A modeling method for ISP topology

To develop a modeling method for Internet topologies, we assume that Internet topologies are designed based on network-cost optimization. In this section, we first define the network-cost and design factors with which we were concerned in Sec.

3.1 and then introduce a modeling method for ISP Internet topologies in Sec. 3.2.

3.1. Designing topologies

We denote network-cost $C(V, E)$ for network $G(V, E)$. Here, V represents the set of nodes and E denotes the set of links in the network. Our design problem for topologies is to obtain a set of links and to determine the link capacity to minimize network-cost $C(V, E)$. To do this, we consider that the network cost is the sum of link costs of each link l. Since the link cost depends on the length and the link capacity, we introduce a cost function f_D, defined as a function of link distance d_l, and a cost function f_B, defined as a function of link capacity b_l Then we regard the product of f_D and f_B as the link cost. A formal description of our design problem is as follows:

$$
\begin{aligned}
\text{minimize } C(V, E) &= \sum_{l \in E} f_D(d_l) \times f_B(b_l), \\
s.t. \quad X &= \mathbf{R} \cdot \mathbf{T}, \quad (3) \\
b_l &= \min_{x_l \le b \in B} b, \quad (4)
\end{aligned}
$$

where $\mathbf{R}$ is a routing matrix of the network represented as $|E| \times |V|^2$. $\mathbf{R}_{l,(i,j)} = 1$ if the traffic demand between nodes i and j goes through link l, otherwise $\mathbf{R}_{l,(i,j)} = 0$. $\mathbf{T}$ is a traffic demand matrix of the network with an order of $|V|^2 \times 1$, where the $((i, j), 1)$-th entry of $\mathbf{T}$ is traffic demand between nodes i and j. Equation (3) determines $X (= (x_0, \ldots, x_l, \ldots, x_{|E|})^T)$, representing the amount of traffic on the links. Equation (4) determines link capacity b_l that is adequate to accommodate amount of traffic x_l passing through link l. In general, the link capacity is discrete, and an upper bound exists due to technology constraints. We therefore define a set of candidates representing link capacities B, and link capacity b_l is selected from B. In our modeling method, we consider 1, 10, 100 Mbps, 1, 2.4, 4.8 Gbps, or 10 Gbps as set of link capacities B. Note that due to this link capacity constraint, we may have to add more links to accommodate the traffic in generating topologies (See Step. 3 of Sec. 3.2).

The above formulation includes three design factors for generating topologies: 1) cost functions f_D and f_B, 2) node locations (that affect the value of d_l), and 3) traffic demands. In Sec. 4, we discuss the structural properties of the generated topologies by changing these three design factors.

3.2. A modeling method to generate realistic router-level topology

We next explain our modeling method based on network-cost optimization. The topology grows incrementally as the BA model does: a new router is added to the network one-by-one. Then we add links to minimize the network cost based on the formulation described in Sec. 3.1.

Unlike previously known modeling methods, our method takes the amount of traffic traversing links into consideration. If we connect a new router to existing routers and more than 10 Gbps, the maximum element in B, traffic passes through a link, and we add more links to divert the traffic to satisfy the link capacity constraint. The details of our algorithm are as follows:

Step 0   Set the number of nodes N and the number of links for each new node m. Make an initial topology that has m nodes and $m \cdot (m - 1)/2$ links.

Step 1   Add new node i.

Step 2   Repeat this step until node i is connected to m different nodes.

> Step 2.1   Calculate network-cost C_j and check the link capacity constraint when node i is connected to existing node j:
>
> $$C_j = C(V, E \cup (i, j)),$$
>
> where V and E are the sets of nodes and links in the topology at this time.

> Step 2.2   Select node u that minimizes network-cost C_j from nodes that satisfy the link capacity constraints, i.e., amount of traffic at all links is less than the upper bound of the capacities. If a node satisfies the above, connect node i to node u, and go back to Step 2. Otherwise, select node u that minimizes the network cost from the existing nodes, and connect node i to node u. Then go to Step. 3 to satisfy the link capacity constraints.

Step 3   Enhance the topology to reduce the amount of traffic on the links. Repeat Step 3.1 until the link capacity constraint on all links is satisfied.

> Step 3.1   Select a link (u, v) that conveys the largest traffic in the network. Calculate each network cost when node u (v) is connected to one of the neighboring nodes of node v (u). Then find a node pair that minimize the network cost and connect them. If all node pairs are already connected to themselves, perform this step for the link that conveys the next largest traffic.

4. Evaluations

In this section, we evaluate our modeling method by comparing structural properties of ISP topology with those of topologies generated by our method and show that our model can generate more realistic ISP topologies than the topologies by the BA and FKP models. In addition, we discuss how the structural properties of the topologies differ depending on the design by changing the cost function, the traffic demands, and the node locations.

For the experiment, we used a topology obtained by measuring the Sprint network [3]. The Sprint topology has 467 nodes and 1280 links, and thus we set N to 467 and m to 2. Then we generate the topologies and add links between the nodes based

on Step. 3 of our model until the topology has the identical numbers of links with the Sprint topology.

We used the following metrics to compare the structural properties of the topologies:

- *Clustering coefficient*: for a node, defined as

$$\frac{2E_i}{k_i(k_i-1)},$$

  where k_i is the degree of node i and E_i is the number of links between the neighbors of node i

- *Average shortest path length*: for a node, average length of shortest paths between node i and all other nodes

- *Betweenness centrality* [12]: for a link, defined as the number of node pairs that passes through the link when minimum hop routing with splittable flow is applied

- *Amount of Traffic on a link*: for a link and a given traffic demand matrix T, defined as the amount of traffic that passes through the link when minimum hop routing with splittable flow is applied

For the node locations, we used geographical information based on the DNS information of each router in the Sprint topology. We normalized the node locations of the Sprint topology into $[0,1]^2$ space from the geographic latitudinal and longitudinal information.

For the traffic demand matrix to generate the topologies in this experiment, we applied a variation of the gravity model [13] to obtain the traffic demand between nodes. The original gravity model is commonly used by social scientists to model the movement of people, goods, or information between geographic areas. In a geographic gravity model of cities, for example, the relative strength of the interaction between two cities is proportional to the product of the populations divided by the squared distance. The model provides accurate estimates of telephone traffic, whose fees increase depending on the communication time and communication distance exchanged between the cities. In our method, since our ISP fee is becoming flat-rate regardless of the communication traffic or distance, we use a distance-independent gravity model to obtain traffic demands. That is, the traffic demand between nodes are only proportional to the product of the populations of the nodes [14] and are given by the following equation: $t_{ij} = \alpha \cdot p_i \cdot p_j$, where t_{ij} is the traffic demand between nodes i and j, p_k is the population of node k, and α is a scaling parameter for a total amount of traffic in the network. We use population information p_k from [15, 16] and set scaling parameter α to a value so that the largest traffic by a link in the Sprint topology becomes 10 Gbps. We assume traffic between nodes passes through the shortest paths between the nodes.

Unless explicitly stated, we use the above node locations and traffic demand matrix to obtain the following results.

4.1. Structural properties of our modeling method

We first investigated the impacts of changing cost functions f_D and f_B on the structural properties of the generated topologies. Figure 1 indicates the structural properties of the topologies by setting link-cost $f_D \times f_B$ to $d_l \log b_l$, $d_l^4 \log b_l$, $d_l^8 \log b_l$, and $d_l b_l$. We also investigated the other link-cost functions, but the results of these functions have a similar tendency and thus are not presented here to facilitate readability of figures.

Let us first discuss the impacts of cost function f_B on the structural properties of the generated topologies. This can be observed from the results of $d_l \log b_l$ and $d_l b_l$ in Fig. 1. Firstly, changing cost function f_B does not give much impact on average shortest path length of the topologies. Secondly, the topology using a logarithm function has high clustering coefficients and a large variation of betweenness centrality as the Sprint topology does, while the topology using a linear function has low clustering coefficients.

This tendency is explained as follows. When a new node joins the network, our method connects it and the existing nodes based on the cost function. Suppose that by adding the new node, the amount of traffic on link l increases by β. Then the network cost related to link l increases by

$$f_D(d_l) \times (f_B(b_l + \beta) - f_B(b_l)),$$

where $f_D(d_l)$ represents the cost for link distance and $f_B(b_l + \beta) - f_B(b_l)$ represents the cost for link capacity.

When f_B is a logarithm function, capacity b_l has more impact on the increase of network cost than case when f_B is a linear function. Thus, the new node tends to connect with nodes already having larger b_l. Since the nodes having larger b_l also tend to be connected to each other, the generated topologies have high clustering coefficients. In addition, as the amount of traffic that traverses link l increases, more links are added around the link to divert the traffic (See Step. 3 of Sec. 3.2), which also leads to higher clustering coefficients. On the contrary, when f_B is a linear function, the cost for link capacity in the network cost is β regardless of b_l. The increase of d_l has more impact on the increase of network cost. Therefore, the new node tends to be connected with near nodes. However, closer nodes are not always connected; connectivity only depends on the distance between the closer nodes. As a result, the generated topologies lack such high clustering coefficients.

We next discuss the impact of cost function f_D on the structural properties by looking at the results of $d_l \log b_l$, $d_l^4 \log b_l$, and $d_l^8 \log b_l$ in Fig. 1. Changing cost function f_D has less impact on average shortest path length of the topologies, as changing cost function f_B does. The topology by setting f_D to d_l has high clustering coefficients, while setting f_D to d_l^8 creates topology with low clustering coefficients because the cost for link distance becomes dominant as the number of multiplications by d_l increases. Therefore, as the number of multiplications by d_l increases, new nodes tend to connect with near nodes to reduce costs of long distance links. As a result, nodes are connected to the nearest nodes that lead to low clustering coefficients, as the number of multiplications by d_l increases.

We next show the amount of traffic on links in Fig. 2. Randomly generated traffic demand matrix is used to calculate the

4

traffic on links for each topology. The figure shows that the result of $d_l^4 \log b_l$ is close to the result of original topology, and shows the same tendency as Fig. 1(c). We obtained the same tendency for betweenness centrality and the amount of traffic on links, so we omitted the result of the amount of traffic on links in further evaluations.

As discussed above, changing cost functions f_D and f_B also changes the structural properties of the generated topologies. However, our method can reproduce a topology that is close to the Sprint topology, especially in terms of the clustering coefficients and betweenness centrality, by setting $d_l^4 \log b_l$ as the link cost. In Fig. 3, we show the degree distribution of the Sprint topology and our topology with $d_l^4 \log b_l$ as the link cost. In this topology, the degree distribution obeys the power-law, as the Sprint topology does. We apply this topology and cost function to subsequent evaluations.

4.2. Comparison between modeling methods

Figure 4 compares our model with the BA and FKP models. For the results of the BA model, we generated 10,000 BA topologies and showed their averaged results. In each BA topology, we set parameter m to 2 and generated the topologies. Then we added links between nodes based on the probability of Eq. (1) until the numbers of links are the same as the Sprint topology. For the FKP model, we set 20 as weight α based on the discussion in Ref. [7] and connected two other nodes to the new node by repeating Step. 3 of the FKP model. We used the DNS information of each node in the topology as the Sprint topology and our topologies to the geographical information. After generating the topology, we added links between the nodes based on minimizing Eq. (2) until the topology has the same number of links as the Sprint topology.

We observe from Fig. 4 that our topology is similar to the Sprint network in terms of the clustering coefficients and betweenness centrality. However, the properties of the topologies by the BA and FKP models are much different from the Sprint topology. For instance, the clustering coefficients are completely dissimilar in Fig. 4(a), and the maximum betweenness centrality is not also reproduced in Fig. 4(c). In addition, in Fig. 4(b), the average shortest path length of the BA topologies is much shorter than the Sprint topology and the topologies by the other modeling methods.

4.3. Effects of traffic demand matrix

To see the impact of the traffic demand matrix on our modeling method, we show the results with our modeling method using a uniform traffic demands in Fig. 5. Firstly, changing the traffic demand matrix has less impact on the average shortest path length of the topologies. However, the generated topology using the uniform traffic demands has lower clustering coefficients and betweenness centrality. When the traffic demands are uniform, there is no difference between the traffic demands among the nodes, so the cost for link capacity has less impact on the network cost than that for link distance. Therefore, new nodes tend to be connected with near nodes. On the other hand, when the traffic demands are given by the distance-independent

gravity model, nodes among which demand is large tend to be connected with each other to prevent accommodating the large traffic indirectly. For the distance-independent gravity model, a new node tends to connect with nodes that have large populations. Since large-population nodes are also connected to each other, the clustering coefficient tends to be high.

4.4. Effects of node locations

Finally, we show the effects of the node locations on the structural properties of the topologies obtained by our modeling method. We generated 500 topologies by setting the node locations randomly in the square space $[0, 1]^2$, and the averaged results are shown in Fig. 6. The clustering coefficient and average shortest path length are much different from the Sprint topology. For example, when the nodes are located at random, the number of nodes whose clustering coefficient is 0 or 1 is smaller than that in the Sprint topology. The average shortest path length also increases when the nodes are located at random.

The reasons for these results are explained as follows. For the node locations based on the DNS information, the nodes of a region (e.g., city) are closely located and connected with short distance links. Among the nodes, one or few nodes become "gateway" nodes that connect regions by long distance links and accommodate the traffic between the regions to reduce the network cost. The existence of such gateway nodes connecting to the long distance links leads to a small-world property, as discussed in Ref. [17]. Thus, the average shortest path length of the topologies decreases in the case of the DNS-based node locations. On the contrary, for the random node locations, the nodes are uniformly distributed in the square space. In this case, the topology has less chance of connecting two nodes with the long distance links and therefore lacks the small-world property.

In summary, the following three design factors all impact the clustering coefficients of the obtained topologies and are important to generate realistic ISP topologies: 1) cost function, 2) traffic demands, and 3) node locations. Moreover, node locations affect the average shortest path length, and traffic demand matrix impacts the betweenness centrality for links.

5. Conclusion

In this paper, we developed a modeling method based on network-cost optimization to obtain realistic ISP topologies and investigated the crucial design policies to form realistic ISP topologies and how the structural properties differ by adding/removing the design policies. With appropriate parameter settings, our modeling method generated a topology with similar structural properties compared to the real ISP network, while still keeping the degree distribution of our topology to follow the power-law as the real ISP network. We also pointed out the importance of modeling cost function, traffic demands, and node locations. These three design factors greatly affected the clustering coefficients of the generated topologies. In addition, we revealed that node locations impacted the average shortest

path length and traffic demands affected the betweenness centrality.

Acknowledgment

References

[1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proceedings of the conference on applications, technologies, architectures, and protocols for computer communication (SIGCOMM '99)*, vol. 29, pp. 251–262, Oct. 1999.

[2] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, "Power laws and the AS-level Internet topology," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 514–524, Aug. 2003.

[3] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2–16, Feb. 2004.

[4] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.

[5] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the Internet's router-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 34, pp. 3–14, Oct. 2004.

[6] R. Fukumoto, S. Arakawa, T. Takine, and M. Murata, "Analyzing and modeling router-level Internet topology," in *Proceedings of the 21st edition of the International Conference on Information Networking (ICOIN 2007)*, Jan. 2007.

[7] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou, "Heuristically optimized trade-offs: A new paradigm for power laws in the Internet," *Lecture Notes in Computer Science*, vol. 2380, pp. 110–122, July 2002.

[8] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.

[9] J. I. Alvarez-Hamelin and N. Schabanel, "An Internet graph model based on trade-off optimization," *European Physical Journal B*, vol. 38, pp. 231–237, Mar. 2004.

[10] T. Bu and D. Towsley, "On distinguishing between Internet power law topology generators," in *Proceedings of Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, vol. 2, pp. 638–647, June 2002.

[11] N. Berger, B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, "Degree distribution of the FKP network model," *Theoretical Computer Science*, vol. 379, pp. 306–316, June 2007.

[12] L. C. Freeman, "Centrality in social networks: Conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.

[13] M. Roughan, A. Greenberg, C. Kalmanek, M. Rumsewicz, J. Yates, and Y. Zhang, "Experience in measuring backbone traffic variability: Models, metrics, measurements and meaning," in *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurment (IMW '02)*, pp. 91–92, Nov. 2002.

[14] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg, "Fast accurate computation of large-scale IP traffic matrices from link loads," *ACM SIGMETRICS Performance Evaluation Review*, vol. 31, pp. 206–217, June 2003.

[15] U.S. Census Bureau, "Census 2000 gateway," Available at `http://www.census.gov/main/www/cen2000.html`.

[16] United Nations Statistics Division, "Demographic and social statistics," Available at `http://unstats.un.org/unsd/demographic/`.

[17] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, June 1998.
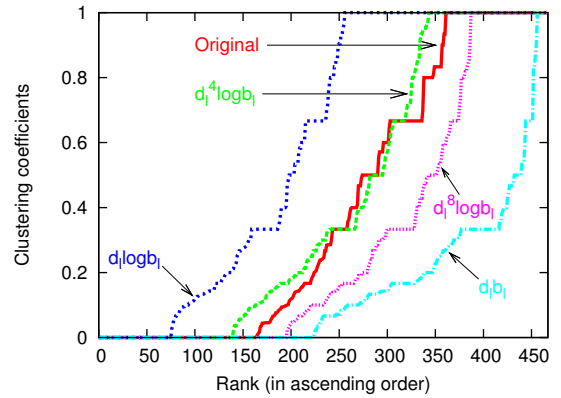
(a) Clustering coefficient
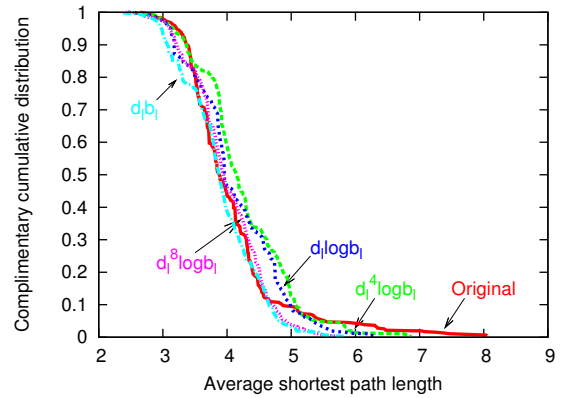
(b) Average shortest path length
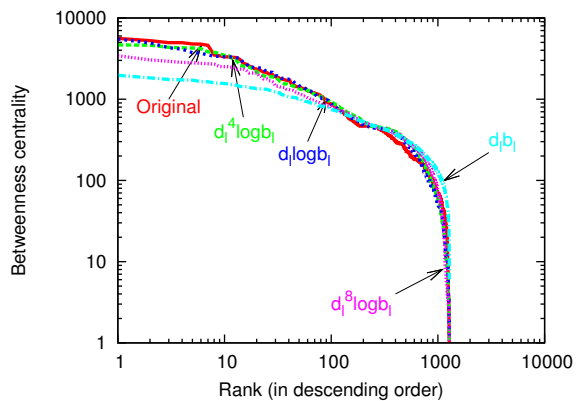
(c) Betweenness centrality

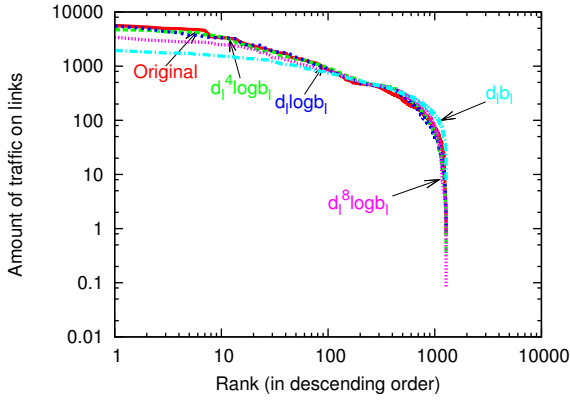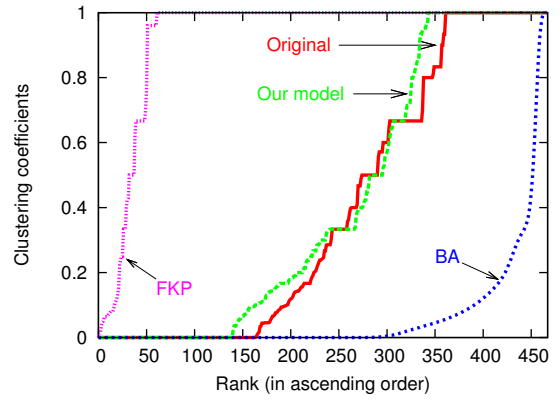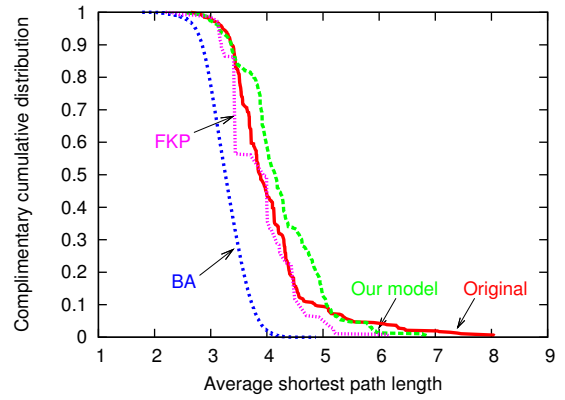Figure 1: Structural properties of our topologies

Figure 2: Amount of traffic on links when randomly generated traffic demand matrix is used. Total traffic demand is set to $N \cdot (N-1)$ where N is the number of nodes ($N = 467$).
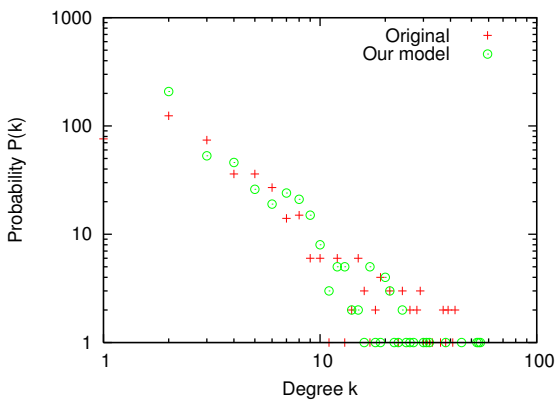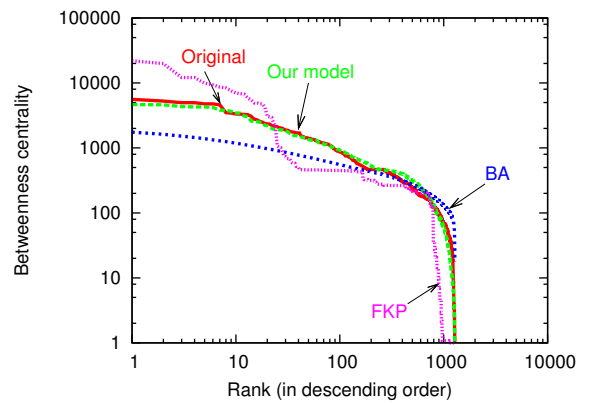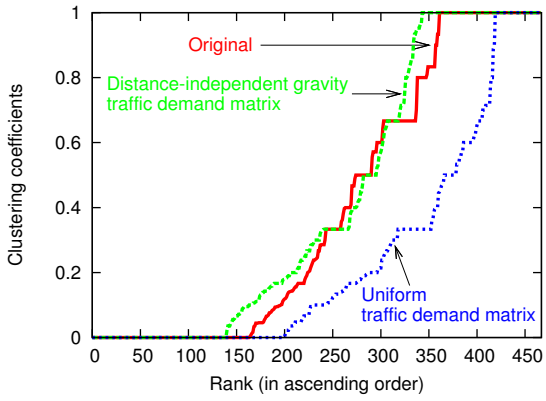

(a) Clustering coefficient

Figure 3: Degree distribution of our topology
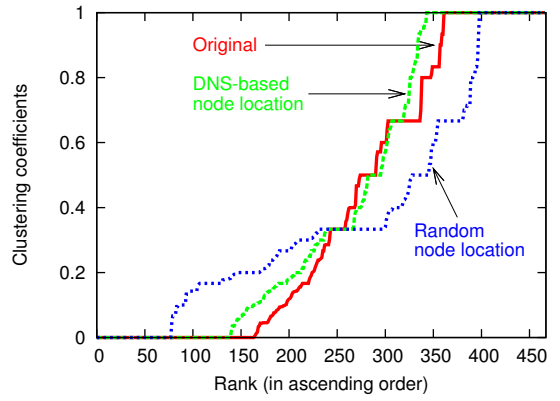

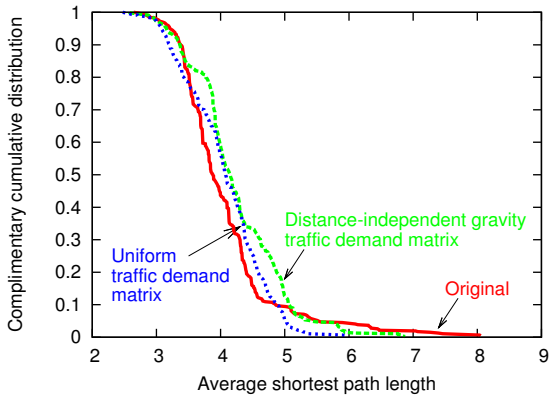(b) Average shortest path length

(c) Betweenness centrality

Figure 4: Comparison between our method and existing modeling methods
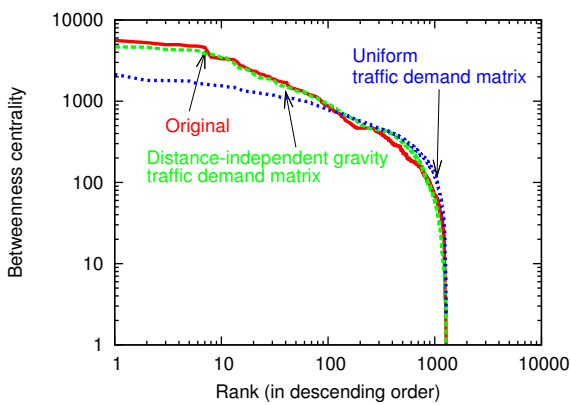
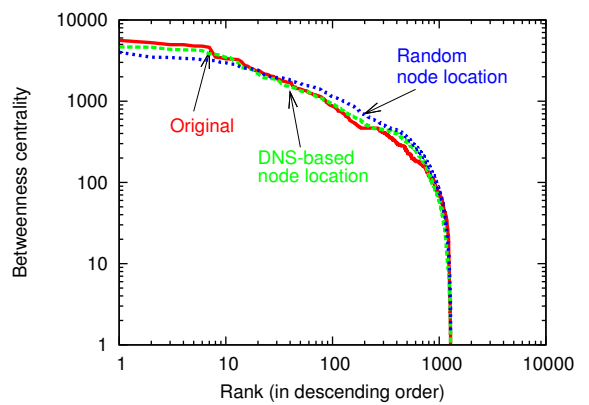(a) Clustering coefficient

(a) Clustering coefficient

(b) Average shortest path length
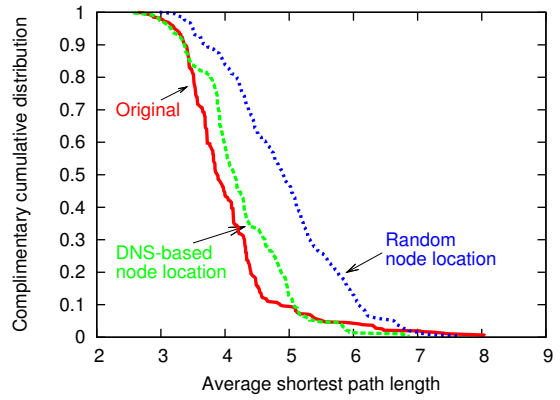
(b) Average shortest path length

(c) Betweenness centrality

(c) Betweenness centrality

Figure 5: Effects of traffic demand matrix on structural properties

Figure 6: Effects of node locations on structural properties