Analysis of the Collaboration Structure in Router-level Topologies

Yu Nakata*, Shin'ichi Arakawa*, and Masayuki Murata*
*Graduate School of Information Science and Technology
Osaka University
Osaka, Japan
{y-nakata, arakawa, murata}@ist.osaka-u.ac.jp

***Abstract*—As the Internet is one of infrastructures, the reliability of the Internet is becoming crucial to survive against failures of network equipment. Physical connectivity of the network is essential to characterize the reliability. There is collaboration structure, which is one of topological structures where two or more nodes are connected with a node, and collaboration is observed in transcriptional regulatory networks and ISP's router-level topologies. The collaboration structure relates to reliability of the network. Here, the main objective of this paper is to indicate whether the increase of collaboration improve the reliability or not. For this purpose, we first categorize the topology into three-level hierarchy; top-level layer, middle-level layer and bottom-level layer. It is apparent that the collaboration between top-level layer and middlelevel layer is much lower in the router-level topologies. We then calculate reliability of each network. The result indicates that reliability of most transcriptional regulatory networks is higher than one of router-level topologies. Finally we confirm that reliability of router-level topologies can be improved by rewiring to increase collaboration between top-level and middle-level layer.**

***Keywords*-power-law network; network reliability; transcriptional regulatory network; router-level topology; collaboration.**

I. Introduction

As the Internet becomes the social and economic infrastructure, the reliability is becoming crucial to survive against failures of network equipment. Many approaches to improve the reliability have been investigated either at the network layer [1] or more higher layer. The reliability is also investigated in the optical communication systems through the protection/restoration techniques [2].

These approaches greatly improve the reliability of the network, however, physical connectivity of the network is more essential to characterize the reliability of the network. That is, if the physical connectivity of the network is easily disrupted by network failures, the approaches to improve the reliability at network layer are no longer effective. In fact, the physical topologies used in the previous studies inherently assume that the physical connectivity is kept after the network failures. In order to design a reliable network, it is important to make the physical topology to be reliable against the network failures. For this purpose, the topological characteristic and topological structure that makes the physical topology more reliable is necessary to investigate.

As for the topological characteristic, Faloutsos et al. [3] demonstrates that the degree distribution of ISP's router-level topologies in the Internet exhibits power-law attribute, meaning that the existing probability $P(k)$ of degree k node that has k links is proportional to $k^{-\gamma}$. The modeling methods of router-level topology are investigated in Refs. [4], [5]. Barabási et al. presents the well-known BA model that generates topologies having power-law degree distribution [4]. Albert et al. also investigated the failure tolerant characteristics of BA topologies generated by the BA model [6]. The results show that BA topology has relatively few high degree nodes. Thus, a random failure of nodes will mostly remove low-degree nodes with little effect on the physical connectivity of the network. However, only the degree distribution does not determine the performance of router-level topologies. In Ref. [5], the author enumerates several topologies that have the same degree distribution but have the different topological structure, and then evaluates the amount of traffic that the network accommodates. Because of the constraints of router's processing capacity and the product lineup of commercial routers, the amount of traffic differs dependent on the topological structure of router-level topology. The results show that router-level topologies have the structure that high-degree nodes connect with low-degree nodes while the topology by the BA model has the structure that high-degree nodes are connected each other.

Li et al. [5] demonstrates that the topological structure greatly affects the network-level performance of router-level topologies. However, Abilene network examined in [5], which is one of scientific networks, is different to the other ISP networks [7]. The main difference may come from the fact that scientific networks like the Abilene network provide fewer opportunities to enhance their network because of budgetary constraints, while ISPs make their efforts on enhancement of network and/or reduce the traffic load on network. The difference on the redundancy of the topology can clearly be seen from the graphs of the Abilene network (Fig. 6 (e) of Ref. [5]) and the Sprint network (Figs. 7 and 8 of Ref. [8]). One of our motivations in this paper is to reveal a topological structure that makes the router-level topologies more reliable.

In this paper, we investigate a collaboration structure in router-level topologies. Here, the collaboration is one of topological structures where two or more nodes are connected with a node. The collaboration contributes robustness or reliability of topologies because it introduces multiple paths between nodes.

The collaboration contributes robustness or reliability of topologies because it introduces multiple paths between nodes. The network where the nodes are more collaborated has a chance to have larger number of paths between nodes. Note that unlike the clustering coefficient defined in [9], we consider the hierarchy of ISP topologies. That is, we first categorize the topology into three layers; top-level layer, middle-level layer, and bottom-level layer. We then investigate the collaboration between the three layers, which clarifies the topological structure of routerlevel topology. The definition of the collaboration structure is defined in Section II.

The collaboration structure is also investigated for transcriptional regulatory network [10]. The transcriptional regulatory network is one of the biological networks where transcription factors are regulated by the gene in the cell. In Ref. [10], the collaboration structure is investigated for several species. The authors show that the complex organism like human is more collaborated than the other organism such as *E. coli* or yeast. One of the main reasons to focus on the collaboration structure in router-level topologies is to answer the question that the current router-level topologies are designed well in terms of biological contexts. Our results show that there is a clear difference on the collaboration between the transcriptional regulatory network and the router-level topologies: The collaboration between top-level layer and middle-level is much lower in the router-level topologies. We therefore investigate the effect to increase the collaboration in the router-level topologies through the rewiring operation to discuss the future direction to design of router-level topologies.

This paper is organized as follows. In Section II, we define the collaboration in the networks. Section III presents the degree of collaboration and reliability against random node failures of router-level topologies, and compares with the collaboration in biological networks. Then, we investigate the effects of collaboration structure on the reliability by changing the physical topology through the rewiring process. Finally, we conclude this paper in Section V.

II. Collaboration in networks

A. Collaboration in biological networks

The collaboration structure in transcriptional regulatory network is investigated in Ref. [10]. In the cell of organisms, there is a transcriptional regulatory network consisted of transcription factors that are a kind of protein. The network transmits information to regulate genes depending on environmental insult. Collaboration in the transcriptional regulatory networks is a co-regulation relationship where two transcription factors regulate a transcription factor. According to the results of [10], more complex organism has more collaboration structure.

There are some analogies between transcriptional regulatory networks and router-level topologies. For example, the degree distributions of both networks exhibit power-low attribute. Another analogy is the hierarchical structure. In the transcriptional regulatory network, there are three level of hierarchy; top-level, middle-level, and bottom-level. Router-level topologies also have the hierarchy in the network; for example, a core network connects with several regions and/or state, a regional network, and an access network. We therefore investigate the collaboration structure in router-level topologies and show the difference, and then examine for changing the collaboration structure to discuss the future direction to design the router-level topologies.

B. Definition of hierarchy in router-level topologies

A key to identify the collaboration structure is to find a hierarchy, i.e., top-level, middle-level, and bottom-level in the router-level topologies. We define top-level, middle-level, and bottom-level nodes in the router-level topologies as follows. We first calculate H_i as the average hop-counts from a node i to the other nodes. Then, we set a directed link from node i to node j when H_i is lower than H_j. That is, when the node is located at the "center" of the network, the node tends to be the top-level node that belongs to the top-level. When the node is located at the "edge" of the network, the node tends to be the bottom-level node that belongs to the bottom.

More precisely, the top-level node is determined through the modularity analysis [11]. We divide the topologies into modules, and when the node has one or more links that connect with other module is classified into the top-level node. Note that, when there is a directional link from middle-node to top-level node, we reverse the direction of the link so as not to have links from lower level layer to the top-level layer. When there is a directional link between top-level nodes, we also change the directional link to be undirected. For the remaining nodes, nodes that have both incoming and outgoing links are classified into the middle-level nodes, and nodes that have only the incoming links are classified into the bottom-level nodes.

C. Definition of collaboration

Degree of collaboration is defined in [10]. It is the fraction of genes that are regulated by multiple transcription factors. In this paper, in order to investigate collaboration inside of topology, we adjusted the definition, that is, degree of collaboration is the fraction of nodes that are regulated by multiple nodes. Degree of collaboration does not depend on number of nodes and links. Bhardwaj et al. [10] introduces two types of degree collaboration. One is degree of collaboration in each layer D^L_{collab} and the other one is degree of collaboration between layers $D^{L_1,L_2}_{collab-betw}$.

1) Degree of collaboration in each layer: Degree of collaboration in each layer D^L_{collab} represents the average of D^i_{collab} for all nodes i in L-level, where D^i_{collab} is the number of nodes that are co-regulated by node i and the another node (A, for instance) divided by the nodes that are regulated by node i. The formal definition of D^i_{collab} and D^L_{collab} is as follows;

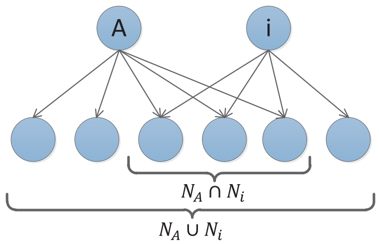

Figure 1. Collaboration between node i and node A: $|N_A \cap N_i|$ is the number of nodes regulated by node A and node i. $|N_A \cup N_i|$ is the number of nodes regulated by node A or node i.

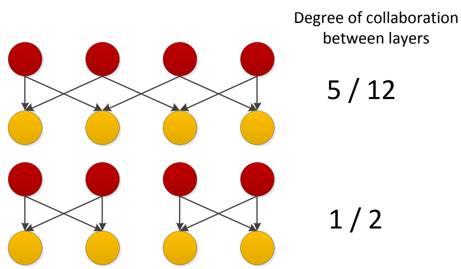

Figure 2. Illustrative example of differences on degree of collaboration between layers even when he number of nodes, links are the same. In the upper topology, the collaboration is $\frac{5}{12}$. In the bottom topology, the collaboration is $\frac{1}{2}$.

$$D^i_{collab} = \frac{\sum_{A \in N} |N_i \cap N_A|}{|N_i|}, \quad (1)$$

$$D^L_{collab} = \langle D^i_{collab} \rangle_i \quad \forall i \in L, \quad (2)$$

where N is a set of nodes in the network, and N_i is a set of nodes that are regulated by node i. Then, $|N_i \cap N_A|$ represents the number of nodes that are regulated by both node i and node A as shown in Fig. 1. $\langle\ \rangle$ represents arithmetic average.

2) Degree of collaboration between layers: Degree of collaboration between layers $D^{L_1,L_2}_{collab-betw}$ indicates fractions of nodes that are co-regulated by the node at L_1-level and the node at L_2-level, and is defined by the following equations.

$$D^{L_1,L_2}_{betw-level-collab} = \frac{\sum_{A \in L_1} \sum_{B \in L_2} \frac{|N_A \cap N_B|}{|N_A \cup N_B|}}{|L_1| \cdot |L_2|}, \quad (3)$$

where $|N_A \cup N_B|$ is the number of nodes regulated either by node A or by node B (see Fig. 1 for illustrative example). $|L|$ is the number of nodes including in L-level. Note that the collaboration defined by Eq. (3) depends on the number of nodes in L-level. To compare with several ISP topologies that have different numbers of nodes/links, we modify the definition of collaboration between layers to represents the number of collaboration:

$$D^{L_1,L_2}_{betw-collab} = \frac{|S_{L_1} \cap S_{L_2}|}{|S_{L_1} \cup S_{L_2}|}. \quad (4)$$

Figure 3 illustrates $S_{L_1} \cap S_{L_2}$ and $S_{L_1} \cup S_{L_2}$. $|S_{L_1} \cap S_{L_2}|$ is the number of nodes regulated by both a node including in L_1-level and another node including in L_2. $|S_{L_1} \cup S_{L_2}|$ is the number of nodes regulated by nodes including in L_1-level or nodes including in L_2-level.

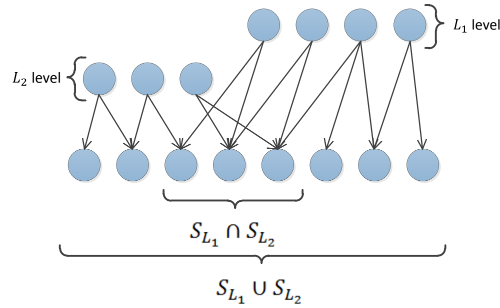

Figure 3. Modification of definition of degree of collaboration between layers. In this case, degree of collaboration between layers is $\frac{3}{8}$

III. Collaboration Structure and Reliability of Router-Level Topologies

A. Degree of collaboration

We first evaluate the collaboration structure in eight router-level topologies; AT&T, Sprint, Ebone, Exodus, Level3, Telstra, Tiscal and Verrio [8]. These topologies are obtained from traceroute-based network measurements, which may require alias resolution. The rocketfuel in Ref. [8] extended Mercator project's method [12] and relaxed the possibility of IP aliasing of routers to some extent. For comparison purpose, we compare the results of router-level topologies and five transcriptional regulatory networks, *E. Coli*, human, mouse, rat and yeast, and model-based two topologies (BA topology and ER topology [13]). For each topology, we calculate the hierarchy and then obtain the degree of collaboration in each layer and degree of collaboration between layers. Note that we do not show the degree of collaboration related to the bottom-level since nodes in bottom layer do not regulate other nodes by our definition of hierarchy.

We show the degree of collaboration in Figs. 4 and 5. From the results of router-level topologies in Fig. 4, we observe that difference between the degree of collaboration of top-level and the degree of collaboration of middle-level is less than 0.1. In contrast, the difference on transcriptional regulatory network is large in general. More distinctive characteristic of router-level topologies can be seen from Fig. 5. In the router-level topologies, the collaboration between top-level and middle-level is marginal, whereas it is not in the transcriptional regulatory network. One possible reason to have such the marginal collaboration is the functionality of middle-level nodes in the router-level topologies. That is, the traffic is first aggregated at the middle-level nodes and then forwarded to the top-level nodes. Thus, there are no consideration on load-balancing between top-level nodes and middle-level nodes.

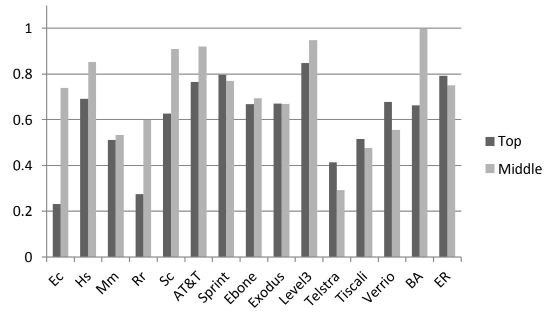

Figure 4. Degree of collaboration in each layer.

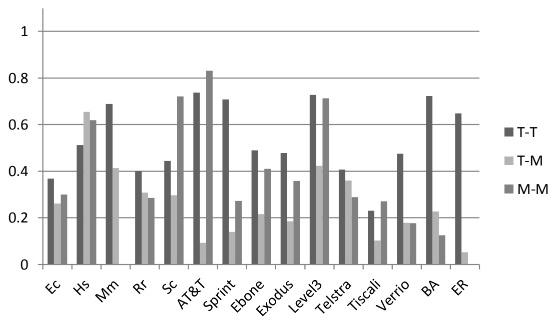

Figure 5. Degree of collaboration between layers.

B. *Reliability*

We next compare the reliability of router-level topologies with that of the transcriptional regulatory networks. The purpose of the comparison is to investigate how the degree of collaboration discussed in previous section relates to the reliability of networks. For this purpose, we consider the random node failures in each network, and we evaluate the ratio of nodes that are reachable from top-level nodes to the number of nodes in the network. Hereafter, we call the ratio as the *reachable node ratio*.

Figure 6 shows the reachable node ratio dependent on the failure ratio. The failure ratio is defined as the number of failed nodes normalized by the number of nodes in the original network. In obtaining the figure, nodes to fail are selected randomly from a set of nodes in top-level or middle-level nodes since bottom-level nodes are located at the edge of the network and removing them does not give the impact on the reachable node ratio. The results of the router-level topologies are depicted as solid-lines without symbols. We also depict the upper bound of the reachable node ratio in the figure. From this figure, we observe that the results of human (Hs), mouse (Mm), and yeast (Sc) is most reliable among the organisms that we investigated, and is close to the results of model-based topologies (BA topology and ER topology). Looking again at Fig. 5, we notice that these organisms exhibit a high collaboration between top-level nodes and middle-level nodes, which increases the number of alternative paths between top-level nodes and bottom-level nodes. That is, it is expected to construct more

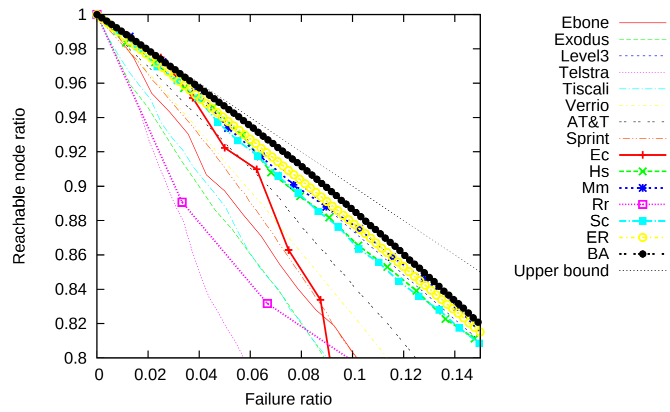

Figure 6. Reliability of networks for the random node failures

reliable network by incorporating such the collaboration structure. In the next section, we will discuss the effect of the collaboration structure on the reliability in detail.

IV. Effects of Collaboration Structure on the Reliability

In the previous section, we show that the transcriptional regulatory networks of human (Hs), mouse (Mm), and yeast (Sc) is most reliable among the organisms that we investigated, and find that these organisms exhibit higher collaboration between top-level nodes and middle-level nodes, while the router-level topologies exhibit lower collaboration between them.

In this section, we investigate effects of collaboration structure on the reliability. More specifically, we increase the collaboration between top-level nodes and middle-level nodes through the link-rewiring, and investigate differences before and after the link-rewiring. Note that the actual ISP network may increment links or capacity of links rather than rewiring the links. However, we still consider to rewire the link because our primal concern here is whether the increase of collaboration improve the reliability or not.

A. *Rewiring to increase the collaboration*

We explain the link-rewiring operation to increase the collaboration between top-level and middle-level. The operation consists from four steps as described below. The illustrative example of each step is shown in Fig. 7.

Step. 1 Find a node X regulated by three or more nodes in the same level. If several nodes are found, a node is selected randomly.

Step. 2 Select a node Y randomly from the several nodes that regulate the node X and that are in the same level.

Step. 3 When the node Y is the middle-level node, find a node Z that is co-regulated only by the top-level nodes. Otherwise, i.e., when the node Y is the top-level node, find a node Z that is regulated only by

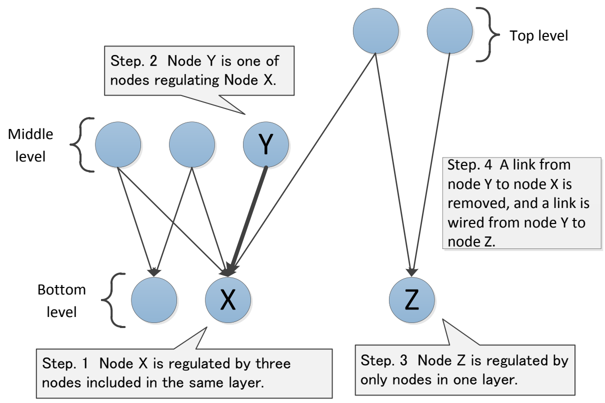

Figure 7. Illustrative example of the link-rewiring operation

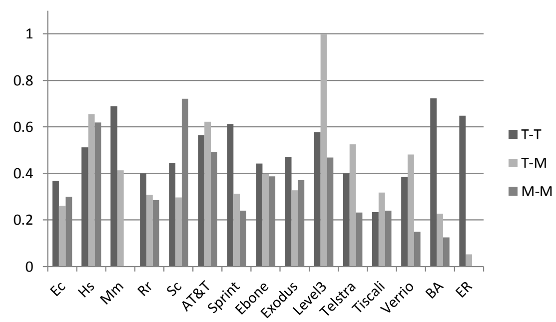

Figure 8. Degree of collaboration between layers after the rewiring operations

the middle-level node. If there are several candidate for the node Z, select the node Z randomly.

Step. 4 Rewire a link between node Y and node X; remove the link from node Y to node X, and wire a link from node Y and node Z.

Note that, in Step. 1, when the node X is regulated by only two nodes, rewiring the link leads to decrease the collaboration in the layer (middle-layer in Fig. 7) that the node Y belongs to.

This rewiring operation is continued until either of following termination conditions is satisfied.

1) There is no candidate for node X or node Z.
2) All nodes are regulated by both of top-level nodes and middle-level nodes. With this case, the collaboration is maximized, thus we do not need the link-rewiring.

The collaboration between layers after the rewiring operation is summarized in Fig. 8, and shows that the operation certainly increase the collaboration between top-level and middle-level.

B. Reliability of topologies after the link-rewiring

Lastly, we investigate the reliability of topologies after the link-rewiring that increases the collaboration between top-level and middle-level. Unlike Fig. 6 that evaluates the connectivity of directed network after the node failures, we present connectivity after the random node failures by using the un-directed links instead of directed links, and evaluate the difference between the original router-level topologies and the topologies after the link-rewiring. In particular, we use the *cover ratio*, which is defined as $\frac{S_i}{N}$ where S_i is the number of nodes in the largest strong component after i-th node failure, as the measure of the reliability.

Figure 9 shows the cover ratio of each topology after the link-rewiring. We examined 300 trials of random node failures and the average of them is plotted in the figure. We can see that the cover ratio improves for the most of router-level topologies except Sprint, Exodus, and Level3 which show a little improvement. The reasons are as follows. The original Level3 topology has a lot of links and is already collaborated well. That is, there is a little space to improvement. The situation for the Sprint and Exodus is different from Level3; the number of rewiring operations in the Sprint and Exodus is 59 and 14 respectively, while it is 221 in the AT&T topology. That is, the Sprint and Exodus have a little opportunity to increase the collaboration between the top-level and middle-level. The results of this section indicate that the collaboration structure of topologies characterizes the reliability, and the reliability improves to some extent by increasing the collaboration.

V. Concluding Remarks

In this paper, we investigated the collaboration structure in router-level topologies, and showed that the collaboration between top-level nodes and middle-level nodes in router-level topologies is less than that in the transcriptional regulated networks. Because of this, the connectivity of router-level topology is easily collapsed when node failures occur. In order to reveal the possible evolution path to improve the reliability of router-level topologies, we demonstrated that the reliability of several topologies improves when the collaboration between the top-level and middle-level increases. However, the improvement of reliability is limited in Sprint and Exodus topologies due to the limited opportunity of link-rewiring operations. That is, the evolution path to construct a reliable network is limited. Our future work is to identify the topological structure to prevent the link-rewiring opportunity in router-level topologies.

Acknowledgement

This work was partly supported by Grant-in-Aid for Scientific Research (A) 21240004 of the Japan Society for the Promotion of Science (JSPS) and the National Institute of Information and Communications Technology (NICT) of Japan. We also thanks to Dr. Nitin Bhardwaj who gives me helpful data of transcriptional regulatory networks and Mr. Shoi Shi of the Tokyo University of Science for fruitful discussion.

References

[1] B. Fortz and M. Thorup, "Robust optimization of OSPF/IS-IS weights," in *Proceedings of International Network Optimization Conference*, pp. 225–230, Oct. 2003.

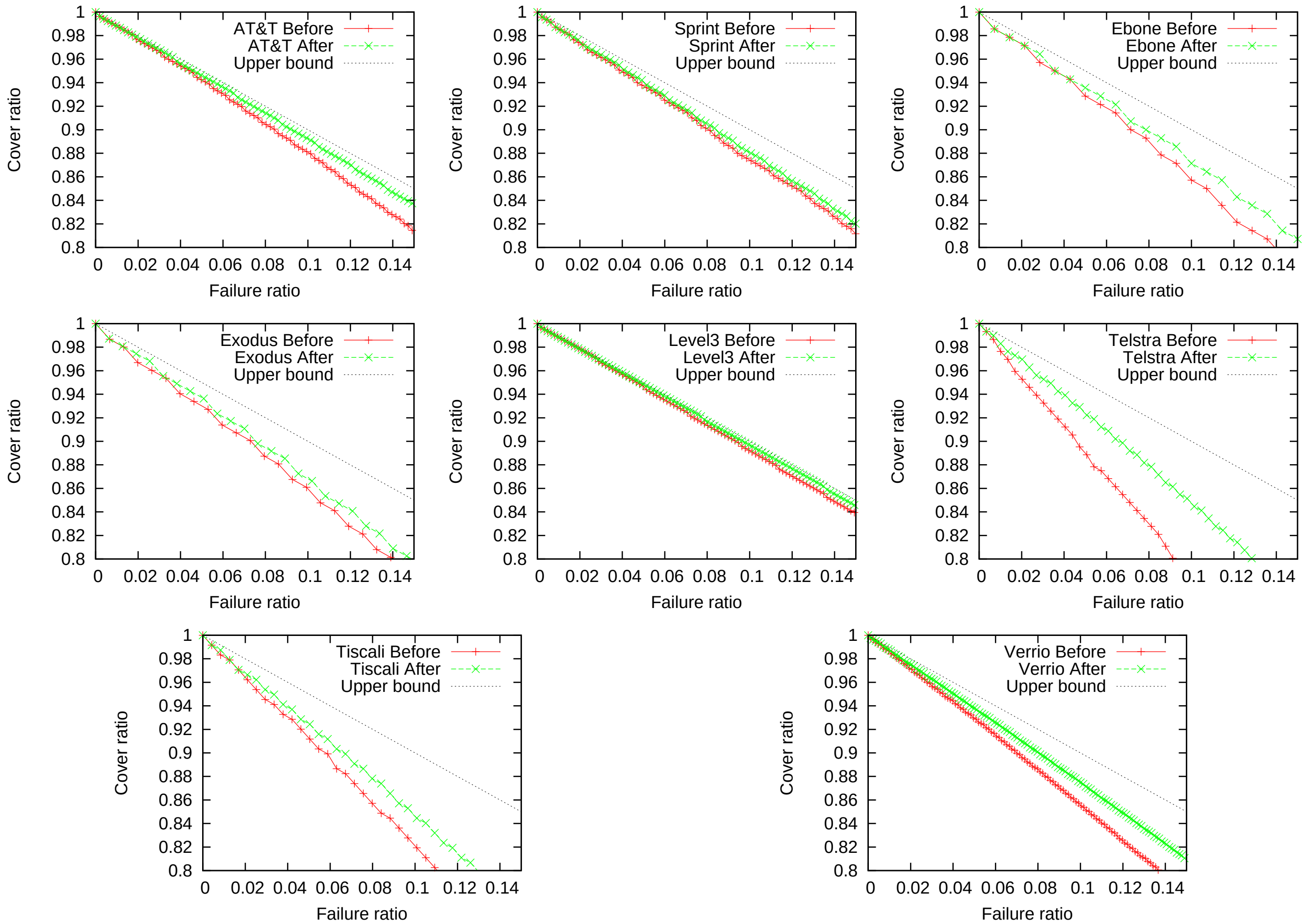

Figure 9. Difference of reliability between topologies before and after rewiring.

[2] R. Munoz, R. Casellas, R. Martinez, M. Tornatore, and A. Pattavina, "An experimental study on the effects of outdated control information in GMPLS-controlled WSON for shared path protection," in *Proceedings of ONDM*, Feb. 2011.

[3] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," *SIGCOMM Comput. Commun. Rev.*, vol. 29, pp. 251–262, Aug. 1999.

[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," *SIGCOMM Comput. Commun. Rev.*, vol. 34, pp. 3–14, Aug. 2004.

[6] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, pp. 378–382, July 2000.

[7] R. Fukumoto, S. Arakawa, T. Takine, and M. Murata, "Analyzing and modeling router-level internet topology," in *Proceedings of the International Conference on Information Networking*, pp. 171–182, Jan. 2007.

[8] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Trans. Netw.*, vol. 12, pp. 2–16, Feb. 2004.

[9] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, June 1998.

[10] N. Bhardwaj, K.-K. Yan, and M. B. Gerstein, "Analysis of diverse regulatory networks in a hierarchical context shows consistent tendencies for collaboration in the middle levels," *PNAS*, vol. 107, pp. 6841–6846, Mar. 2010.

[11] M. E. J. Newman, "Modularity and community structure in networks," *PNAS*, vol. 103, pp. 8577–8582, June 2006.

[12] K. Claffy, T. E. Monk, and D. McRobb, "Internet tomography," *Nature*, Jan. 1999.

[13] P. Erdös and A. Rényi, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.