Evaluating the AS-level Internet models: beyond topological characteristics*

Fan Zheng-Ping(范正平)†

Department of Automation, Sun Yat-Sen University, Guangzhou 510275, *China*

(Received 16 August 2011; revised manuscript received 31 August 2011)

A surge number of models has been proposed to model the Internet in the past decades. However, the issue on which models are better to model the Internet has still remained a problem. By analysing the evolving dynamics of the Internet, we suggest that at the autonomous system (AS) level, a suitable Internet model, should at least be heterogeneous and have a linearly growing mechanism. More importantly, we show that the roles of topological characteristics in evaluating and differentiating Internet models are apparently over-estimated from an engineering perspective. Also, we find that an assortative network is not necessarily more robust than a disassortative network and that a smaller average shortest path length does not necessarily mean a higher robustness, which is different from the previous observations. Our analytic results are helpful not only for the Internet, but also for other general complex networks.

Keywords: AS-level Internet, scale-free model, robustness

PACS: 89.20.Hh, 05.70.Ln, 87.23.Ge **DOI:** 10.1088/1674-1056/21/2/028902

1. Introduction

Performances of Internet algorithms and protocols are usually tested on an Internet model before their implementations. It has been observed that the Internet topology has a significant effect on the performances of almost everything processed on the Internet.[1−4] Therefore, it is very important and even absolutely necessary to determine which one is better from among the existing Internet models.

Currently, a widely used approach to determining whether a model is suited for the Internet is by examining whether it can reproduce some fundamental topological characteristics of the Internet. Goh *et al.*[5] and Tangmunarunkit *et al.*[6] used the power-law degree distribution to evaluate whether a model is suited for the Internet. Serrano *et al.*[7] compared their proposed model to the Internet in terms of clustering coefficient and assortative-mixing coefficient. Zhou[8] used the rich-club connectivity to evaluate the positive-feedback preference (PFP) model and other models. Piraveenan *et al.*[9] used local assortativity to validate autonomous system (AS)-level Internet models. Wang and Loguinov[10] compared the wealth-based evolution model with other models by examining whether they can reproduce the clustering coefficient and characteristic distance features presented in the Internet. Although a large number of models have been proposed and a variety of topological characteristics have been used to evaluate different Internet models, it is still difficult to judge which models are better to model the Internet.

To address this issue, in this paper we suggest a suitable Internet model that should at least be heterogeneous and have a linearly growing mechanism based on the observation on the AS-level Internet topology. A comparison of robustness performance in resisting intentional attacks between the Internet and different models shows that the approach using topological characteristics to evaluate Internet models may lead to some unreliable results. Also, it shows that some topological characteristics such as assortative-mixing feature and average shortest path length are not good indicators for characterizing the Internet resilience against intentional attacks.

2. Invalidating models by Internet evolving dynamical behaviours

Clearly, at the AS level, the Internet is an evolving self-organizing system. Therefore, a suitable Internet model should capture such evolving behaviours pre-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 60704045 and 60804012) and the Fundamental Research Funds for the Central Universities (Grant No. 09Lgpy57).

†Corresponding author. E-mail: fanzhp@mail.sysu.edu.cn

© 2012 Chinese Physical Society and IOP Publishing Ltd

sented in the Internet, instead of only capturing some features on a particular snapshot.

Figure 1 shows the evolving behaviour of network size of the real Internet and that predicted by the Fluctuation-driven model[5] which is an exponentially growing Internet model. Here, we use the Internet topology data[11] collected by University of California, Los Angeles (UCLA) because they can cover most ASs and almost all customer-provider links in the Internet. From this figure, we can see that from 2004 to 2008, a time period of total 5 years, the Internet has a linearly growing tendency. Clearly, the number of nodes predicted by the Fluctuation-driven model is much larger than that of the real Internet. Indeed, it is even beyond the maximum capacity that can be assigned by Internet Assigned Numbers Authority (IANA) at the end of 2008.

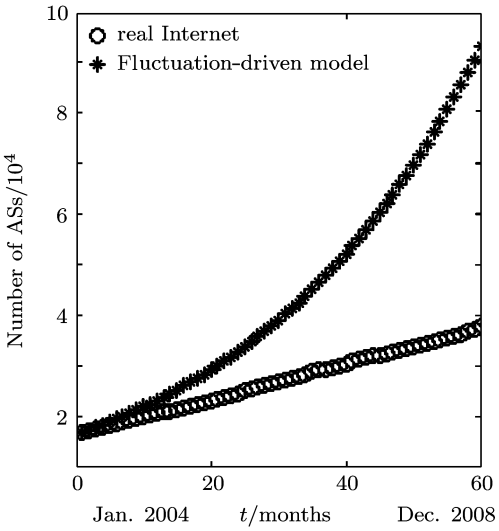


Fig. 1. Number of ASs in the Internet from 2004 to 2008 and that predicted by the Fluctuation-driven model.

Figure 2 shows the number of links of the real Internet and that predicted by the Fluctuation-driven model. From this figure, we can see that they have similar behaviours. Our observations on the linear growing tendency of the Internet confirmed this finding that the Internet growth has shifted from exponentially to linearly.[12] Therefore, those exponentially growing Internet models such as the Fluctuation-driven model should be excluded for modeling the Internet.

The previous observation[13] has shown that there are some invariances over the Internet evolution. Figure 3 shows the cumulative degree distributions of the Internet on the first day of every year from 2004 to 2011. We can see that although the Internet has grown remarkably in the past 8 years, with nodes increasing from 16573 to 46674, the power-law exponents are almost constant, showing only oscillation between 2.14 and 2.21. Although it was argued[14] that the Internet degree distribution does not follow exactly a power-law but a Weibull form or something else, it is surely that the Internet is heterogeneous, that is, a few nodes in the Internet have many connections while the majority of nodes has only a few connections. Such a heterogeneous feature of the Internet can be also observed by using other data sources.[15] Therefore, those homogeneous models such as the random model are excluded for modeling the Internet.

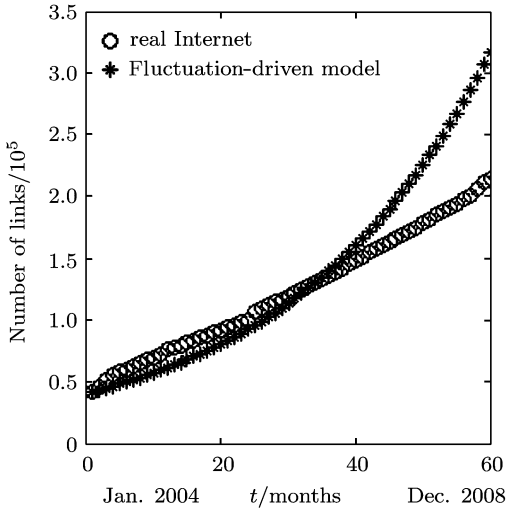


Fig. 2. Number of links in the Internet from 2004 to 2008 and that predicted by the Fluctuation-driven model.

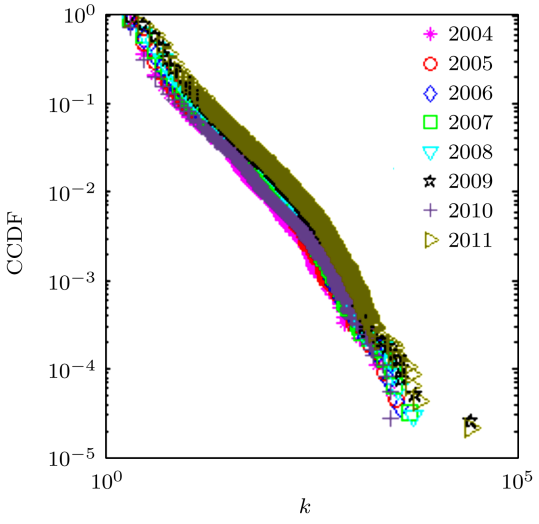


Fig. 3. Cumulative degree distributions of the AS-level Internet on the first day of every year in a time period from 2004 to 2011.

3. Evaluating Internet models in terms of robustness performance

Although topological characteristics such as power-law distribution play an important role in evaluating Internet models, their abilities to distinguish models are not necessarily unlimited. Indeed, the Internet is essentially an engineering network and it is particularly designed to perform some tasks and functions.[16,17] From this perspective, a suitable Internet model should behave as the Internet in terms of performances, instead of only capturing some topological characteristics of the Internet. Therefore, the issue on whether models are suited to sustain the Internet performances should be considered with a high priority when judging Internet models. Unfortunately, in most work this issue is never touched and only topological characteristics are used to evaluate Internet models. For example, the average clustering coefficient and the average shortest path length are used to validate different Internet models;[14] the PFP model is thought as the best Internet model because it can reproduce some topological characteristics, particularly the Rich-club connectivity.

Note that failures frequently occur in the today's Internet, owing to some intentional operations such as distributed Denial-of-service (DoS) intentional attack. Thus, here the network performance in resisting intentional attacks is taken as an indicator to judge whether models are properly evaluated by topological characteristics.

Figure 4 shows the performances of the Internet in resisting intentional attacks against the Barabasi–Albert (BA) model,[18] extended Barabasi–Albert (EBA) model,[19] Fitness model,[20] multi-local-world (MLW) model,[21] positive-feedback preference (PFP) model and Random model. As usual, the ratio S_f/S_0 is used to measure the communication capability of the network, where S_f represents the size of the largest connected component after a fraction of nodes, f, has been intentionally removed from the network.

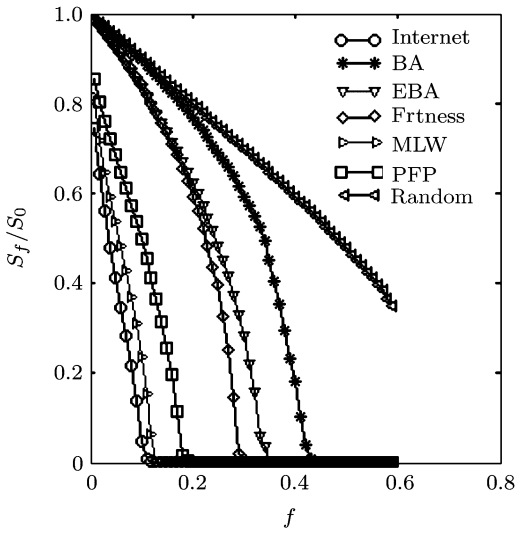


Fig. 4. Robustness comparisons between the Internet and other different models.

For comparison, the UCLA data collected on Jan. 1 2008 are used to evaluate different Internet models. The heterogeneous models considered here, namely BA, EBA, Fitness, MLW and PFP models, if possible, will use the same values as the observed Internet in three significant characteristics: average degree, degree distribution and joint degree distribution. The reason is that these three metrics are responsible for characterizing most fundamental topological characteristics of the Internet.[21] For the random model, it will have the same numbers of nodes and links as those of the Internet. Some topological characteristics of the Internet and all models considered are summarized in Table 1. It should be emphasized that our goal here is not to claim which model is the best to model the Internet. As a result, it has no effect on our conclusions to whether or not to tune the model parameters to best fitting to the Internet. It will become clearer as our presentation develops.

Table 1. Topological parameters for the Internet and models considered. N represents the number of nodes, $\bar{k}$ the average degree, γ the power-law exponent, r the assortative-mixing coefficient and $\bar{d}$ the average shortest path length.

	Internet	BA	EBA	Fitness	MLW	PFP	Random
N	33297	33297	33297	33297	33297	33297	33297
$\bar{k}$	10.44	10.00	10.42	10.00	10.24	10.00	10.45
γ	2.21	3.0	2.83	2.43	2.23	2.23	null
r	–0.15	–0.02	0.03	–0.08	–0.03	–0.15	0.00
$\bar{d}$	3.42	3.99	3.95	3.44	3.60	3.29	4.70

From Fig. 4, we can see that heterogeneous and homogeneous models are remarkably different in resisting intentional attack. Clearly, a phase transition exists for the Internet and heterogeneous models but not for the Random model. Again, it indicates that homogeneous models should be excluded for modeling the Internet because they cannot represent the Internet in terms of not only topological characteristics but also network performances such as robustness in resisting intentional attacks.

For heterogeneous models, we can also see from Fig. 4 that the PFP model is worse than the MLW model in representing the Internet performance of resisting intentional attacks. However, the PFP model is better than the MLW model in representing the Internet topological characteristics, characterizing by the average degree, power-law degree exponent and joint degree distribution as shown in Table 1. It means that although a model can capture some fundamental topological characteristics of the Internet, it cannot represent the Internet in terms of some certain performances. Indeed, an Internet model can present only one subset of the Internet topological features. Meanwhile, the relation between the Internet topology and the concerned network performance is unknown. As a result, the approach to the evaluation of Internet models by using topological characteristics is questionable though it has been widely used in both physics and engineering communities. Using such an approach without constraints may thus lead to some incorrect conclusions. In this sense, the effects of topological characteristics on evaluating Internet models are overestimated, particularly in the engineering community. For example, the performance of a new routing protocol is usually tested by simulation tools such as Opnet and Network Simulator. However, the models embedded in such simulation tools are usually determined by examining whether they can reproduce some topological characteristics of the Internet.[22] Obviously, the reliability of the obtained results is questionable because the used model may have a remarkable difference in routing performance from the Internet, despite capturing some topological characteristics of the Internet.

The robustness comparison between the Internet and other models, shown in Fig. 4, can also allow us to investigate the effects of topological characteristics on the Internet resilience against intentional attacks. A previous study[23] suggested that the robustness of a network can be enhanced by increasing its assortative-mixing feature. However, we can see from Fig. 4 that although the PFP model is more disassortative than the MLW model, it is more robust than the MLW model. It indicates that an assortative network is not necessarily more robust than a disassortative network in terms of resisting intentional attacks. On the other hand, the PFP model has less connectivities than the MLW model. It is confirmed that deleting some links can surprisingly enhance the network robustness to a certain extent.[24] This observation suggests that there are other factors that can significantly affect the Internet resilience against intentional attacks. For example, the power-law degree distribution is thought as one of the most important factors for the network resilience against intentional attacks.[25] From Fig. 4, we can see that the feature of power-law degree distribution is surely important for the Internet performance in resisting intentional attacks. However, we can also see that the PFP and the MLW models, which have the same power-law exponent values, each can have a mutually different behaviour in resisting intentional attacks. It means that the power-law degree distribution solely cannot characterize the robustness of the Internet against intentional attacks. The reason is that two networks with the same degree distribution may be remarkably different in other topological characteristics,[26] which can significantly affect the ability of the Internet to resist intentional attacks. In the recent work[25,27] it was suggestd that the smaller the average shortest path length, the better the network performance in resisting intentional attacks will be. From Fig. 4, we can see that the Internet has the worst performance in resisting intentional attacks, in all the heterogeneous models. However, the average shortest path length of the Internet is smallest, except for the PFP model, which is shown in Table 1. Therefore, a smaller average shortest path length does not necessarily mean a higher resilience against intentional attacks, which is very different from the previous observations.

4. Conclusions

In this paper, we suggest that from topological characteristics perspective, it is better for an Internet model to be heterogeneous and have a linearly growing mechanism. We are completely clear that these two factors are not enough to model the Internet. However, it is an important step to build a more suitable Internet model. A comparison of robustness

performance between the Internet and different models shows that the effects of topological characteristics on evaluating Internet model may be over-estimated, particularly in the engineering community. Such a comparison can also allow us to investigate the effects of topological characteristics on the internet resilience against intentional attacks. We show that an assortative network is not necessarily more robust than a disassortative network. Interestingly, a smaller average shortest path length does not necessarily mean a higher robustness, which is different from the previous observations. In our investigation, the Internet is treated as a graph network, regardless of its nature. Therefore, our discussion can easily be extended to other general complex networks.[28]

References

[1] Adamic L A, Lukose R, Puniyani A and Huberman B A 2001 *Phys. Rev.* E **64** 046135

[2] Li Y, Cao H D, Shan X M, Ren Y and Yuan J 2009 *Chin. Phys.* B **18** 1721

[3] Liu Y L, Zhu J and Luo X S 2009 *Chin. Phys.* B **18** 3772

[4] Ding D W, Zhu J and Luo X S 2008 *Chin. Phys.* B **17** 105

[5] Goh K I, Kahng B and Kim D 2002 *Phys. Rev. Lett.* **88** 108701

[6] Tangmunarunkit H, Govindan R, Jamin S, Shenker S and Willinger W 2002 *ACM Comput. Commun. Rev.* **32** 147

[7] Serrano M, Boguna M and Diaz-Guilera A 2005 *Phys. Rev. Lett.* **94** 038701

[8] Zhou S 2006 *Phys. Rev.* E **74** 016124

[9] Piraveenan M, Prokopenko M and Zomaya A Y 2009 *Eur. Phys. J.* B **70** 275

[10] Wang X and Loguinov D 2010 *IEEE Trans. Networking* **18** 257

[11] Oliveira R, Pei D, Willinger W, Zhang B and Zhang L 2010 *IEEE Trans. Networking* **18** 109

[12] Zhang G Q, Quoitin B and Zhou S 2011 *Comput. Commun.* **34** 649

[13] Trajkovic L J 2010 *IEEE Circ. Sys. Mag.* **10** 48

[14] Clegg R G, Cairano-Gilfedder C D and Zhou S 2010 *Comput. Commun.* **33** 259

[15] Dhamdhere A and Dovrolis C 2011 *IEEE Trans. Networking* **19** 1420

[16] Zhang L M, Deng X H, Yu J P and Wu X S 2011 *Chin. Phys.* B **20** 048902

[17] Zhou S Y, Wang K, Zhang Y F, Pei W J, Pu C L and Li W 2011 *Chin. Phys.* B **20** 080501

[18] Barabasi A L and Albert R 1999 *Science* **286** 509

[19] Albert R and Barabasi A L 2000 *Phys. Rev. Lett.* **85** 5234

[20] Bianconi C and Barabasi A L 2001 *Europhys. Lett.* **54** 436

[21] Fan Z P, Chen G R and Zhang Y N 2009 *Phys. Lett.* A **373** 1601

[22] Ayyoub M A and Gupta H 2010 *IEEE Trans. Mobile Computing* **9** 553

[23] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701

[24] Beygelzimer A, Grinstein G, Linsker R and Rish I 2006 *Phys. Lett.* A **357** 593

[25] Rodrigues F A, Costa L D F and Barbieri A L 2011 *Mol. BioSyst.* **7** 1263

[26] Alderson D L and Li L 2007 *Phys. Rev.* E **75** 046102

[27] Berche B, Ferber C V, Holovatch T and Holovatch Y 2009 *Eur. Phys. J.* B **71** 125

[28] Qu Z H, Wang P, Song C M and Qin Z G 2010 *Chin. Phys.* B **19** 110504