

Evolution of the Internet and its cores

To cite this article: Guo-Qing Zhang *et al* 2008 *New J. Phys.* **10** 123027

View the [article online](#) for updates and enhancements.

Related content

- [Symbiotic effect: A guideline for network modeling](#)
Guo-Qing Zhang, Guo-Qiang Zhang, Su-Qi Cheng *et al.*
- [Evaluating the AS-level Internet models: beyond topological characteristics](#)
Fan Zheng-Ping
- [A network growth model based on the evolutionary ultimatum game](#)
L L Deng, C Wang, W S Tang *et al.*

Recent citations

- [Khan Mohammad Habibullah *et al*](#)
- [Nonlinear growth: an origin of hub organization in complex networks](#)
Roman Bauer and Marcus Kaiser
- [Flow interaction based propagation model and bursty influence behavior analysis of Internet flows](#)
Xiao-Yu Wu *et al*

Evolution of the Internet and its cores

Guo-Qing Zhang^{1,2,5}, Guo-Qiang Zhang^{1,2}, Qing-Feng Yang^{1,2},
Su-Qi Cheng^{1,2} and Tao Zhou^{3,4,5}

¹ Institute of Computing Technology, Chinese Academy of Sciences,
Beijing 100190, People's Republic of China

² Graduate University of Chinese Academy of Sciences, Beijing 100190,
People's Republic of China

³ Department of Modern Physics, University of Science and Technology
of China, Hefei 230026, People's Republic of China

⁴ Department of Physics, University of Fribourg, Chemin du Muse 3,
CH-1700 Fribourg, Switzerland

E-mail: gqzhang@ict.ac.cn and zhutou@ustc.edu

New Journal of Physics **10** (2008) 123027 (11pp)

Received 15 September 2008

Published 18 December 2008

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/10/12/123027

Abstract. In this paper, we empirically study the evolution of large scale Internet topology at the autonomous system (AS) level. The network size grows in an exponential form, obeying the famous Moore's law. We theoretically predict that the size of the AS-level Internet will double every 5.32 years. We apply the k -core decomposition method on the real Internet, and find that the size of a k -core with larger k is nearly stable over time. In addition, the maximal coreness is very stable after 2003. In contrast to the predictions of most previous models, the maximal degree of the Internet is also relatively stable versus time. We use the edge-exchange operation to obtain the randomized networks with the same degree sequence. A systematical comparison is drawn, indicating that the real Internet is more loosely connected, and both the full Internet and the nucleus are more disassortative than their randomized versions.

⁵ Author to whom any correspondence should be addressed.

Contents

1. Introduction	2
2. Exponential growth of the Internet	3
3. The Internet is loosely connected compared with the randomized graph	4
4. Size stability of k-cores	6
5. Evolving properties of the nucleus	8
6. Conclusion and discussion	8
Acknowledgments	10
References	10

1. Introduction

The last few years have witnessed tremendous activity devoted to the understanding of complex networks. Particularly, functions and performance of the Internet, such as routing [1, 2], traffic statistics [3, 4], navigation [5, 6] and information retrieval [7, 8], attract more and more attention for their significance in modern society. Extensive evidence [9] has shown that the knowledge of topology is a crucial prerequisite for understanding and optimizing Internet performance. A seminal work by Faloutsos *et al* [10], revealing the heterogenous degree distribution of the Internet, has induced an avalanche of works aiming at uncovering the structural architecture of the Internet, including the degree–degree correlation [11], hierarchical organization [12], fractal properties [13], loop structure [14], rich-club phenomena [15], clique-degree distribution [16], and so on.

Although the immediate number of neighbors, namely the degree, is widely used as an indicator of the importance of each node, this is over simplified thus may lead to some misunderstanding since the networks of very different structures and functions could have exactly the same degree sequence [17]. Actually, the degree represents minimal local information. Based only on this information, one cannot distinguish whether a node belongs to the central part or is located in a peripheral position. As an effective tool to extract the central part of large scale networks, the k -core decomposition [18] has recently attracted much attention and has been extensively used to analyze various networks [19]–[22], especially those of heterogenous degree distribution. For the Internet at the autonomous system (AS) level, k -core decomposition is usually used as a basis for visualization [23, 24], and the invariant statistical properties of k -cores with different sizes (i.e. different values of k) indicate the self-similar nature of the Internet [25]. Very recently, Carmi *et al* [26] studied the structural properties of the most central part of the Internet at the AS level, namely the *nucleus*, which is defined as the smallest k -core (i.e. the k_{\max} -core with highest index k_{\max}).

In this paper, based on the empirical analysis of the temporal evolution of Internet maps, we show that (i) the size of a k -core with larger k is nearly stable over time (with some fluctuation), in contrast to the exponential growth of the full graph size; (ii) the maximal coreness is very stable after 2003, in contrast to the prediction from the configuration model; (iii) the maximal degree is relatively stable versus time, in contrast to the prediction of mainstream Internet models; (iv) the Internet is loosely connected compared with its randomized version; (v) both the Internet and its nucleus are more disassortative than the corresponding randomized networks.

Table 1. The basic topological properties of the Internet at AS level for about five years with sampling interval of six months. Here, N and E are the total number of nodes and edges, C denotes the average clustering coefficient, $\langle d \rangle$ is the average distance, r is the assortative coefficient [27] quantifying the degree–degree correlation, and k^* denotes the maximal degree among all nodes. Note that another symbol, k_{\max} , usually representing the maximal degree in the literature, is used to denote the maximal core index (also called coreness, with definition given below) in this paper. N_n denotes the size of the nucleus, that is to say, the number of nodes in the k_{\max} -core.

Time	N	E	C	$\langle d \rangle$	r	k^*	k_{\max}	N_n
2001–12	12666	25672	0.296	3.62	−0.199	2609	15	18
2002–6	13631	27749	0.292	3.65	−0.190	2692	15	24
2002–12	14625	29057	0.257	3.70	−0.193	2591	14	40
2003–6	15740	32263	0.264	3.71	−0.199	2507	17	40
2003–12	16691	35604	0.257	3.72	−0.199	2436	21	46
2004–6	17861	39637	0.266	3.72	−0.194	2437	24	47
2004–12	19085	42175	0.273	3.74	−0.200	2424	24	54
2005–6	20349	44016	0.275	3.75	−0.202	2462	23	37
2005–12	21588	45917	0.260	3.79	−0.196	2456	23	41
2006–6	22960	48545	0.242	3.82	−0.196	2460	23	78
2006–12	24403	52826	0.242	3.82	−0.196	2467	25	76

2. Exponential growth of the Internet

Up to now, the most prominent passive measurement for collecting AS-level Internet topology is the *Routeviews Project* [28], which set up several (about 40) border gateway protocol (BGP) collectors, each peering with dozens of BGP speakers residing in different ASs, to collect BGP tables and BGP updates. We collected the routing data from December 2001 to December 2006, with an interval of half a year. Hence, we have in total 11 AS-level Internet graphs. An AS graph is not a simple snapshot of the Internet, but a result of merging ten snapshots uniformly distributed within the same month. In this way, we hope an AS graph can give a more representative view of the Internet than a single snapshot can. We do not incorporate the trace-route data such as DIMES to complement the AS graph because the process of mapping IP level paths to AS paths still remains an open issue and could involve many false links, as indicated in [29]. The basic topological properties are presented in table 1.

As shown in figure 1, the size of the AS-level Internet grows very fast, actually, it obeys the famous *Moore's law* as $N(t) \sim 10^{0.0283t} \sim e^{0.0652t}$. We denote $\lambda = 0.0652$, the growing rate and Δt , the period during which the size doubles. Clearly, $\Delta t = \frac{1}{\lambda} \ln 2 \approx 10.64$. Since the time unit is half a year here, we predict that the size of AS-level Internet topology will double every 5.32 years (in the current framework, the maximal number of ASs is 2^{16} ; therefore, this prediction is just of theoretical meaning unless a new framework/protocol is established in the future that allows more ASs). The number of edges also grows in an exponential form. Indeed, it scales as $E \sim N^\gamma$, with $\gamma = 1.11 \pm 0.04$. In a short period, this relation can be well approximated as a linear function. In comparison, the number of links in the World-Wide-Web grows much faster as $E \sim N^{1.29}$ [30], exhibiting a remarkable effect of accelerating growth [31].

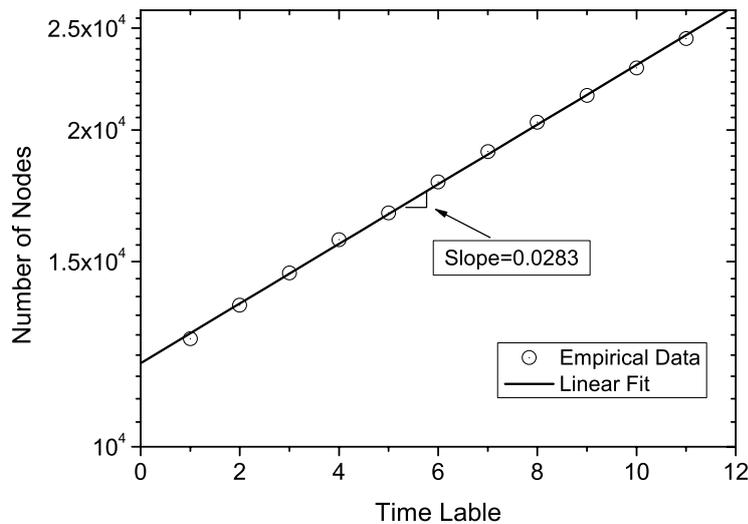


Figure 1. Number of nodes versus time. The growing tendency of the size of the AS-level Internet follows an exponential form with exponent 0.0283 ± 0.0001 . The y-axis is in a logarithmic scale. The time labels, from 1 to 11, correspond to December 2001 to December 2006 with six month intervals.

3. The Internet is loosely connected compared with the randomized graph

From table 1, one can get some structural information about the Internet, for example, it has very short average distance and large clustering coefficient, hence displays the so-called small-world phenomenon [32]. However, the topological measurements shown in table 1 are not independent of each other. For example, a network having very large clustering coefficient is, statistically, of longer average distance [33] since the abundant local connections are less helpful for reducing distances. Note that the temporal data reported here exhibit a negative correlation between clustering coefficient and average distance, but this is not in conflict with [33], since in [33], the networks are of the same size. Another example is that a network of very heterogenous degree distribution tends to have negative assortative coefficient [34]. Actually, based on extensive numerical analysis, Zhou and Mondragón [34] found that the assortative coefficient of a connected network having the same degree sequence as the Internet is always close to -0.2 (see figure 3 of [34]). That is to say, given such a degree sequence, one can try any optimization algorithms to enlarge or depress the assortative coefficient, however, the resulting value cannot be far from -0.2 , indicating that the assortative coefficient is not independent of degree distribution.

To filter out the structural bias induced by the heterogeneity of degree distribution, Maslov and Sneppen [35] proposed an edge-exchange operation, based on which a randomized network, having exactly the same degree sequence as the original network, can be obtained. Then, one can compare the topological properties between the original and the randomized networks, and this method can highlight the topological features besides degree distribution. As shown in figure 2, the procedure of the edge-exchange operation goes as follows: (i) randomly pick two existing edges $e_1 = (v_1, v_2)$ and $e_2 = (v_3, v_4)$, with all four vertices (v_1, v_2, v_3, v_4) being different; (ii) exchange these two edges to obtain $e'_1 = (v_1, v_4)$ and $e'_2 = (v_2, v_3)$. To ensure the

Table 2. Topological properties of randomized networks. Every data point is obtained by averaging over ten independent realizations, and in each realization, the number of exchanges is set as ten times of the number of edges.

Time	N	E	C	$\langle d \rangle$	r	k^*	k_{\max}	N_n
2001–12	12666	25672	0.178	3.42	-0.195	2609	19	47
2002–6	13631	27749	0.165	3.47	-0.185	2692	19	50
2002–12	14625	29057	0.154	3.52	-0.187	2591	18	52
2003–6	15740	32263	0.153	3.52	-0.192	2507	21	50
2003–12	16691	35604	0.155	3.51	-0.191	2436	24	52
2004–6	17861	39637	0.151	3.52	-0.185	2437	26	46
2004–12	19085	42175	0.149	3.53	-0.190	2424	26	55
2005–6	20349	44016	0.145	3.55	-0.192	2462	26	55
2005–12	21588	45917	0.132	3.61	-0.186	2456	25	54
2006–6	22960	48545	0.128	3.61	-0.185	2460	26	60
2006–12	24403	52826	0.123	3.63	-0.184	2467	27	70

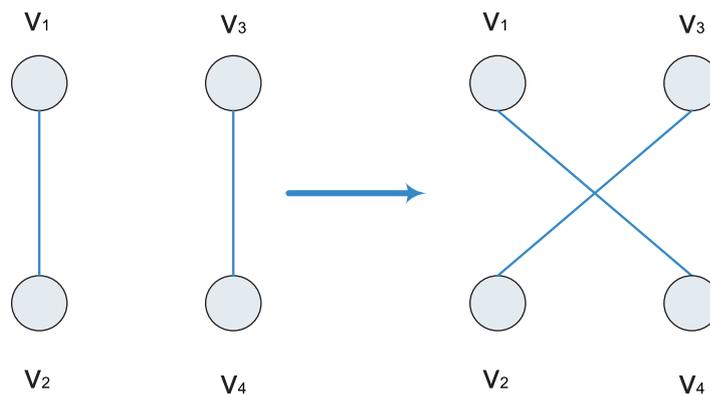


Figure 2. The edge-exchange operation used in the randomizing process [35]. Two edges are chosen randomly and the two vertices of edges are exchanged with each other. Multiple edges and self-edges are prohibited. This illustration is a glorified copy of figure 1 in [36].

operations are sufficient to get a fully randomized network, we set, in each realization, the number of exchanges equal to ten times of the number of edges.

Some basic topological measurements of the randomized networks are reported in table 2. The average distance in the real Internet is remarkably longer than its randomized version, indicating that the Internet is loosely connected. This is because in the Internet, the density of edges connecting nodes with very large degrees is lower than the randomized graph (see also the empirical results in [37]). This also makes the real Internet more disassortative (it has more edges connecting large-degree and small-degree nodes) and of larger clustering coefficient (the small-degree nodes determine the magnitude of the clustering coefficient in a network with very heterogenous degree distribution [12], therefore more edges between small-degree nodes make the clustering coefficient increased). Previously, the Internet is expected to be very compactly

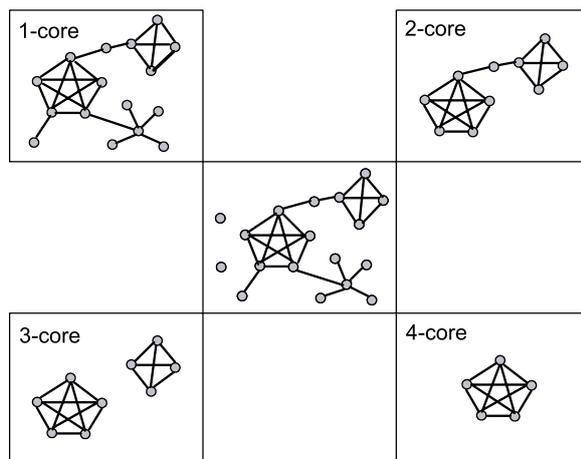


Figure 3. Illustration of k -core decomposition.

connected, showing the rich-club phenomenon [15]. The comparison between the Internet and its randomized version provides a completely different picture, that is, the Internet is, indeed, loosely connected. Colizza *et al* [37] claims that although large-degree nodes in the Internet are tightly connected, the connections between large-degree nodes are even sparser than those in the randomized networks. This conclusion is in accordance with our finding.

4. Size stability of k -cores

Uncovering the growing tendency of the central part of the Internet has great significance since those central nodes actually govern the global functions of the Internet. In this paper, we apply the k -core decomposition [18] method to extract the central part, where a k -core is obtained by recursively removing all the nodes of degree smaller than k , until the degree of all remaining nodes is larger than or equal to k . Here, k is called the core index or the coreness.

An illustration of k -core decomposition is shown in figure 3. Given a graph, the 0-core is exactly this graph itself, and the 1-core is the subgraph excluding all the isolated nodes. A node of degree larger than or equal to k may not appear in the k -core since some of its neighbors could be previously removed. Generally speaking [25, 26], a core of higher coreness is considered to be more central. We denote by $N(k)$ the number of nodes in the k -core, the highest coreness, k_{\max} , is defined as the maximal k that keeps $N(k)$ larger than zero. That is to say, k_{\max} is the highest coreness corresponding to a nonempty core.

Very recently, Carmi *et al* [26] studied the topological properties of the nucleus (i.e. the k_{\max} -core) of the Internet at the AS level. Based on a growing Internet model (scale-free configuration networks with parameters similar to the real Internet), their numerical results show that both the size of the nucleus, as well as k_{\max} , grows as a power of N (see figure 3(a) of [26]). We have applied the k -core decomposition method on the real data, as shown in table 1, there is no clear evidence of the exponential growth of k_{\max} versus time, which is not in accordance with the theoretical prediction by the configuration model. In particular, k_{\max} is very stable after 2003. Notice that, as shown in table 2, even for the randomized graphs, the k_{\max} remains stable after 2003. The size of the nucleus, N_n , exhibits large fluctuations versus time (see table 1), and no

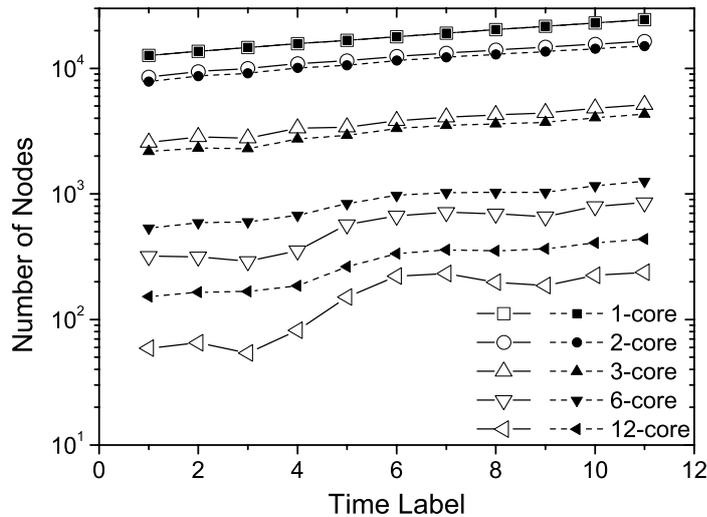


Figure 4. The sizes of k -cores versus time. The time labels, from 1 to 11, correspond to December 2001 to December 2006 with six month intervals. The large open symbols denote the empirical results of the real Internet, whereas the filled small symbols denote the numerical results of randomized networks. Those numerical results are obtained by ten realizations, and in each realization, the number of exchanges is set as ten times of the number of edges.

clear scaling can be observed (the fluctuations of N_n versus time are smaller in the randomized graphs, however, for both cases, there is no observable scaling behavior).

Since the values of k_{\max} are different for different samples, the direct comparison of the size of the nucleus versus time may not be relevant. Instead, we investigate the size of the k -core, $N(k)$, for a given k . As shown in figure 4, $N(1)$ and $N(2)$ display almost the same scaling as the size of the full Internet, N . $N(3)$ also shows clear increment, however, the data points have obvious fluctuations and cannot be well fitted by an exponential function. When k gets larger than 3, the clear increasing tendency is destroyed by large fluctuations. Comparing with the explosion of the full map of the Internet (see figure 1), the sizes of k -core ($k > 3$) are relatively stable, especially after the year 2003. Two typical examples, $N(6)$ and $N(12)$, are shown in figure 4. Those empirical results suggest that the explosion of the Internet is mainly a result of growth in the periphery, and the central part may undergo a far different evolutionary mechanism compared with the periphery nodes.

Figure 4 also reports the numerical results of the sizes of k -cores of the randomized graphs. When k is very small ($k \leq 3$), the temporal tendencies of the sizes of k -cores for randomized graphs are almost the same as those of the real Internet. In contrast, when k gets larger ($k > 3$), the k -cores of randomized graphs exhibit far different growing behavior from those of the real Internet. Firstly, for large k , $N(k)$ in the randomized graph is obviously larger than that in the real Internet. Secondly, $N(k)$ in the randomized graph grows monotonically without observable fluctuations. Even for the randomized networks, we could not find a clear scaling/fitting, however, a weak but solid conclusion can be drawn, that is, the size of the k -core (for large k) in the real Internet grows more slowly than that in the randomized graphs.

5. Evolving properties of the nucleus

Since the nucleus is the most central part in a network, to uncover its evolving properties is of significant importance. In figure 5, we report the empirical results for some fundamental topological characteristics. Firstly, we would like to see if the content of the nucleus changes much during its evolution. We define $V_n(t)$ as the set of nodes in the nucleus at time t , and $\rho(t)$ as the fraction of nodes in $V_n(t)$ that stay in $V_n(t+1)$, namely $\rho(t) = \frac{|V_n(t) \cap V_n(t+1)|}{|V_n(t)|}$, where $|A|$ denotes the number of elements in set A . Clearly, small ρ corresponds to inconstant content of the nucleus. As shown in figure 5(a), in most cases, $\rho(t)$ is larger than 80%. That is to say, the content of the nucleus does not change much.

Figure 5(b) reports the density of edges in the nucleus, which is defined as $D_E = \frac{2E_n}{N_n(N_n-1)}$, where E_n is the number of edges in the nucleus. In most cases, D_E is larger than 0.5, indicating that the nucleus is very tightly connected. In the nucleus, the degree of every node is no less than k_{\max} , and thus $2E_n \geq k_{\max} \times N_n$. Therefore, there exists a lower bound of D_E , namely $D_E \geq \frac{k_{\max}}{N_n-1}$. This lower bound is also shown in figure 5(b). It is observed that the change of edge density in the nucleus exhibits the same tendency as the theoretical lower bound. However, the value of D_E is much larger than the lower bound, again indicating that the nucleus is very tightly connected. As observed in figures 5(e) and (f), the clustering coefficients and distances (including both diameters and average distances) of the nucleus are very close to those of the randomized networks (distances of nucleus and the randomized version are exactly the same), in contrast, the nuclei are more disassortative than their randomized version. In figure 5(c), we show the nodes' average degree in the original Internet (i.e. the full Internet). Compared with the typical value of average degree (about 4, see table 1) and the lower bound degree for a node to be included in the nucleus, i.e. k_{\max} , one can say that most of the nodes in the nucleus have high degrees (for example, more than 70% of nodes in the nucleus in December 2006 have degree larger than 100, while the average degree of the full graph is about 4 and k_{\max} is 25). It is interesting that D_E , $\langle k \rangle$, r and C are strongly positively correlated, whereas they are negatively correlated with average distance. The correlation between edge density and average distance is easily understood. Actually, in the high-density case, the shortest path of length larger than 2 can be ignored (the diameter reported in figure 5(f) has already demonstrated that no shortest path has length larger than 2), and only if two nodes are not directly connected is their distance 2. Therefore, $\langle d \rangle = 1 \times D_E + 2 \times (1 - D_E) = 2 - D_E$, which exhibits a completely negatively correlation with edge density. We have checked that this analytical result is exactly the same as the numerical result. However, other correlations cannot be simply explained, for example, additional links could simultaneously increase the number of triangles and the degrees of relevant nodes, thus it is hard to say the network with higher density must have larger clustering coefficient. In addition, the additional links may connect nodes of large degrees thus increase r , however, they may connect large-degree nodes and low-degree nodes, thus depress r . Thus far, It is not clear for us whether the correlations found in this paper represent some specific topological characters of the nucleus, or whether they are just a trivial phenomenon. We here report this empirical phenomenon, and leave the possible explanation as an open question.

6. Conclusion and discussion

In this paper, we empirically study the evolution of large scale Internet topology at the AS level. The network size grows in an exponential form, and will double approximately every 5.32 years.

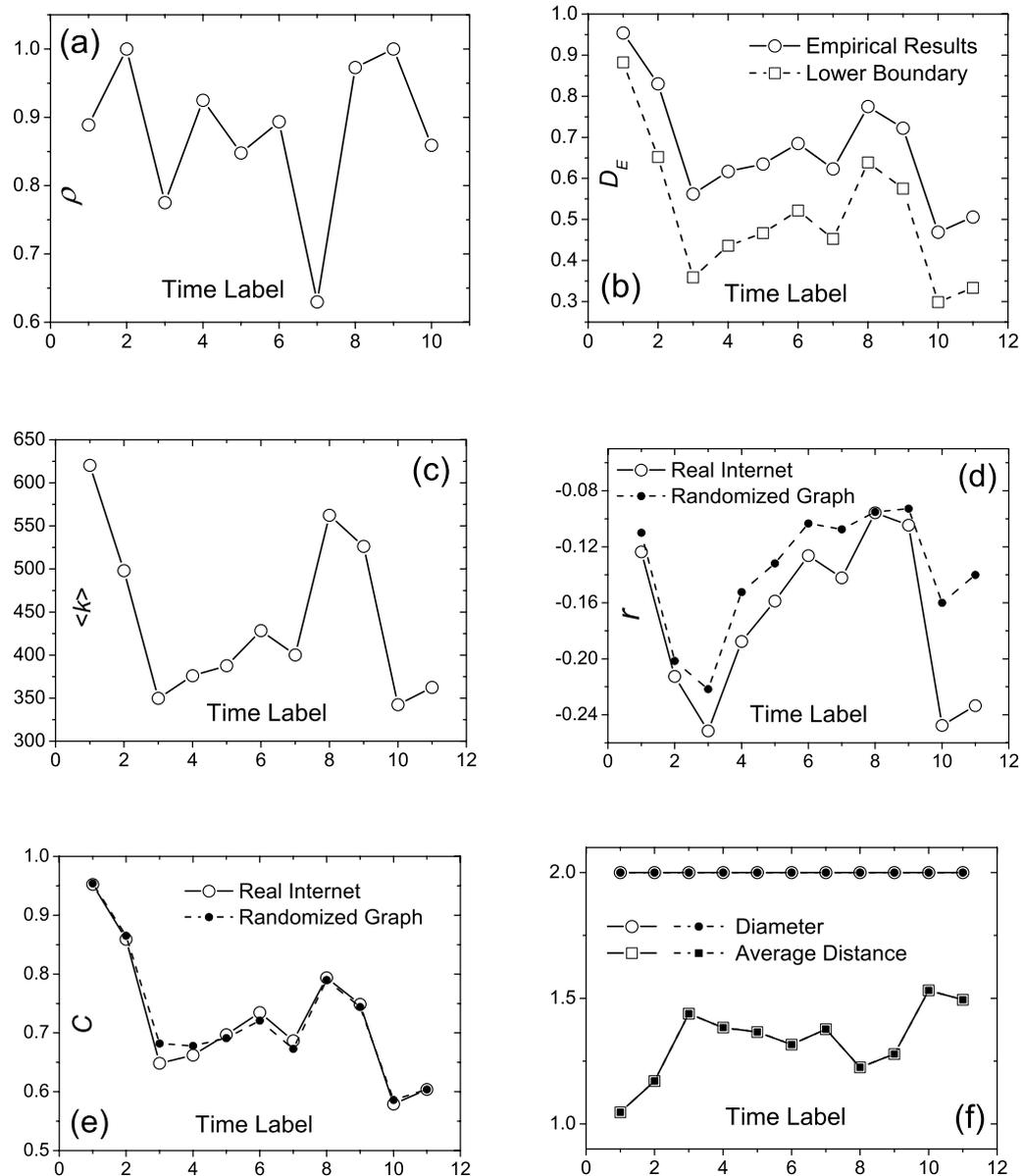


Figure 5. Statistical properties of the nucleus versus time. (a) The fraction of nodes still in the nucleus at the next sampling time, $\rho(t)$, versus time t . The time labels, from 1 to 10, correspond to December 2001 to June 2006 with six month intervals. (b) The edge density of empirical data (\circ) and the corresponding theoretical lower bound (\square). (c) The nodes' average degree in the full Internet. (d) The assortative coefficient. (e) The clustering coefficient. (f) The average distance (denoted by \square) and the maximal distance (namely diameter, denoted by \circ). In the panels (b)–(f), the time labels, from 1 to 11, correspond to December 2001 to December 2006. In the plots (b), (d), (e) and (f), the large open symbols denote the empirical results of the nucleus, whereas the solid and small symbols denote the numerical results of the randomized networks. Those numerical results are obtained by 10 realizations, and in each realization, the number of exchanges is set as 10 times of the number of edges.

Although in the current framework, the maximal number of ASs is 2^{16} , our finding is of theoretical interest. Correspondingly, the number of edges shows a weakly accelerating growth as $E \sim N^{1.11}$. Different from the theoretical prediction by a simple configuration model [26], the sizes of k -cores with larger k are relatively stable compared with the growth of the full Internet map, and the highest coreness, k_{\max} , is also stable, especially for the data after 2003. Those results suggest that the central part and the periphery of the Internet should be governed by different evolutionary mechanisms. Actually, the majority of the new nodes and edges that contribute to the explosion of the Internet appear in the periphery.

Up to our knowledge, the most accurate Internet model (judged by a number of topological parameters) is the so-called positive-feedback preference (PFP) model [38]–[41]. In this model, a node's ability to acquire new links increases as a feedback loop of the node's degree, thus the maximal degree, k^* , increases very fast (faster than the Barabási–Albert model [42]) as the network size increases. However, as shown in table 1, the maximal degree of the Internet is also relatively stable versus time, indicating the existence of some hidden evolving mechanisms instead of or in addition to the PFP mechanism. Indeed, most of the previous models embedded in the preferential attachment mechanism could not reproduce the stability of the maximal degree. The aging effect can lead to an evolving network with relatively stable maximal degree [43], however, there is no clear evidence indicating an aging mechanism in the real Internet. The limitation of traffic capacity in an individual level may cause a boundary of the individual connectivity. Another candidate that may contribute to the statistical properties reported here is the mutual interaction among existing nodes [44]: according to the transportation demand of information packets, new edges between existing nodes may be created while some existing edges may disappear or be rewired. In addition, we systematically compared the structures of the real Internet and its randomized version, and found that the real Internet is more loosely connected, which is in accordance with the empirical results reported in [37]. We believe this work can provide insights into Internet topology, as well as some evidence of the mechanism that governs the evolution of the Internet. In particular, it gives some important criteria for modeling the Internet.

Acknowledgments

We acknowledge Dr Jian-Guo Liu for his assistance in preparation of the manuscript. This work is partly supported by the National Natural Science Foundation of China under grant no 60673168 and the Hi-Tech Research and Development Program of China under grant no 2006AA01Z207 and 2008AA01Z203. TZ acknowledges the National Natural Science Foundation of China under grant no 10635040, as well as the 973 Program 2006CB705500.

References

- [1] Yan G, Zhou T, Hu B, Fu Z-Q and Wang B-H 2006 *Phys. Rev. E* **73** 046108
- [2] Zhang G-Q, Wang D and Li G-J 2007 *Phys. Rev. E* **76** 017101
- [3] Barthélemy M, Gondran B and Guichard E 2002 *Phys. Rev. E* **66** 056110
- [4] Tadić B, Thurner S and Rodgers G J 2004 *Phys. Rev. E* **69** 036102
- [5] Kleinberg J M 2000 *Nature* **406** 845
- [6] Zhou T 2008 *Physica A* **387** 3025
- [7] Zhou T, Ren J, Medo M and Zhang Y-C 2007 *Phys. Rev. E* **76** 046115

- [8] Zhou T, Jiang L-L, Su R-Q and Zhang Y-C 2008 *Europhys. Lett.* **81** 58004
- [9] Pastor-Satorras R and Vespignani A 2004 *Evolution and Structure of the Internet* (Cambridge: Cambridge University Press)
- [10] Faloutsos M, Faloutsos P and Faloutsos C 1999 *Comput. Commun. Rev.* **29** 251
- [11] Pastor-Satorras R, Vázquez A and Vespignani A 2001 *Phys. Rev. Lett.* **87** 258701
- [12] Ravasz E and Barabási A-L 2003 *Phys. Rev. E* **67** 026112
- [13] Caldarelli G, Marchetti R and Pietronero L 2000 *Europhys. Lett.* **52** 386
- [14] Bianconi G, Caldarelli G and Capocci A 2005 *Phys. Rev. E* **71** 066116
- [15] Zhou S and Mondragón R J 2004 *IEEE Commun. Lett.* **8** 180
- [16] Xiao W-K, Ren J, Feng Q, Song Z-W, Zhu M-X, Yang H-F, Jin H-Y, Wang B-H and Zhou T 2007 *Phys. Rev. E* **76** 037102
- [17] Doyle J C, Alderson D L, Lun L, Low S, Roughan M, Shalunov S, Tanaka R and Willinger W 2005 *Proc. Natl Acad. Sci. USA* **102** 14497
- [18] Pittel B, Spencer J and Wormald N 1996 *J. Combin. Theory* **67** 111
- [19] Bader G D and Hogue C W V 2003 *BMC Bioinform.* **4** 2
- [20] Wuchty S and Almaas E 2005 *Proteomics* **5** 444
- [21] Dorogovtsev S N, Goltsev A V and Mendes J F F 2006 *Phys. Rev. Lett.* **96** 040601
- [22] Schwartz J M, Liu A J and Chayes L Q 2006 *Europhys. Lett.* **73** 560
- [23] Baur M, Brandes U, Gaertler M and Wagner D 2004 *Proc. 12th Int. Symp. on Graph Drawing* (Berlin: Springer-Verlag) pp 43–8
- [24] Alvarez-Hamelin J I, Dall'Asta L, Barrat A and Vespignani A 2006 *Adv. Neural Inform. Process. Syst.* **18** 41
- [25] Alvarez-Hamelin J I, Dall'Asta L, Barrat A and Vespignani A 2008 *Netw. Heterogen. Media* **3** 371
- [26] Carmi S, Havlin S, Kirkpatrick S, Shavitt Y and Shir E 2007 *Proc. Natl Acad. Sci. USA* **104** 11150
- [27] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
- [28] <http://www.routeviews.org>
- [29] Oliveira R, Zhang B and Zhang L 2007 *Proc. ACM SIGCOMM 2007* (New York: ACM Press)
- [30] Broder A, Kumar R, Moghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A and Wiener J 2000 *Comput. Netw.* **33** 309
- [31] Dorogovtsev S N and Mendes J F F 2001 *Phys. Rev. E* **63** 025101
- [32] Watts D J and Strogatz S H 1998 *Nature* **393** 440
- [33] Zhao M, Zhou T, Wang B-H, Yan G, Yang H-J and Bai W-J 2006 *Physica A* **371** 773
- [34] Zhou S and Mondragón R J 2007 *New J. Phys.* **9** 173
- [35] Maslov S and Sneppen K 2002 *Science* **296** 910
- [36] Wang B, Zhou T, Xiu Z-L and Kim B-J 2007 *Eur. Phys. J. B* **60** 89
- [37] Colizza V, Flammini A, Serrano M A and Vespignani A 2006 *Nat. Phys.* **2** 110
- [38] Zhou S and Mondragón R J 2004 *Phys. Rev. E* **70** 066108
- [39] Zhou S 2006 *Phys. Rev. E* **74** 016124
- [40] Zhou S, Zhang G-Q and Zhang G-Q 2007 *IET Commun.* **1** 209
- [41] Zhang G-Q and Zhang G-Q *ICC 2007 Proc.* (Piscataway, NJ: IEEE Press)
- [42] Barabási A-L and Albert R 1999 *Science* **286** 509
- [43] Dorogovtsev S N and Mendes J F F 2001 *Phys. Rev. E* **62** 1842
- [44] Wang W-X, Hu B, Zhou T, Wang B-H and Xie Y-B 2005 *Phys. Rev. E* **72** 046140