Computer Communications xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies

Guoqiang Zhang [a,*], Bruno Quoitin [b], Shi Zhou [c]

[a] *Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China*
[b] *Networking Lab, Université de Mons, Place du Parc, 20, 7000 Mons, Belgium*
[c] *Department of Computer Science, University College London, United Kingdom*

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Internet topology
Autonomous systems (AS)
Network evolution
Phase change
IPv4 and IPv6

ABSTRACT

In this paper, we investigate the evolution of the IPv4 and IPv6 Internet topologies at the autonomous system (AS) level over a long period of time. We provide abundant empirical evidence that there is a phase transition in the growth trend of the two networks. For the IPv4 network, the phase change occurred in 2001. Before then the network's size grew exponentially, and thereafter it followed a linear growth. Changes are also observed around the same time for the maximum node degree, the average node degree and the average shortest path length. For the IPv6 network, the phase change occurred in late 2006. It is notable that the observed phase transitions in the two networks are different, for example the size of IPv6 network initially grew linearly and then shifted to an exponential growth. Our results show that following decades of rapid expansion up to the beginning of this century, the IPv4 network has now evolved into a mature, steady stage characterised by a relatively slow growth with a stable network structure; whereas the IPv6 network, after a slow startup process, has just taken off to a full speed growth. We also provide insight into the possible impact of IPv6-over-IPv4 tunnelling deployment scheme on the evolution of the IPv6 network. The Internet topology generators so far are based on an inexplicit assumption that the evolution of Internet follows non-changing dynamic mechanisms. This assumption, however, is invalidated by our results. Our work reveals insights into the Internet evolution and provides inputs to future AS-Level Internet models.

1. Introduction

The Internet has experienced rapid growth in the past 30 years, evolving from a simple laboratory test-bed network to a gigantic ecosystem. It is often considered as the most complex technological network ever made by human beings. From the highest level, this ecosystem can be represented by a graph, where nodes represent the autonomous systems (ASes), and two nodes are connected if and only if the two ASes are engaged in a business relationship to exchange data traffic.

Since late 1990s', various research activities are devoted to the mapping, characterisation and modelling of the Internet [1–11,44,45]. These efforts have indeed uncovered intriguing features of the Internet, e.g., power-law degree distribution [1], rich-club phenomenon [12], disassortative mixing [13], self-similarity [40], etc. These discoveries are further followed by proposals of different network models that try to reproduce these distinctive topological properties [14–18,42]. Readers can refer to [46] for a survey of network modelling and generation.

However, despite the significant amount of efforts, existing studies still face several problems and challenges:

- Firstly, although tremendous Internet measurement projects are set up, we still cannot have a comprehensive and accurate view of the real AS topology [19–22]. This is because the AS topology inference methods, either BGP-based or traceroute-based, suffer a common problem of systematic loss of a non-trivial fraction of links, mostly peer-to-peer links between periphery nodes.
- Secondly, most studies are carried out on particular *snapshots* of the Internet topology or over short-term historic data (less than 5 years), e.g., topological properties are uncovered for particular snapshots, and network models are validated by particular observed snapshots. Relatively few efforts have been put to the evolutionary study of the Internet topology over a long time period.
- Thirdly, of the limited number of studies on the evolution of the Internet, researchers often do not determine the real causes for observed topology changes. Some of the changes may not due to real evolution events but originate from the variation of monitors [2,23–25]. This makes their claims questionable.
- Finally, the Internet now is experiencing a gradual transition from the IPv4 network to the IPv6 network due to a number of reasons including the shortage of IP addresses. A natural question is whether these two networks show similar or different evolutionary trends. Yet, to the best of our knowledge, very

---

* Corresponding author. Tel.: +86 1062600721.
E-mail address: guoqiang@ict.ac.cn (G. Zhang).

0140-3664/$ - see front matter Crown Copyright © 2010 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2010.06.004

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004

few work has been done to study the evolution of the IPv6 network, let alone a side-by-side study of the two networks. Without this study, problems such as how the IPv6-over-IPv4 tunnelling impacts on the evolution of the IPv6 network could not be properly understood.

Motivated by these, in this paper we undertook an in-depth side-by-side study of the evolution of the IPv4 and IPv6 AS-Level Internet topologies over a long period of time. We aim to answer questions such as: whether the Internet has a uniform evolution process, or experiences different evolution stages? whether its featured structural properties keep unchanged, or evolve over time? and whether the existing network models are capable of modelling the real evolution process of the Internet? More specifically, our original contributions are:

1. We are the first to carry out a long-term side-by-side evolutionary study of the IPv4 and IPv6 network topologies at the AS level.
2. Based on historic routing data, we show ample empirical evidence that both the IPv4 and IPv6 networks have experienced a phase change in their evolution, but with different transition patterns. The IPv4 network has evolved into a stable structure, whereas the IPv6 network has just entered a stage of rapid growth. Notably, it is the first time in the literature to discover phase change in the evolution of the IPv6 network.
3. We have discussed the impact of IPv6-over-IPv4 tunnelling deployment scheme on the evolution of the IPv6 network.
4. We point out the fundamental impact of the phase changes of the Internet evolution on designing and evaluating future Internet models.

The following of the paper is organised as such. Related work is discussed in Section 2. Section 3 presents the data sets and approaches we use for this study. Section 4 gives the side-by-side evolution study of the IPv4 and IPv6 AS-Level topologies. We discuss our findings in Section 5. Finally, we conclude the paper in Section 6.

2. Related work

The last decade has witnessed a surge of research activities related to network topology measurement, characterisation and modelling. Various projects are set up to map the Internet topology. The BGP table dumps archived by Routeviews [10] and RIPE [11] offer good feeds for the study of AS-Level Internet topology. The outcome of the active measurement projects, such as CAIDA [8] and DIMES [9], on the other hand, provides input to studies for both the AS-Level and router-level Internet topologies.

These data sources provide researchers with an unprecedented opportunity to uncover the unique structural properties as well as evolutionary mechanisms of this complex man-made system. Various topological features are discovered for specific topology snapshots, e.g., power-law degree distribution [1], assortative mixing [13], rich-club phenomenon [12], extremely large maximum degree [17,25], high clustering coefficient [26], and self-similarity between regional AS subgraph and the global AS graph [40]. These analysis were followed by a number of graph-theory based generative models to reproduce the observed characteristics and try to explain the evolution of networks, e.g., BA [14], AB [15], GLP [16], PFP [17,18].

Recently, there is a growing trend to study the Internet from an evolutionary perspective. Based on the early day's data from Routeviews, it was shown in [2] and [23] that the AS-Level Internet topology was densifying and its effective diameter was shrinking. In [27], the authors grouped the BGP data from Routeviews and RIPE into three sets to evaluate the effects of different monitors, i.e., data from a single monitor, data from a fixed number of monitors that are present throughout the entire measurement period, and data from all monitors. It was shown that after a short exponential revealing period, the network follows a constant birth rate. In [24], the authors carried out a 10-year study of the evolution of the Internet (the longest time period among this kind of studies before this paper) and it was shown that the number of ASes as well as the number of CP (customer-provider) links follow similar growth trends, that is, both grow exponentially from Nov, 1997 to May, 2001, and then enter into a linear growth mode. In our recent work [25], we reported that the maximum degree remains nearly invariable in recent years, and the so-called *k*-core property is stable over time.

Only recently, there has been some effort towards characterising and modelling the IPv6 network. CAIDA's Ark project began to perform continuous large-scale active measurement of the IPv6 network since Dec, 2008 [29]. In [30], it was shown that although the IPv6 AS topology obeys power-law, its degree exponent is much smaller than the IPv4 counterpart, and a novel model was proposed to reproduce this smaller degree exponent phenomenon. Eddy [47] took a three-year-long evolutionary study of the IPv6 AS-Level topology. However, since the author only studied the data from May, 2003 to Sep, 2006, he did not observe the phase transition that took place in 2006 as we will report in the following.

3. The data set

In this study, we used the data set offered by Routeviews and RIPE since they are the only public sources that archive historic BGP data. We do not use the AS topological data derived from traceroute measurements due to issues in converting router paths to AS paths [27,31,32]. We used an approach similar to [27] to group the data into different sets to evaluate the effects of different monitors. For IPv4, we built three different data sets:

- **OIX**: data from the single Routeviews collector *route-views.route-views.org*, which is extensively used in early day's AS topology analysis. We collected the data from the starting date of the collector, i.e., Nov, 1997.
- **Set52**: data from a set of 52 monitors in both Routeviews and RIPE. These 52 monitors reside in 36 ASes that persist all the time since Jul, 2004. According to the AS taxonomy provided by CAIDA [48,37], the 36 ASes contain 11 tier-1 ASes, 19 tier-2 ASes, 2 NICs, and 4 abstained ASes.[1]
- **ALL**: data from all the collectors of Routeviews (except route-views6) and RIPE that started prior to Jul, 2004.

For IPv6, we built two data sets:

- **Set4**: data from 4 monitors (residing in the following four ASes: AS2497, AS2914, AS7660, AS30071) that occurred most frequently since May, 2003, with each occurring more than 60 times out of the 77 months.
- **ALL**: data from the route-views6 collectors since May, 2003.

We collect the data on a monthly basis. Each month, we collected one snapshot from each collector in the last day of the month with collection time as close as possible for different monitors, and then synthesised the AS paths from different monitors to construct the corresponding data set. AS paths that contain AS set, private ASNs, or loops were filtered out from the graph construc-

---

[1] Abstained means the algorithm [37] fails to make a predication of the AS class.

tion. Although collecting only one snapshot in a month can miss some hidden links that could be revealed at a later time, merging all the snapshots over a relatively long time period, however, can potentially introduce the problem of stale links [21,22,27]. We thus focused on an *instant operating view* of the AS-Level Internet topology by merging the snapshots from various monitors.

In our following study of network evolution, for each topological property, we will first leverage different data sets to make a rough judgement on whether the property in question is sensitive to the number and set of monitors, and then choose the appropriate data set for further reasoning. Taking the IPv4 network as an example, for those properties that are insensitive to the number of monitors or can be gracefully characterised by existing monitors, we will focus on the OIX data set to supply a comprehensive view of the evolutionary trends of these properties. While for those properties that are sensitive to monitors, we will primarily rely on the Set52 data set to make our conclusions and use other data set conservatively, hoping to minimise the effect of biased sampling due to monitor variation.

4. Evolution of the IPv4 and IPv6 AS-Level Internet topologies

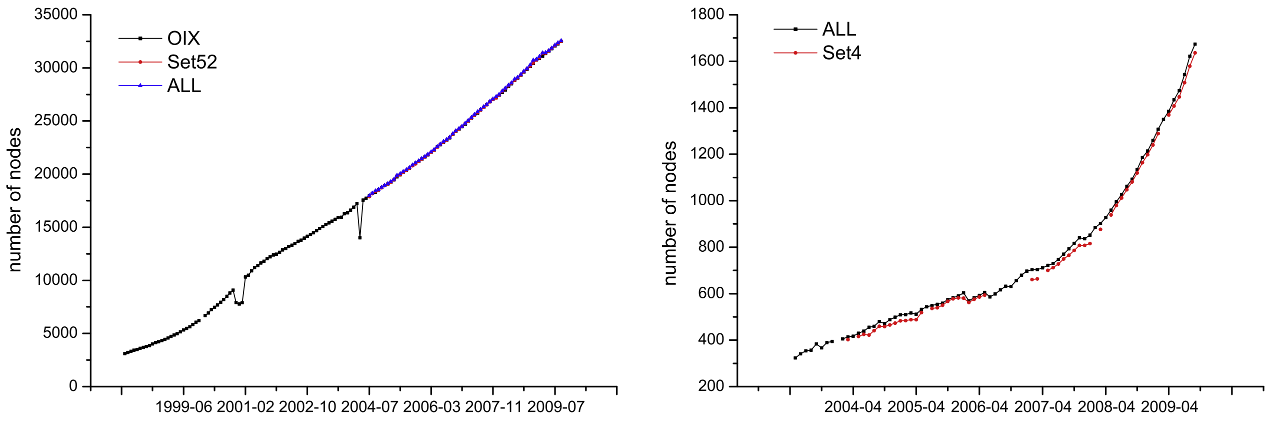
Here, we perform a side-by-side evolutionary study of several important graph properties of the IPv4 and IPv6 AS-Level topologies. These properties include network size (number of nodes and edges), degree properties (maximum degree, average degree, and degree distribution), average shortest path length, clustering coefficient and assortative coefficient.

4.1. Network size

The first and foremost question of network evolution is how the network size evolves over time. Network size consists of two aspects: the number of nodes and the number of edges. However, the limited number of BGP monitors has significant impact on the number of edges that can be discovered. It has long been recognized that the AS topology inferred from BGP data will systematically lose a large fraction of peer-to-peer links [3,20–22,19]. Nevertheless, the monitor issue almost has no effect on the number of nodes that can be detected [3,22,19,33]. These perceptions are obviously confirmed by Figs. 1 and 2, from which we observe that whatever the data set is, the number of nodes observed are similar, albeit the number of detected edges can show significant differences.

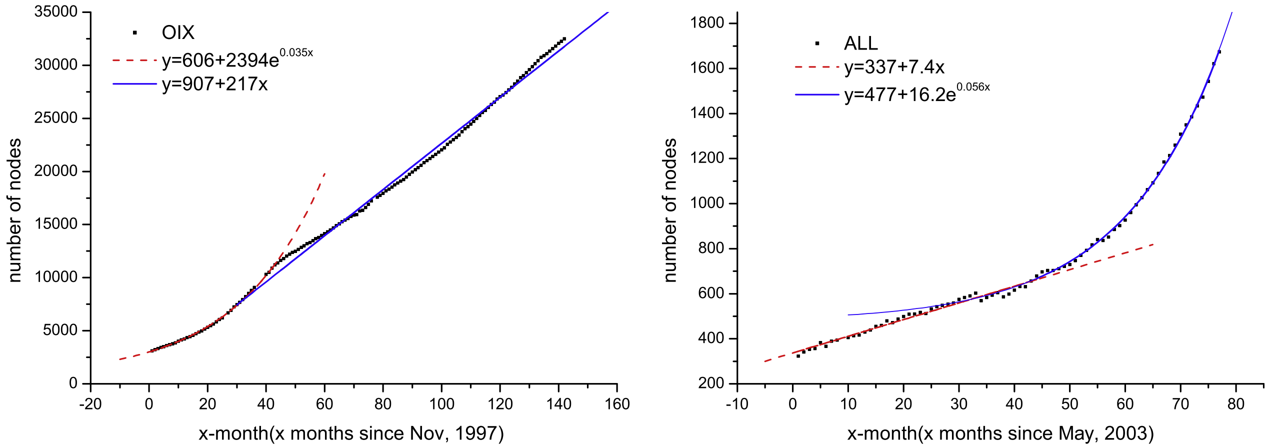
Since the number of monitors has little effect on the number of ASes that can be detected, we can rely on the OIX data set to make a long time study of the evolution of the number of nodes over the past 12 years. It is easy to find that from 1997 to 2001, the number of nodes obeyed an exponential growth rate, but after that, it can be better described by a linear growth process. This effect has also been reported in [24].

The IPv6 network is different from the IPv4 network. In the early days from 2003 to 2006, the number of nodes grew linearly,



(a) IPv4

(b) IPv6

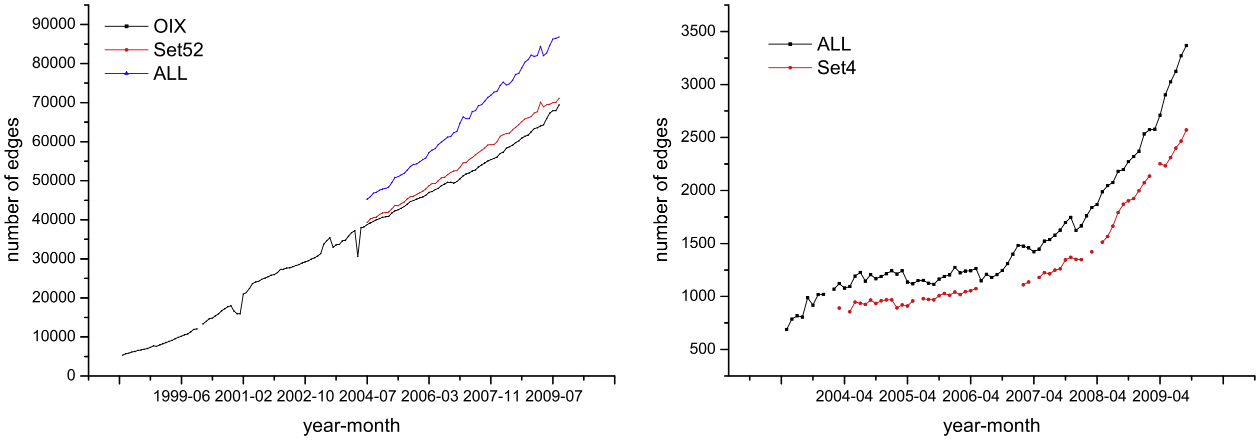


(c) Curve fittings for IPv4 (OIX)

(d) Curve fittings for IPv6 (ALL)

Fig. 1. Evolution of the network size (number of nodes) of IPv4 and IPv6 networks.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004



(a) IPv4 (b) IPv6

Fig. 2. Evolution of the number of edges of the IPv4 and IPv6 networks.

which was also reported in [47]. However, after 2006, the number of nodes grows exponentially. Fig. 1(c) and (d) presents the fitting functions of the two curves of IPv4's OIX data set and IPv6's ALL data set (we excluded some apparently exceptional points in the OIX data set during the fitting, and renumbered the months by sequential numeric numbers). The result is that, in the IPv4 network, the leading portion of the curve grows exponentially with $y \sim e^{0.035x}$, and the rest grows linearly with $y \sim 217x$. While in the IPv6 network, the leading portion grows linearly with $y \sim 7.4x$, and the rest grows exponentially with $y \sim e^{0.056x}$. It is interesting to observe that the exponential growth rate of IPv6 after 2006 is even faster than the exponential growth rate of IPv4 before 2001.

The number of edges has similar growth trends, however, since the number and set of monitors vary over time, any conclusions made on the edges should be taken cautiously.

The difference between the growth patterns in the number of ASes in the IPv4 and IPv6 networks is an indication of the different development stages of these two networks. The IPv4 network, after a rapid exponential growth, enters into a more stable stage, whereas the IPv6 network is still in the exponential growth stage.

4.2. Maximum degree

Fig. 3 reports the evolution of maximum degrees in the two networks. The maximum degree is a particularly important topological property in the AS topology because it is often far larger than what the typical preferential attachment models would predict and hence plays a crucial role in ensuring the network connectivity. It is observed that despite the different sets of monitors used to construct the graphs, the growth trends of maximum degree are similar. This means the maximum degree is largely unaffected by the limitation of the current monitoring system. In fact, the nodes with the highest degrees are always the tier-1 transit ASes. As is shown in [22], the current public view of BGP monitors are sufficient to detect all the neighbouring links of these tier-1 ASes in the IPv4 network, so the maximum degree is largely unaffected by deploying more monitors. In comparison, there is still a slight gap of the maximum degree between Set4 and ALL data set in the IPv6 network, which implies that 4 monitors are insufficient to capture all the neighbouring links of tier-1 ASes in the IPv6 network.

In the IPv4 network, the maximum degree grew rapidly from 1997 to 2001, after which it remained relatively stable. Our previous analysis on the data from Dec, 2001 to Dec, 2006 also confirmed that the maximum degree of IPv4 AS-Level topology remains quite stable [25]. In this paper, we give a more comprehensive picture of its evolution over a much wider temporal spectrum. In the IPv6 network, the maximum degree grew slowly from 2003 to 2007 (similar result was also reported in [47]), while after that it entered into a rapid growth stage. The maximum degree growth pattern is another indication that the IPv6 network is currently in the rapid expansion stage.

4.3. Average degree

The density of connectivity in a network can be indicated by the average degree of nodes, which can be given as $2L/N$ where L is the number of edges and N is the number of nodes. Fig. 4(a) shows the evolution of average degree of the IPv4 and IPv6 networks. For the IPv4-OIX data set, the average degree was increasing until 2001 and then it remained relatively stable. For the IPv4-Set52 data set, the average degree is also very stable in recent years. These observations do not support previous claims that the AS-Level Internet topology was a so-called accelerating network [34,35], or the Internet followed the so-called densification law [2,23], i.e. the number of edges grows faster than the number of nodes, or equivalently, the average node degree increases. This claim may be correct before the phase change, but it stops accelerating after the phase change. For the IPv4-ALL data set, the average degree was much larger and was still increasing. The larger average degree is due to the larger number of monitors, but it is not clear whether the still increasing average degree is also due to the increasing number of monitors. Nevertheless, the average degree of the IPv6 network, however, exhibits an remarkably different evolutionary trend, where the average degree is in fact decreasing rapidly in recent years. This is the trend for both the IPv6-Set4 and IPv6-ALL data sets.

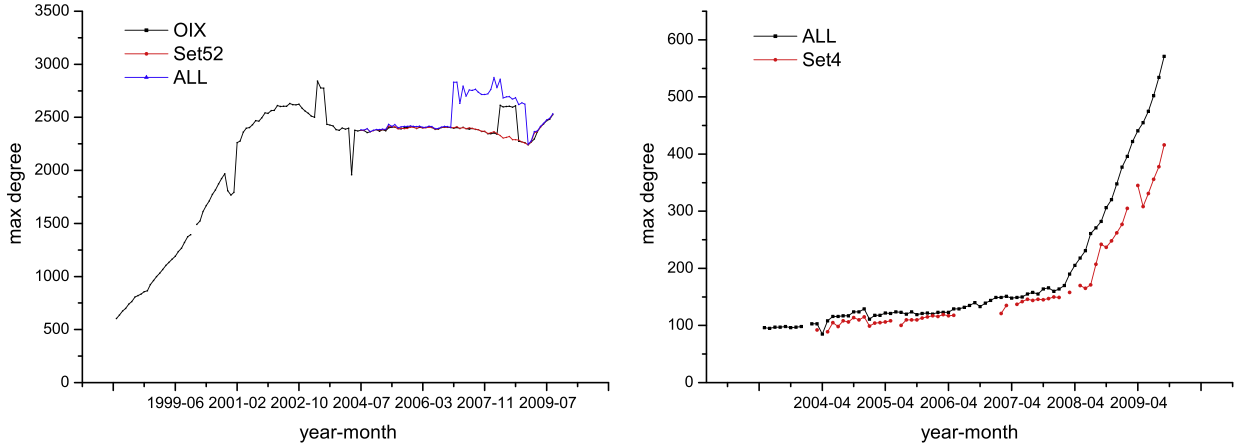
4.4. Shortest path length

A key topological property of a network is the average shortest path length between any pair of nodes, which indicates a network's routing efficiency if all traffic follows the shortest path available.[2]

---

[2] In reality, inter-AS paths are also constrained by the routing policies [36] and real AS paths may be inflated compared with the shortest paths [6], but minimising the number of hops is still a major criterion in path selection and the shortest path also reflects the best achievable routing efficiency.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004



(a) IPv4 (b) IPv6

Fig. 3. Evolution of the maximum degree of the IPv4 and IPv6 networks.

It is reported in [2] that the average shortest path length of the IPv4 AS graph was shrinking, whereas it was reported in [25] that this measure was increasing. This contradiction arises from the fact that these two works investigated different stages of the Internet evolution. In the former, the data was collected from 1999 to 2001, just before the phase change; while in the latter, the data was from 2001 to 2006, just after the phase change.

We show in Fig. 5(a) the average shortest path length of the IPv4 network evolution over a much longer period encompassing the two previous works. The OIX data set shows that following a few years of decreasing, the average shortest path started to grow in 2001. The decreasing of the average shortest path in the early years is arguable because it is unclear whether it was merely due to the lack of monitors. Nevertheless, all three data sets show that the average shortest path length of the IPv4 network is increasing in recent years.

Taking this phenomenon with the evolutionary trend of the maximum degree, we can conjecture that before 2001, the IPv4 network was at an evolution stage when there was a boom of newly born ASes. The new ASes tended to connect to the most-connected ASes (or tier-1 ASes), and the most-connected ASes rapidly enriched their mutual peering relationships (i.e. rich-club phenomenon [12]). As a result, the maximum degree increased rapidly, and the average shortest path length *might* shrink. After 2001, the core of the IPv4 became relatively stable in terms of the number of tier-1 ASes and the edges among them [25], and newly born ASes primarily connected to tier-2 or tier-3 regional service providers. Therefore the maximum degree stopped increasing, and the average shortest path length started to increase.

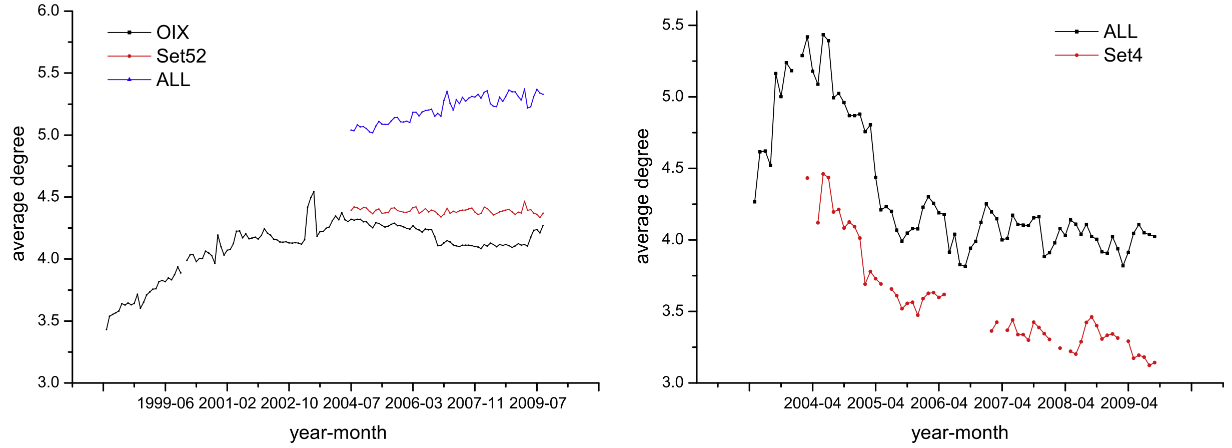
4.5. Degree distribution

Degree distribution is a frequently cited macroscopic topological property. Fig. 6 shows the complementary cumulative degree distribution (CCDF) of the IPv4 and IPv6 networks. It is clear that the networks follow a power-law degree distribution, $p(k) \sim k^{-r}$, as repeatedly reported before [1,14,40,26,47,30].

For the IPv4 network, the CCDF curves of different snapshots overlap with each other with a stale power-law exponent. For the IPv6 network, we see that the curves shift to the left as time goes with a slightly increasing power-law exponent.

4.6. Clustering coefficient

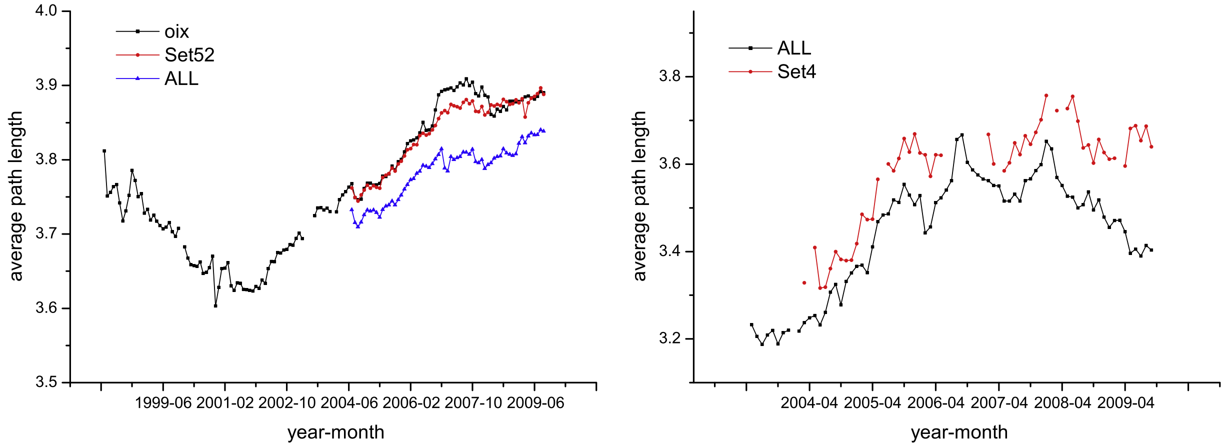
The clustering coefficient of a network is defined as three times the ratio of the total number of triangles to the total number of connected vertex triples in the network [49]. It measures the density of triangles in a network, which is relevant to alternative path



(a) IPv4 (b) IPv6

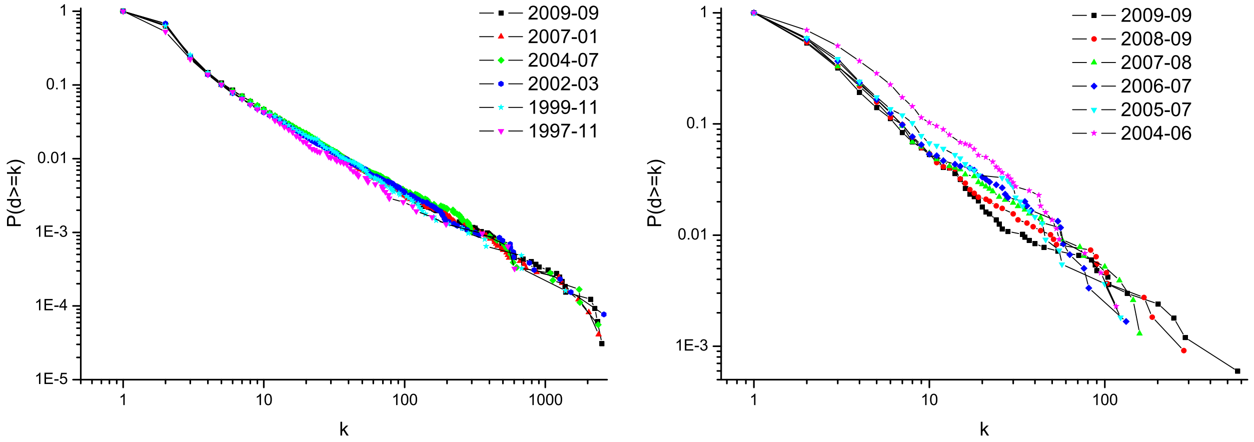
Fig. 4. Evolution of the average degree of the IPv4 and IPv6 networks.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004



(a) IPv4 (b) IPv6

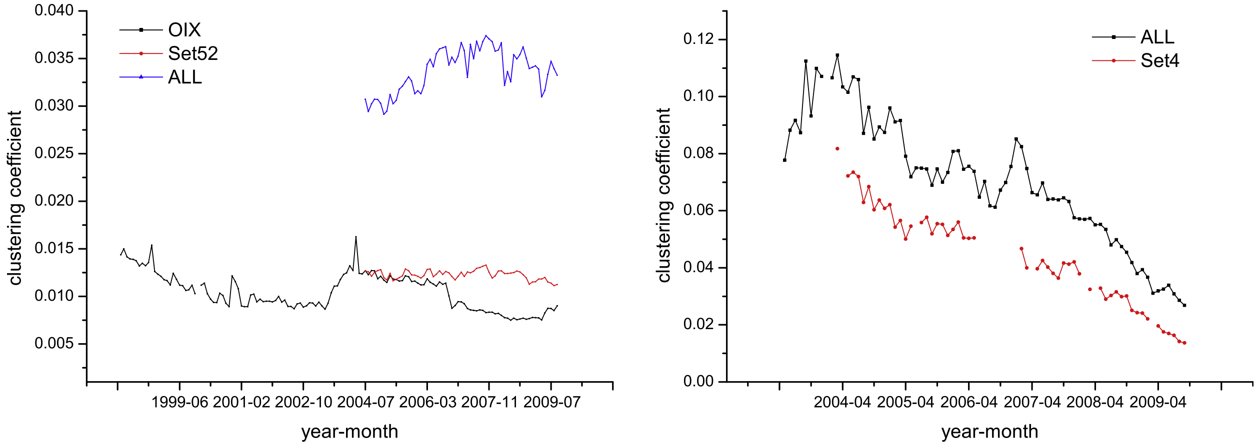
Fig. 5. Evolution of the average shortest path length of the IPv4 and IPv6 networks.



(a) IPv4 (OIX) (b) IPv6 (ALL)

Fig. 6. Evolution of degree distributions in the IPv4 and IPv6 networks.

and redundancy. Fig. 7 shows the evolution of clustering coefficient of the IPv4 and IPv6 networks. As reported in [22] the number of monitors has significant impact on the clustering coefficient. Here, we discuss the IPv4-Set52 and IPv6-Set4 data sets because they maintain the same monitors during their measurement time. The IPv4-Set52 data set shows that the clustering coefficient in the



(a) IPv4 (b) IPv6

Fig. 7. Evolution of clustering coefficients in the IPv4 and IPv6 networks.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004

IPv4 network is relatively stable in the past 5 years. In comparison, the clustering coefficient in the IPv6 network decreases gradually in recent years.

4.7. Assortative coefficient

Assortative coefficient [13] measures whether nodes tend to connect to nodes of similar degrees. AS-Level Internet topologies are known to be disassortative mixing with a negative assortative coefficient value, i.e., high-degree nodes tend to connect to low-degree nodes and vice versa. Fig. 8 shows the evolution of assortative coefficient of the IPv4 and IPv6 networks. We can see that both networks are disassortative mixing. The IPv6 network is more disassortative than the IPv4 network. An interesting observation is that the assortative coefficient of each of the two networks remains relatively stable over all time, and the variation in monitors has little influence on this metric. Hence, disassortative mixing could be viewed as an invariant for the AS-Level Internet topology.

5. Discussions

5.1. Internet evolutionary phase changes

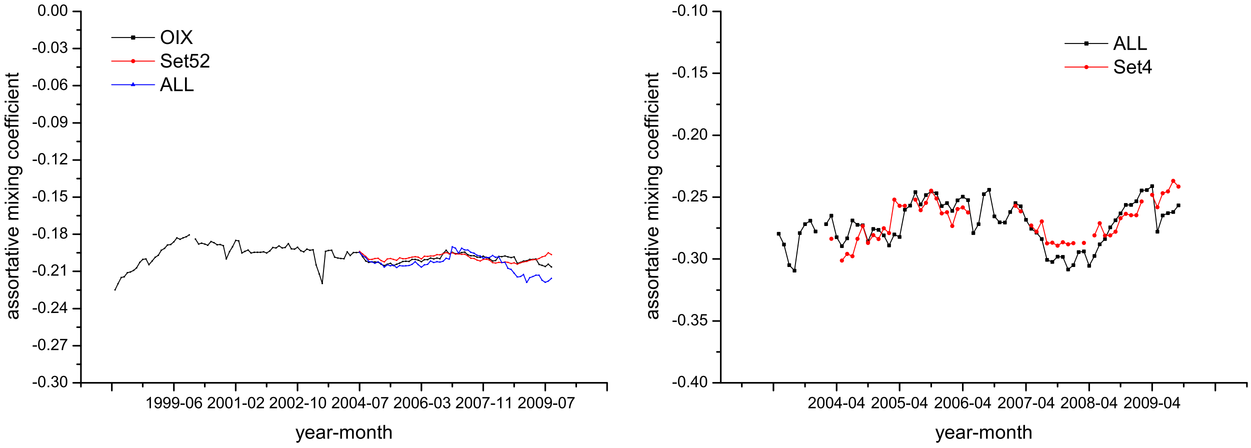
We summarise the above results in Table 1. It is clear that both the IPv4 and the IPv6 networks have experienced an evolutionary phase change. The phase change of the two networks, however, happened at different times with different transition patterns.

For the IPv4 network, the phase transition took place around year 2001 when the network changed from a process of rapid growth to a stage of slow growth with relatively stable structure. One possible reason could be the burst of the dot-com bubble at the beginning of this century which slowed down the investment on Internet. There might be technical reasons as well, such as the near exhaustion of AS numbers and the increasing size of BGP routing tables.

For the IPv6 network, the phase transition took place in year 2006 when the network changed from a stage of relatively slow growth to a process of rapid expansion. This may relate to a number of events happened around that time, including a boom of IPv6 deployment projects around the world, such as the CNGI project in China, and the plan to phaseout the 6bone, which is an IPv6 network that extensively relied on the IPv6-over-IPv4 tunnelling technique [51]. The exact reasons for the phase changes will be investigated in our future work.

5.2. IPv6-over-IPv4 tunnelling

One of the fundamental differences between the evolution of IPv6 and IPv4 is that the growth of the IPv4 topology followed the growth of the physical infrastructure of the Internet, whereas the growth of the IPv6 topology was more a matter of deploying the IPv6 technology over the existing (and changing) infrastructure. On the other hand, for the same reason the two networks are also related. The IPv6-over-IPv4 tunnelling, for example, has been widely used in the early years of IPv6 deployment. This technique allows two ASes on the IPv6 network appear to be directly connected with each other whereas in fact there might be a number of hops between them on the underlying IPv4 network. IPv6 network is more disassortative than the IPv4 network (see



(a) IPv4 (b) IPv6

Fig. 8. Evolution of the assortative coefficient of the IPv4 and IPv6 networks.

Table 1
Summary of the Internet evolutionary phase changes.

	IPv4 before 2001	IPv4 after 2001	IPv6 before 2006	IPv6 after 2006
Number of nodes	Exponential growth	Linear growth	Linear growth	Exponential growth
Number of edges	Exponential growth	Linear growth	Slow growth	Rapid growth
Maximum degree	Rapid growth	Stable	Slow growth	Rapid growth
Average degree	Steady growth	Stable	Rapid decreasing	Slow decreasing
Shortest path length	Decreasing	Increasing	Increasing	Stable
Degree distribution	Power-law		Power-law	
Power-law exponent	Stable		Slightly increasing	
Clustering coef.	Stable		Decreasing	
Assortative coef.	Stable		Stable	

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004

Fig. 8(b)), which may partly arise from the tunnelling deployment of IPv6.

To study the impacts of the tunnelling technique on the evolution of the IPv6 network, we use the following approach: for an edge (AS1, AS2) on an IPv6 snapshot, if AS1 and AS2 are also present in the corresponding IPv4 snapshot, we compute the shortest path length between AS1 and AS2 on the IPv4 snapshot. Intuitively, if IPv6-over-IPv4 tunnelling is prevalently used, then the shortest path length between AS1 and AS2 on the IPv4 snapshot will have high probability to exceed 1. We plot the evolution of the average shortest path length (ASPL) as well as the distribution of path length, d, for all such AS pairs in Fig. 9(a) and (b) respectively. It can be seen that the ASPL is indeed well above 1 and it decreases as time goes by. Also the probability that a connected pair of ASes on IPv6 are also directly connected on the IPv4 network (i.e. the columns with $d = 1$) increases over time. This suggests that the IPv6 gradually shifts from the tunnelling phase to the genuine IPv6 connectivity phase. This is in accordance with the IETF's phaseout planning of the 6bone in 2004 [51]. We expect this trend towards the deployment of genuine IPv6 sessions will bring a diminishing difference between the length of IPv6 and IPv4 paths.

5.3. Internet measurement monitors

As any work based on Internet measurement data, our work would still be affected by the limited number of vantage points. All along the paper, we try however as much as possible to show how the number of monitors affects each metric. In general, we note that data from one collector is insufficient for the IPv4 network. Set52 and ALL data sets often capture very similar network structures although the ALL data set contains significantly more monitors. For the IPv6 network, the ALL data set is more appropriate than the Set4 data set.

5.4. Generative AS-Level Internet models

A number of generative models have been proposed to reproduce and explain the evolution of networks. These models are different in many ways but they are all based on a common assumption that a network obeys a non-changing, uniform growth mechanism throughout its evolution. This, however, is clearly not the case for the IPv4 and IPv6 networks.

Taking the maximum degree as an example, Fig. 10 shows that the maximum degree grows monotonically with the number of nodes for each of the four typical AS-Level generative models, whereas the maximum degree of the IPv4 network (OIX data set) exhibits an evolutionary phase change where after a critical point the maximum degree became relatively stable. This is not surprisingly as the models were not designed to reproduce such phase change.

Our work highlights that it is not sufficient to validate a generative Internet model against a few snapshots of the network. Rather, we should validate a model against long-term evolution data. Our results on the evolutionary phase changes of the Internet networks provide new input for designing and validating future Internet models.

Indeed, pure graph-theory based generative models have already been questioned in both the router-level and AS-Level topologies [41,43,38,28,39]. It has been recognized that the router-level topology can be more accurately modeled by optimization-driven approaches, e.g., HOT [43] and IGen [41], and could be designed in a cost-effective manner [50]. However, these challenges are not raised from the perspective of the phase change. A possible direction for future research could be to borrow some ideas from the optimization-driven approaches to the development of AS-Level generative models that capture the phase change.

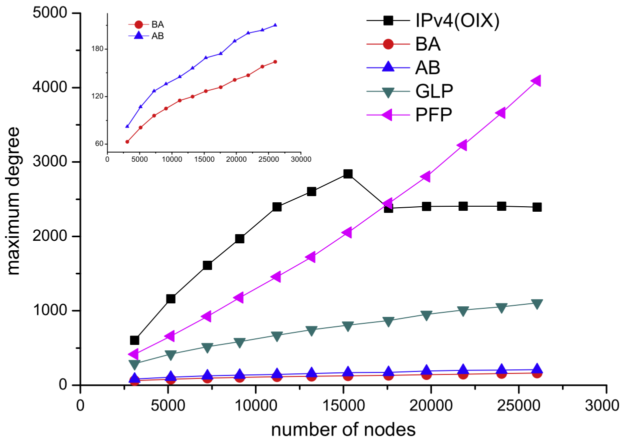


Fig. 10. Evolution of maximum degree of real AS topology (OIX) and four typical generative models. BA network is generated with $m = 2$ [14]. AB network is generated with $p = 0.2$, $q = 0.3$ and $m = 2$ [15]. GLP network is generated with $p = 0.4695$ and $\beta = 0.6447$ [16]. PFP network is generated with $p = 0.4$ and $\delta = 0.021$ [18]. Evolution of maximum degrees in BA and AB is further illustrated in the inset since the growth trends are not legible in the original plot due to the relatively small absolute values.

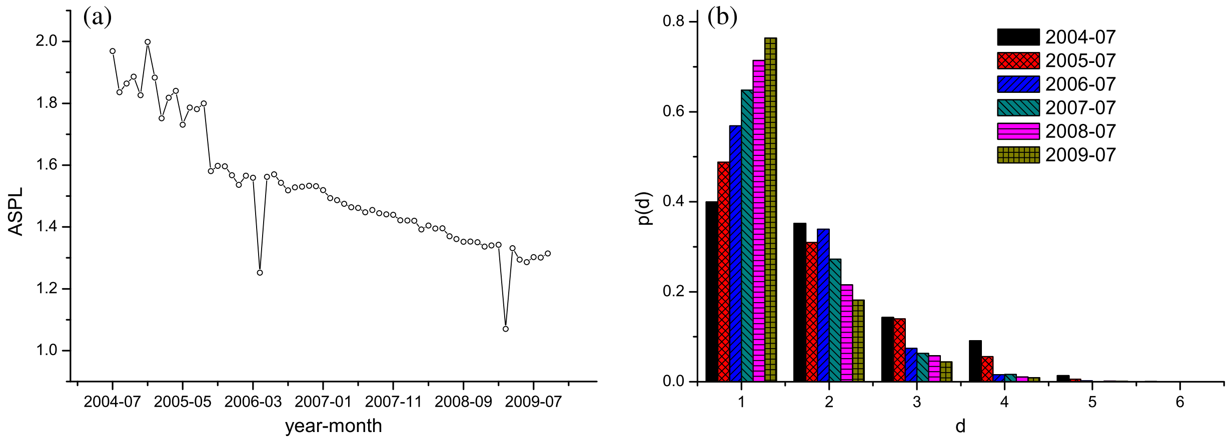


Fig. 9. The impact of IPv6-over-IPv4 tunnelling on the Internet evolution. (a) Evolution of the average shortest path length (ASPL) of the connected IPv6 AS pairs on the IPv4 network. (b) Evolution of the distribution of shortest path length of the connected IPv6 AS pairs on the IPv4 network, where $p(d)$ is the probability that a connected pair of ASes on the IPv6 network has path length d on the corresponding IPv4 network.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004

6. Conclusion

In this paper, we performed an in-depth side-by-side study of the evolution of the IPv4 and IPv6 Internet topologies at the autonomous system level based on historic data over a long period of time. Ample evidence shows that both networks have undergone a phase change in their evolution process. For the IPv4 network, the approximate phase transition occurred around 2001; while for the IPv6 network, the phase transition took place around late 2006. The phase transition pattern of the two networks are quite different. While the IPv4 slowed down from a rapid growth, the IPv6 has just engaged in a fast expansion. We also found that the IPv6-over-IPv4 tunnelling deployment scheme partly affects the evolution of the IPv6 network.

Our work fundamentally changes our knowledge on the Internet topology evolution. It provides valuable input for refining existing network models or developing new models. It also opens interesting questions for future work, such as the exact reason for the evolutionary phase changes of the IPv4 and IPv6 networks and the possibility of phase changes in the future.

Acknowledgement

GZ is supported by National Natural Science Foundation of China under Grant No. 60673168, and the Hi-Tech Research and Development Program of China under Grant No. 2008AA01Z203. SZ is supported by the Royal Academy of Engineering and EPSRC (UK) under Grant No. 10216/70.

References

[1] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the Internet topology, in: Proc. ACM SIGCOMM'99, 1999.
[2] J. Leskovec, J. Kleinberg, C. Faloutsos, Graphs over time: densification laws, shrinking diameters and possible explanations, in: Proc. ACM SIGKDD'05, 2005.
[3] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, The Internet AS-Level topology: three data sources and one definitive metric, ACM SIGCOMM Computer Communication Review 36 (2006).
[4] B.C. Zhang, R. Liu, D. Massey, L.X. Zhang, Collecting the Internet AS-Level topology, ACM SIGCOMM Computer Communication Review (2005).
[5] H. Chang, R. Govindan, S. Jamin, S.J. Shenker, W. Willinger, Towards capturing representative AS-Level Internet topologies, in: Proc. ACM SIGMETRICS'02, 2002.
[6] L.X. Gao, F. Wang, The extent of AS path inflation by routing policies, in: IEEE Global Internet Symposium, 2002.
[7] G.Q. Zhang, G.Q. Zhang, S.Q. Cheng, T. Zhou, Symbiotic effect: a guideline for network modeling, Europhysics Letters 87 (2009) 68002.
[8] Availabel from: <www.caida.org>.
[9] Available from: <www.netdimes.org>.
[10] Available from: <www.routeviews.org>.
[11] Available from: <www.ripe.net>.
[12] S. Zhou, R. Mondragón, The rich-club phenomenon in the Internet topology, IEEE Communications Letters 8 (2004) 180–182.
[13] M.E.J. Newman, Assortative Mixing in Networks, Physical Review Letters 89 (2002) 208701.
[14] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 09512.
[15] R. Albert, A.L. Barabási, Topology of evolving networks: local events and universality, Physical Review Letters 85 (2000) 5234.
[16] T. Bu, D. Towsley, On distinguishing between Internet power-law topology generators, in: Proc. IEEE INFOCOM'02, 2002.
[17] S. Zhou, R. Mondragón, Accurately modeling the Internet topology, Physical Review E 70 (2004) 066108.
[18] S. Zhou, Understanding the evolution dynamics of the Internet topology, Physical Review E 74 (2006).
[19] G.Q. Zhang, Measuring the impacts of sampling bias on Internet AS-Level topology inference, in: Proc. 2009 International Conference on Communications and Mobile Computing, 2009.
[20] R. Cohen, D. Raz, The Internet dark matter-on the missing links in the AS connectivity map, in: Proc. IEEE INFOCOM'06, 2006.
[21] Y. He, G. Siganos, M. Faloutsos, S. Krishnamurthy, Lord of the links: a framework for discovering missing links in the Internet topology, IEEE/ACM Transactions on Networking 17 (2009).
[22] R. Oliveira, D. Pei, W. Willinger, B.C. Zhang, L.X. Zhang, In search of the elusive ground truth: the Internet's AS-Level connectivity structure, in: Proc. ACM SIGMETRICS'08, 2008.
[23] D. Chakrabarti, C. Faloutsos, Graph mining: laws, generators, and algorithm, ACM Computing Surveys 38 (2006).
[24] A. Dhamdhere, C. Dovrolis, Ten Years in the evolution of the Internet ecosystem, in: Proc. IMC'08, 2008.
[25] G.Q. Zhang, G.Q. Zhang, Q.F. Yang, S.Q. Cheng, T. Zhou, Evolution of the Internet and its cores, New Journal of Physics 10 (2008) 123027.
[26] M.E.J. Newman, The structure and function of complex networks, SIAM Review 45 (2) (2003) 167–256.
[27] R. Oliveira, B.C. Zhang, L.X. Zhang, Observing the evolution of Internet AS topology, in: Proc. ACM SIGCOMM'07, 2007.
[28] H. Chang, S. Jamin, W. Willinger, To peer or not to peer: modeling the evolution of the Internet's AS-Level topology, in: Proc. IEEE INFOCOM'06, 2006.
[29] B. Huffaker, K. Claffy, Visualizing IPv6 AS-Level Internet topology 2008. Available from: <http://www.caida.org/research/topology/as_core_network/ipv6.xml>.
[30] B. Xiao, L.D. Liu, X.C. Guo, K. Xu, Modeling the IPv6 Internet AS-Level topology, Physical A 388 (2009) 529–540.
[31] Y. Hyun, A. Broido, K.C. Claffy, On third-parth addresses in traceroute paths, in: Proc. of Passive and Active Measurement Workshop (PAM), 2003.
[32] Z.M. Mao, J. Rexford, J. Wang, R.H. Katz, Towards an accurate AS-Level traceroute tool, in: Proc. ACM SIGCOMM'06, 2006.
[33] Y. Zhang, Z. Zhang, Z.M. Mao, Y.C. Hu, B.M. Maggs, On the impact of route monitor selection, in: Proc. IMC'07, 2007.
[34] S.N. Dorogovtsev, J.F.F. Mendes, Effect of the accelerating growth of communications networks on their structure, Physical Review E 63 (2001) 25101.
[35] D.M.D. Smith, J.P. Onnela, N.F. Johnson, Accelerating networks, New Journal of Physics 9 (2007) 181.
[36] L.X. Gao, On inferring autonomous system relationships in the Internet, IEEE/ACM Transactions on Networking 9 (6) (2001).
[37] X. Dimitropoulos, D. Krioukov, G. Riley, K. Claffy, Revealing the autonomous system taxonomy: the machine learning approach, in: Proc. Passive and Active Measurement Conference (PAM), 2006.
[38] H. Chang, S. Jamin, W. Willinger, Internet connectivity at the AS-Level: an optimization-driven modeling approach, in: Proc. ACM SIGCOMM Workshop on MoMeTools, 2003.
[39] H. Haddadi, S. Uhlig, A. Moore, Modeling Internet topology dynamics, ACM SIGCOMM Computer Communication Review 38 (2008).
[40] S. Zhou, G.Q. Zhang, G.Q. Zhang, Chinese Internet AS-Level topology, IET Communications 1 (2007) 209–214.
[41] B. Quoitin, V. Van den Schrieck, P. Francois, O. Bonaventure, IGen: generation of router-level Internet topologies through network design heuristics, in: Proc. 21st International Teletraffic Conference, 2009.
[42] H. Tangmunarunkit, Ramesh Govindan, Sugih Jamin, Network topology generators: degree-based vs. structural, in: Proc. ACM SIGCOMM'02, 2002.
[43] L. Li, D. Alderson, W. Willinger, J. Doyle, A first-principles approach to understanding the internet's router-level topology, in: Proc. ACM SIGCOMM'04, Oregon, 2004.
[44] W. Mühlbauer, S. Uhlig, B. Fu, M. Meulle, O. Maennel, In search for an appropriate granularity to model routing policies, in: Proc. ACM SIGCOMM'07, Kyoto, Japan, 2007.
[45] W. Mühlbauer, O. Maennel, S. Uhlig, Building an AS-topology model that captures route diversity, in: Proc. ACM SIGCOMM'06, Pisa, Italy, 2006.
[46] H. Haddadi, M. Rio, G. Iannaccone, A. Moore, R. Mortier, Network topologies: inference, modeling, and generation, IEEE Communications Surveys and Tutorials 10 (2) (2008).
[47] W.M. Eddy, Basic Properties of the IPv6 AS-level topology, ACM SIGMETRICS Performance Evaluation Review 36 (3) (2008) 50–57.
[48] Available from: <http://www.caida.org/data/active/as_taxonomy/>.
[49] S.N. Dorogovtsev, Clustering of correlated networks, Physical Review E 69 (2004) 027104.
[50] G.Q. Zhang, On cost-effective communication network designing, Europhysics Letters 89 (2010) 38003.
[51] R. Fink, R. Hinden, 6bone(IPv6 Testing Address Allocation) Phaseout, in: IETF RFC 3701, 2004.

Please cite this article in press as: G. Zhang et al., Phase changes in the evolution of the IPv4 and IPv6 AS-Level Internet topologies, Comput. Commun. (2010), doi:10.1016/j.comcom.2010.06.004