

The Effect of AS's Geographic Locations on Internet's Stability

Sun Jing-hao^{1,a}, Diao Ding¹, Huang Ling-jin¹ and Meng Ya-kun¹

¹School of Computer and Communication Engineering, Northeastern University, Qinhuangdao, China

^ajhsun@mail.neuq.edu.cn

Keywords: Complex Network; AS's Geographic Location; Hypergraph; Network Stability; Toughness

Abstract. This paper analyzes how the AS's geographic locations and transmission distances affect on Internet's stability. In contrast to the traditional studies focusing on the future AS network which based on nowadays, this paper emphasizes particularly on the change of AS network's performance under the effect of the economic adjusting or the technology innovation to help make adjusting measures and to lead the technological trend. A double-deck network model of geographic hypergraph and AS network is presented to simulate the geographic locations, which is not reflected in traditional models. The simulate results of our model show that Internet's stability ascends with the increase of transmission distances. However, if AS's geographic locations are random, the transmission distances has no obvious effects on Internet's stability.

Introduction

Internet AS-layer network has network topological properties such as scale-free and high clustering. Research shows that, the topological characteristics of congestion properties and stability have a great impact on the key of network performance. Therefore, a large number of scholars around the AS network topology causes and characteristics evolution conduct research. The previous studies considered from the view of influencing factors can be grouped into two broad categories:

1) The study based on economic factors: In 1999, three Faloutsos brothers found AS scale-free networks, Barabasi and Albert attributed it to the results of the economic factors "Matthew" effect. Since then, a large number of scholars put forward numerous connectivity-based network model which is based on economic factors^{[7]-[15]}.

2) The study based on geographical factors: as we all know, the AS network of clustering structure is mainly influenced by the geographical factors. But the number of studies based on the geographical factors is relatively small^{[16]-[19]} and we find that these studies in the AS model abstraction ignored the geographic distribution, so geographical factors' effects on network performance can not correctly be reflected.

Therefore, this article will focus on the establishment of network model which can correctly reflect the AS geographic distribution, to explore the impact of geographical factors on network performance. In addition, we will conduct a more important try: from a new research point of view, then ask new questions.

Previous studies have mainly focused on the status and development trend of AS network topology in today's economic and technological environment. However, these studies can not answer this question:

I . Once the macro-control on the economy has changed the geographic distribution of the AS, how will AS network topology and performance be affected?

II .If breakthrough in the link laying and maintenance technology makes the length of the connection increases and costs reduced in routers built, which kind of changes will happen on the topology and performance of network?

Unlike previous research point of view, the above problems under the assumption that the Internet network environment factors largely change, and then explore the network topology and performance. The study of these problems can provide a theoretical basis for the development of the macro-control measures and grasping network technology oriented.

The structure of the paper is organized as follows: First, introduce the AS network geographical factors needed to be taken into account, and point out the flaws of the previous model. Then, give a structure of AS geography properties, and model the AS networks as the geographic hypergraph. Finally, experiments with geographical factors are given.

Geographical factors of AS Networks

AS geographical distance between built even and geographical distribution are the main AS network geographical factors. We found that the two geographical factors will affect network performance. However, previous studies have ignored the geographic distribution of AS, only considering geographical distance threshold when they establish a connection between ASs, each AS scope are confined to a small fixed area, connection between the ASs were divided into regional connections and inter-regional connections^{[16][19]}, as shown in figure(1-1).

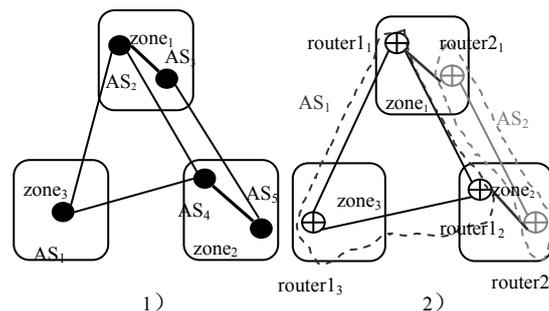


Fig.1 Differences in geographical distribution between traditional model and the fact

In fact, AS is constituted by routers, AS router includes border router and internal router .Two AS through each border router can connect with each other. In addition, an AS border router may be distributed in different geographic regions, rather than be confined to a fixed area^{[18][20]}, but only when the two AS boundary routers are in the same area, it is possible to build the edges between two AS^{[17][18][21]},as shown in figure(1-2).

Therefore, AS geographical distribution is a reflection of border routers located in different geographic regions in fact. Below we will elaborate on relationship among the router geographical distance threshold to establish a connection, AS boundary router and internal router, discussing the influence of geographical factors.

2. 1 Threshold value of the geographical distance for connecting routers

Previous studies considered impacts of geographical factors on evolution of network topology, as typical representative of geographical distance^{[16][19],[21]}. Study shows that in Internet network there exists a threshold value, more than 75% of routers to establish a connection between distances are within range of the threshold value^{[17][18]}. We assume all the routers to establish the distance connection will not exceed the threshold value. Between AS needed to connect through border router, the number of the AS border router

can directly connect AS are proportional to a threshold value, as shown in Figure 2, threshold value $R_1 < R_2$, then $routeNum(S_1) < routeNum(S_1 \cup S_2) \cdot routeNum(S)$ represents the number of routers in area S.

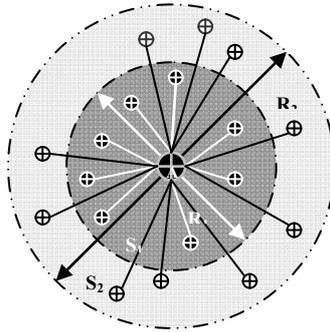


Fig.2 The number of the AS border router can directly connect AS are proportional to a threshold value

2.2 Border router

An AS border router is the router which directly connected to routers in other AS. Because of the threshold limit distance in different AS to build even between router to establish a connection, each border router must be distributed in same geographic area. Accordingly, the geographical distribution boundary router determines the geographical scope AS connection.

2.3 Internal router

Internal router is the router which not directly connected to routers in other AS. The role of internal router is to make connections between AS boundary routers. Because of threshold limits even, the farther distance away border router are, the more internal router you need to make connection, the higher price of link connection and maintenance to pay. In addition, we expect to exhaust less AS boundary router to provide services area as much as possible. Therefore, the larger the distance between adjacent an AS, the more extensive area single border router can provide services, show in figure 3.

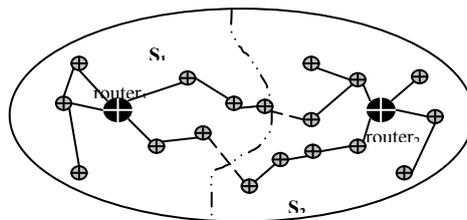


Fig. 3 Interconnection of border router in fixed area

AS needed make trade-offs between the area adjacent to geographical distance between border routers and coverage range. In the following analysis we will introduce AS boundary router's regional expansion mode to characterize this trade-off.

Geographic hypergraph structure

According to the description of the second, the service area of an AS is a collection of all its border router built even regional, service area collection of all AS is AS network border router built even regional clustering. In the field of operations research and optimization, hypergraph theory is usually used to study the clustering optimization problems. This paper presents the Geographical Hypergraph Structure (GHS) to characterize the AS network geographic distribution.

The world limited geographic area is divided into N zones, $ZONE = \{zone_1, zone_2, \dots, zone_n\}$ adjacent on the actual geographic locate adjacently in hypergeographic area.

Define a hypergraph on Internet AS-layer network, $G = (AS_1, AS_2, \dots, AS_m)$ as a limited subset of the cluster on zone, meet the conditions:

$$(1) AS_i \neq \emptyset \quad (i = 1, 2, \dots, m) \quad (2) \bigcup_{i=1}^m AS_i = ZONE$$

In hypergraph G , we define the element $zone_1, zone_2, \dots, zone_n$ in $ZONE$ as node, the set AS_1, AS_2, \dots, AS_m as edge. We do not consider isolated nodes. The graphical representation of the geographic hypergraph structure shown in Figure 4.

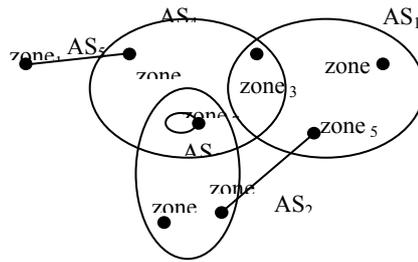


Fig.4 Geographic hypergraph structure

Model

Based on the geographical hypergraph and AS network interaction mechanism, this paper establish AS network and geographic hypergraph double-deck network model are as follows:

(1) Initialization

Geographical hypergraph: Initial geographic hypergraph contains n nodes $zone_i$ ($i = 1, 2, \dots, n$) and sides $AS_i = \{zone_i\}$ ($i = 1, 2, \dots, n$). AS network: initial AS network is a complete graph which contains n_0 AS nodes;

(2) At each time step h

a. Add a new AS node

Geographical hypergraph: Add new edges called AS_i , select node $zone_j$ randomly that make the node $zone_j \in AS_i$. AS network: Add new nodes called AS_i , and select m nodes randomly from $zone_j$. The nodes belong to set $\{AS_t | t = 1, 2, \dots, n \wedge AS_t \neq AS_i\}$ and build connection with AS_i .

b. Increase n internal edges

Geographical hypergraph: execute step c ($p = 1$), select the edge AS_i and the node $zone_j$, to make $zone_j \in AS_i$ and make $\Omega = \{AS_t | AS_t \cap AS_i \neq \emptyset\}$. AS network: Select n nodes from randomly from set $\Omega = \{AS_t | AS_t \cap AS_i \neq \emptyset\}$ and build connection with AS_i .

c. With probability p were node AS geographical area expansion (p is an adjustable parameter)

AS network: According to the probability $q = \frac{k_i}{\sum_j k_j}$, choose node AS_i , in which k_i means the AS_i

degree value. Geographical hypergraph: If this modeling is set by random expansion mode: select node $zone_j$ randomly that make the node $zone_j \in AS_i$;

(3) repeat step (2), until the geographic hypergraph meets the definition of 2.1, and the AS network reach the required size.

Simulation

In this section, we will analyse two cases about the AS to geographical area random expansion and neighbor expansion, the effect of the number $zoneNum$ of geographical area on characteristics and stability of the network. Among them, the random expansion portrays that the AS border router settings is a macro-control influence; Neighbors expansion considers the costs of internal router maintenance and link laying; And $zoneNum$ distance threshold is inversely proportional, distance threshold also reflects the technology level of router link laying. Therefore, $zoneNum$ reduce said that between routers link laid technology to improve.

By adjusting the number $zoneNum$ of geographical area, the influence of the network characteristics can be concluded, as shown in figure 6.

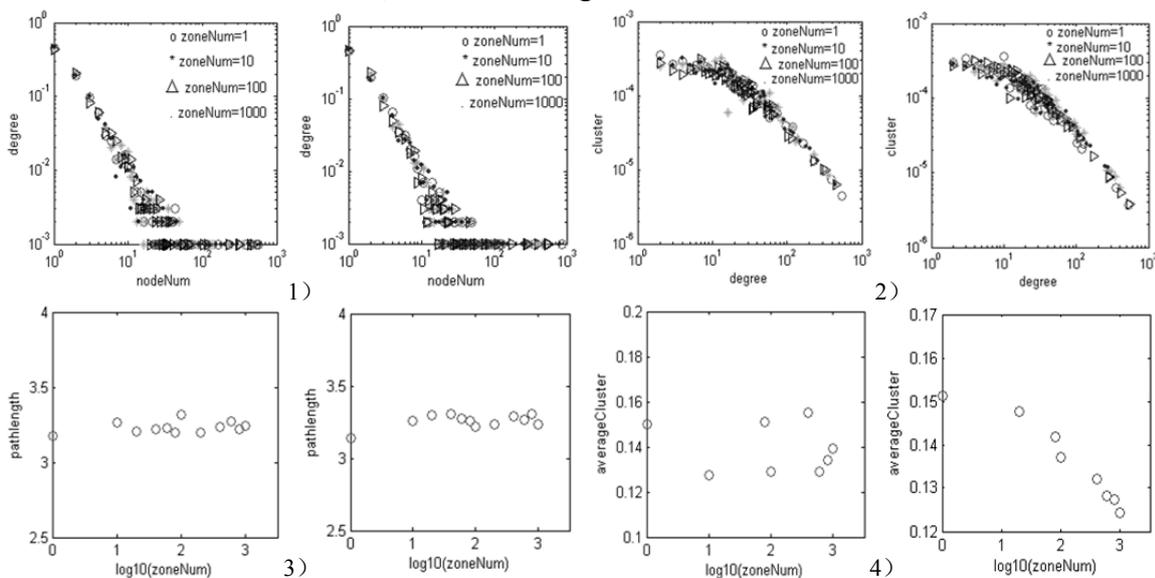


Fig. 6: the influence of number $zoneNum$ of area on network characteristics

In the picture, on the left side said AS take random expansion mode, on the right side said AS take neighbor expansion mode. Among them, 1) the influence of $zoneNum$ on degree distribution; 2) the influence of $zoneNum$ on degree and concentration coefficient correlation; 3) the influence of $zoneNum$ on Characteristic path length; 4) the influence of $zoneNum$ on average concentration coefficient.

From figure 6, we can see, for the network degree distribution, degree and concentration coefficient correlation and characteristic path length of these features, no matter AS takes random expansion or neighbor expansion, $zoneNum$ almost does not have apparent effect on them. The average concentration coefficient for network, when AS takes random expansion, $zoneNum$ has no obvious influence; When AS takes neighbors expansion, with increasing of $zoneNum$, the network average concentration coefficient show attenuation trend.

Acknowledgments

This research was performed at Northeastern University and was supported by the Science and Technology Research and Development Program of Qinhuangdao No.2012021A101. The support is gratefully acknowledged.

References:

- [1] M Faloutsos, P Faloutsos, C Faloutsos. On power-law relationships of the Internet topology. ACM SIGCOMM Computer Communication Review, 29(4): 251~262 (1999)
- [2] G Siganos, P Faloutsos, C Faloutsos. Power laws and the AS-level Internet topology. IEEE/ACM Transactions on Networking, 11(4): 514~524 (2003)
- [3] T Bu, D Towsley. On Distinguishing between Internet Power Law Topology Generators. IEEE INFOCOM, 2:638~647(2002)
- [4] K Klemm, VM Eguiluz. Highly Clustered Scale-Free Networks. Physical Review E,65(3) 036123~036128(2002)
- [5] Z Toroczkai, KE Bassler, Jamming is limited in scale-free systems. Nature, 428:716~717(2004)
- [6] R Albert, H Jeong, AL Barabasi. Error and attack tolerance of complex networks. Nature,406: 387~482 (2000)
- [7] J Winick, S Jamin. Inet-3.0: Internet topology generator. Technical report , Department of EECS, University of Michigan,(2002)
- [8] Albert R, Barabasi A L. Topology of evolving networks: Local events and universality. Phys. Rev. E, 70:066~108(2004)
- [9] A Medina, A Lakhina, I Matta, J Byers. BRITE: An approach to universal topology generation. Proceedings of MASCOTS, Washington, 346~353(2001)
- [10] T Bu, D Towsley. On distinguishing between Internet power law topology generators. Proceeding of INFOCOM, New York,2: 638~647(2002)
- [11] S. Zhou and R. J. Mondragon, Accurately modeling the Internet topology. Phys. Rev. E, 70:030~50(2004)
- [12] S. Zhou and R. J. Mondragon. Towards modeling the Internet topology-the interactive growth model. Teletraffic science and engineering, 5: 121~130(2003)
- [13] ST Park, DM Pennock, CL Giles. Comparing static and dynamic measurements and models of the Internet's topology. Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies, 3: 1616~1627(2004)
- [14] B Sagy, G Mira, W Avishai. An incremental super-linear preferential Internet topology model. Proc 5th Annual Passive and Active Measurement Workshop, LNCS 3015, 53~62(2004)
- [15] Wang Xiaoming, L Dmitri. Wealth-Based Evolution Model for the Internet AS-Level Topology. Proceeding of INFOCOM, New York(2006)
- [16] B Sagy, G Mira, W Avishai. A geographic directed preferential Internet topology model. Arxiv: CS, (2005)
- [17] A Lakhina, JW Byers, M Crovella, I Matta. On the geographic location of Internet resources(short abstract), Proc. SIGCOMM Internet Measurement Workshop(IMW' 02), 249~250 (2002)
- [18] A Lakhina, JW Byers, M Crovella, I Matta. On the geographic location of Internet resources. IEEE Journal on Selected Areas in Communications, 21: 934~948(2003)
- [19] SH Yook, H Jeong, AL Barabasi. Modeling the Internet's large-scale topology. Proc. Nat. Acad. Sci., 9(21):13382~13386(2002)

- [20] A Aditya, C Shuchi, K Arvind, S Srinivasan. On the Scaling of Congestion in the Internet Graph. ACM SIGCOMM Computer Communication Review, 34(3):43~56(2004)
- [21] H Tangmunarunkit, J Doyle, R Govindan, S Jamin, S Shenker, W Willinger. Does AS size determine degree in AS topology. ACM SIGCOMM Computer Communication Review, 31(5): 7~10 (2001)