Correlation Properties and Self-similarity of Renormalization Email Networks

Lianming Zhang[1], Sundong Liu[2], Yuling Tang[1], and Hualan Xu[1]

[1] College of Physics and Information Science, Hunan Normal University, Changsha, 410081, P.R. China
lianmingzhang@gmail.com, tyl_ok@sina.com, hnhlxu@gmail.com

[2] Shenzhen Institute of Information Technology, Shenzhen, 518029, P.R. China
liusd@sziit.com.cn

Abstract. A degree-thresholding renormalization method is recently introduced to find topological characteristics of some complex networks. As a matter of fact, the applicability of these characteristics depends on the level or the type of complex networks. Here, a modified version of this original algorithm is presented to unravel ubiquitous characteristics of observed email networks and obtain correct understanding of underlying evolutionary mechanism. Some topology metrics of the email networks under renormalization were analyzed. The results show that renormalization email networks have the power-law distribution with double exponents, are disassortative and become assortative after half of total renormalization steps, have high-clustering coefficients and rich-club phenomena. These characteristics are self-similar both before and after renormalization until half of total renormalization steps, otherwise are self-dissimilar.

Keywords: complex networks, email networks, topological properties, self-similarity, renormalization.

1 Introduction

The correct understanding of Internet topology is important to generate real network topologies, construct accurate network simulation environments and promote the development of network applications, protocols and routing architectures [1].

There are three different layers at which to describe Internet topology [2]: the link layer, the network layer, and the application layer topologies. The network layer topology can itself be seen at five different levels: the IP interface, the router, the Point of Presence (PoP), the autonomous system (AS), and the Internet Service Provider (ISP) levels. At the application layer, there are the World Wide Web (WWW) linkage structure, the email network and the peer to peer (P2P) topology, and so on. However, some people also assign the email network as a social network, the WWW as an information network, and the other level topologies, such as the AS level and the router level topologies, as technological networks [3].

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1846–1859, 2009.

In the past decade, the Internet topological properties, especially at the AS level, the router level, and the WWW of the application layer, have attracted substantial attention. Research shows that some properties appear at various level topologies or different type networks, while other properties only exist in the specific level topologies or the some type networks. For example, at the three aforementioned levels, the Internet topologies exhibit the properties of the power-law distribution, the small-world behavior [4][5]. The AS level and the WWW structure both show disassortative mixing on their degrees [6], and further, the AS level also has a rich-club phenomenon [7]. Recently, the self-similarity of the WWW structure had been revealed under a length-scale transformation [8], but the AS level and the router-level topologies are not self-similar [9]. Even for them, there have some different properties [10][11].

The properties of the email network were analyzed firstly in [12]. The nodes correspond to email addresses and the links correspond to emails using data from server log files. The resulting network exhibits the power-law distribution with negative exponent $\gamma = -1.81 \pm 0.10$, and pronounced the small-world behavior with high local clustering coefficient $C = 3.44 \times 10^{-2}$ and small mean shortest path length $l = 4.95 \pm 0.03$. Email is currently a most used service of Internet, so the email virus is widely used as a major form of computer virus. The spreading of email viruses is dissimilar in networks with different topological properties [13]. It is greatly facilitated in real email networks with power-law distribution compared to random networks [6]. With the flood of email viruses in today's communication experience, it is important to reveal the other potential properties of the real email networks and to deeply understand the underlying evolutionary mechanism of the real email networks that leads to these common properties.

To obtain further understanding of the real email networks, features of the *renormalization email networks* (REMNs) based on a *degree-thresholding renormalization* (DTR) method was investigated. In this paper, the average degree, the power-law distribution, the assortativity coefficient, the clustering coefficient and the rich-club connectivity with different values of degree threshold of the REMNs were analyzed.

2 Degree-Thresholding Renormalization Algorithm

Renormalization is a useful method for analyzing the structure and form of a complex system. Its essence is to change the amount of coarse-graining of the complex system in measurement, in order to analyze or reveal some features and laws of the changing process of physical quantities of the complex system.

The box-covering renormalization method was firstly applied to some complex networks [8]: the WWW, the social network (*Actors*), the biological networks (*Esche-richia coli* and *Homo sapiens*) and the cellular networks (*Archaeoglobus fulgidus E. coli* and *Caenorhabditis elegans*), and the results show that there is a power-law distribution of self-similar structure: $N_B \sim l_B^{-d_B}$, where N_B is the number of boxes needed to cover the complex network, l_B is the size of the box, and d_B is the fractal dimension or box dimension. A power-law distribution for all boxes with different length-scales was found as follows: $P(k) \sim k^{-\gamma}$, where $P(k)$ is the degree distribution, γ is known as the degree exponent of the power-law distribution. However, there are many coarse-graining methods for a complex network, and the structure of the complex network is self-similar for some coarse-graining methods but it is self-dissimilar for others [14].

Recently, an attempt of a new coarse-graining method to apply a DTR procedure to the Internet AS and the social networks (*airport networks*) and to the randomized observed networks preserving the degree distribution was provided in [15], and the results show that degree distributions and degree-degree correlations are self-similar both before and after renormalization with different values of threshold, and clustering of the observed networks is self-similar, but one of the randomized observed networks is not.

In this section, a detailed study of the DTR method used to calculate quantity charactering the topology of email networks was presented. For a *observed original email network* (OOEMN) G and threshold k_T, the REMNs $G(> k_T)$ induced by nodes with degrees $k > k_T$ and the hidden REMNs $G'(\leq k_T)$ induced by nodes with degrees $k \leq k_T$ are extracted from G. For $k_T = 0$, then $G(> k_T) = G$, and the hidden REMN $G'(\leq k_T) = \emptyset$, where are no nodes without links; $G(> k_T) = \emptyset$, and $G'(\leq k_T) = G$ for $k_T > k_{max}$, where k_{max} is the maximum degree of the OOEMN. A modified version of the DTR algorithm is implemented as follows:

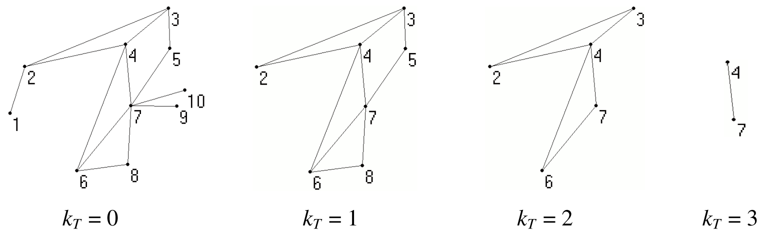
1. Assign a unique *id* form 1 to n to all nodes of the OOEMN, and given a threshold k_T.
2. Apply a two-dimensional array $A[x_i, y_i]$ to store the link $x_i \rightarrow y_i$ of the OOEMN, x_i and y_i represent separately the *id* of a node form 1 to n.
3. Sort the array. Firstly, sort ascending by the value of the first column x_i, and if the value is the same, then sort ascending by the value of the second column y_j.
4. Eliminate the rows for $x_i = y_i$, and eliminate the rows for $x_i = y_j$ and $x_j = y_i$, generate the new array $B[x_i, y_i]$.
5. Calculate the maximum degree k_{max} of the OOEMN.
6. Set the row $r = 1$, the new array $C = [\,]$ and $D = [\,]$. Repeat the following until r is equal to the length of the array $B[\,]$.

6.1. Calculate the occurrence number num_1 and num_2 of the array element $B[r, 1]$ and $B[r, 2]$, respectively.

6.2. If $num_1 \leq k_T \mid num_2 \leq k_T$, append the array elements $B[r, :]$ to the array $C[\,]$ of the REMN $G(> k_T)$, else append the array elements $B[r, :]$ to the array $D[\,]$ of the hidden REMN $G'(\leq k_T)$.

6.3. Increase r by 1.

In Figure 1 a typical example (10 nodes and 13 undirected links) to illustrate the DTR method was presented. The first graph with $k_T = 0$ depicts the observed original network. The following resulting networks with $k_T = 1$, 2 or 3 exhibit that the method is applied to it, respectively. For instance, in the case of $k_T = 3$, there is a resulting sub-network with 2 nodes and 1 link after the method applied to it, and 8 nodes and 12 links are eliminated. The eliminated

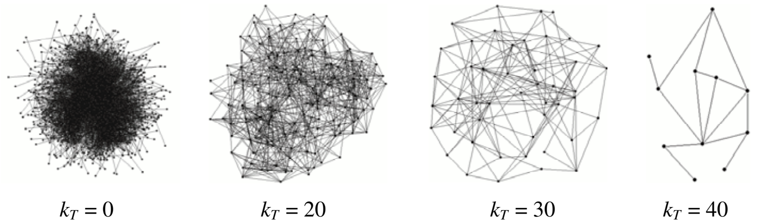


$k_T = 0$ $k_T = 1$ $k_T = 2$ $k_T = 3$

Fig. 1. Illustration of the DTR method

nodes and links constitute a hidden sub-network (10 nodes and 12 undirected links).

In Figure 2 the method was applied to the OOEMN ($k_T = 0$), and the resulting REMNs are constituted for all $k_T = 1, 2, 3, \cdots$, and k_{max}. Here, the REMNs in the case of $k_T = 20, 30, 40$, respectively, were given.



$k_T = 0$ $k_T = 20$ $k_T = 30$ $k_T = 40$

Fig. 2. The stages of REMNs with k_T = 0, 20, 30, 40

3 Properties of Renormalization Email Networks

3.1 Data Source

The data source being available from [16] was applied to construct email networks at the application layer of the Internet. The data source consists of 1133 nodes and 10903 directed links [17].

After analyzing this data source, the result shows that there is one looping back link, which is a start point (*id*=1133) of a link as well as an end point (*id*=1133) of the link, and that there are 5451 links, which overlap backward the others. In this paper, an undirected and unweighted OOEMN with 1133 nodes and 5451 links by eliminating looping and overlapping back links was firstly obtained, and then the DTR algorithm was applied to it, and to obtain the REMNs with $k_T = 0, 1, 2, \cdots, k_{max} - 1$. In Figure 2 the REMNs with $k_T = 0$ (1133 nodes and 5451 links), $k_T = 20$ (145 nodes and 887 links), $k_T = 30$ (46 nodes and 151 links), and $k_T = 40$ (11 nodes and 15 links) were mapped with the pajek (*http://pajek.imfm.si/*).

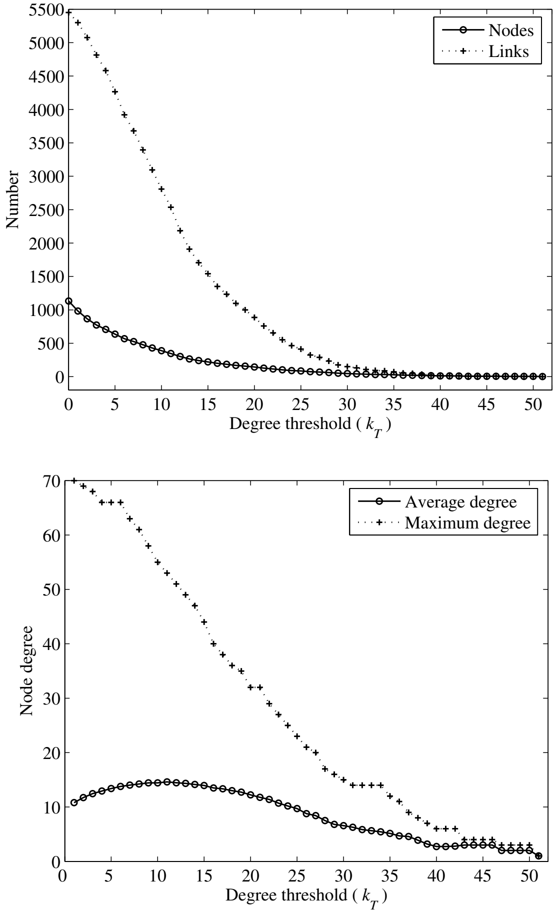


Fig. 3. *Top*: The number of nodes n and links m versus threshold k_T. *Bottom*: The average degree $\overline{k}$ and the maximum degree k_{max} versus threshold k_T.

3.2 Basic Properties

The number of nodes n, the number of links m, the average degree $\overline{k} = 2m/n$ and the maximum degree $k_{max} = max\{k_1, k_2, \cdots, k_n\}$ are the four basic topological properties of the REMNs. The first two properties are the approximate scale characteristic. REMNs with higher n and m are likely to experience more evolution time. The last two properties are the coarsest connectivity characteristic. REMNs with higher $\overline{k}$ and k_{max} are likely to be more robust.

Four basic properties of the REMNs as a function of threshold in the DTR method was plotted in Figure 3. Here, two new metrics was introduced as follows: $\rho_{node} = \Delta n/(nk_T)$, and $\rho_{link} = \Delta m/(mk_T)$, where Δn and Δm are respectively the number of eliminated nodes and links of the REMN with threshold k_T, so ρ_{node} and ρ_{link} response to the changing trend of eliminated nodes and links, respectively.

Figure 3 (top) shows that the number of email addresses and links reduce with the threshold increasing, and the rate of reduction is getting smaller and smaller. And furthermore the rate of links reducing is more than one of email addresses in the REMNs with low-threshold. The metric ρ_{node} in the four intervals is respectively 6.57%, 4.36%, 3.20%, and 2.48%, but the metric ρ_{link} is 4.85%, 4.20%, 3.24%, and 2.49%, respectively. In other words, the structure evolution speed of the email networks is getting increasingly great.

With the threshold increasing the average degree increases slowly from 10.8 to the first maximum 14.6 in the interval [0,11], and then decreases to 2.7 in [11,40], and again increases to the second maximum 3 and then reduces to 1 in [40,51] in Figure 3 (bottom). It shows that the average degree is self-similar in the REMNs with low and high thresholds, but they are self-dissimilar in the REMNs with middle thresholds.

It is concluded that the identities of the initially infected email addresses are more important in sparsely connected REMNs with low average degree than densely connected REMNs with high average degree [13], so the importance of them increases slightly firstly, and then decreases secondly and increases thirdly and decreases finally in the REMNs with the threshold increasing.

3.3 Power-Law Distribution

The degree distribution is the ratio of the number of nodes with degree k to the total number of nodes: $P(k) = n(k)/n$, where $n(k)$ is the number of nodes with degree k. The degree distribution contains more information about connectivity in a given REMN than four basic metrics of topological properties. For example, the average degree can be derived from the degree distribution.

Figure 4 shows PDFs of the degree distribution of the REMNs with different threshold k_T versus degree k in log-log scale. The results are shown that each of the PDF of the REMNs with $k_T = 1, 5, 10, 20, 30$ has a maximum value, where the corresponding degree $k_{inflexion}$ is 2, 6, 9, 10 and 6. The degree distribution of the REMNs with high-degree $k > k_{inflexion}$ is the power-law with a negative exponent: $P(k) \sim k^{-\gamma_1}$, and one of the REMNs with low-degree $k < k_{inflexion}$ is the power-law with a positive exponent: $P(k) \sim k^{-\gamma_2}$. Comparing the curves with different thresholds k_T, it is concluded that the degree distribution with high-degree in the REMNs follows power-law with a negative exponent and is self-similar. But one with low-degree does power-law with a positive exponent and is approximately self-similar.

In the power-law REMNs, a few email users have many links connected to other email users, while many email users do a few links. The power-law exponent depicts quantitatively the twisted phenomena of the REMNs topologies, and affects email viruses propagation. An email virus initially propagates faster on power-law REMNs with large exponent, and then dose subsequently slower than

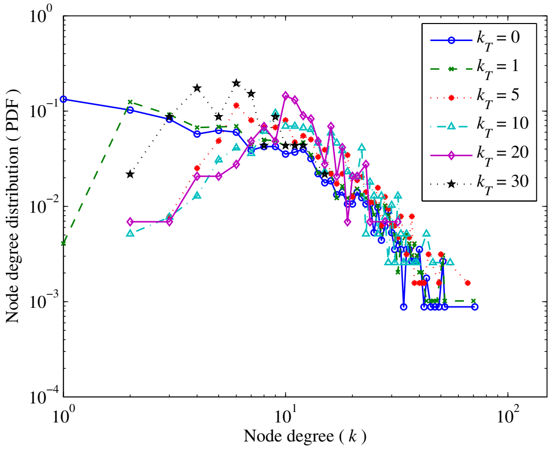


Fig. 4. Log-log plots of the PDFs of REMNs with different threshold k_T versus degree k

on those with small exponent [13]. As a result, the behavior of email viruses propagating on REMNs with similar exponents is self-similar.

3.4 Mixing Characteristic

The correlation of the connection probability and the degree from the power-law distribution cannot be obtained, while it can do from the *joint degree distribution* (JDD), which is the ratio of the number of links connecting nodes with degree k' and k to the total number of links: $P(k', k) \sim m(k', k)/m/2$. The disadvantage of the JDD is that it is intuitive, and the network topologies are not differentiated with the metric parameters. The *average neighbor connectivity* (ANC) of nodes can depict the correlation of networks [18]: $k_{nn}(k) = \sum_{k'=1,2,\cdots,k_{max}} k'P(k' \mid k)$, where the conditional probability $P(k' \mid k)$ is the probability that a given node with degree k' is connected to a node with degree k, and the following equation is shown: $P(k' \mid k) = \overline{k}/kP(k', k)/P(k)$. It can be used to measure the mixing characteristic of networks, namely the connected tendency of nodes with high-degree [19].

Figure 5 (Top) shows the normalized value $k_{nn}(k)/(n-1)$ of the ANC of the REMNs with different threshold k_T versus degree k in log-log scale. The ANC increases with increasing k_T. One of the REMNs with low-threshold k_T increases slightly with increasing k, where the high-degree email addresses are inclined to connect with other similar email addresses. One of the REMNs with large-threshold k_T fluctuates up and down for different degree k, where the connection tendency of high-degree email addresses was vivid. The standard deviation of the normalized value of the ANC of the REMNs with $k_T = 0$, 1, 5, 10, 20, 30 is 0.0015, 0.0018, 0.0025, 0.0057, 0.0124 and 0.0131, respectively.

Figure 5 (bottom) shows that the average value and the maximum value of ANC increase slightly with increasing k_T, and there have the jumping change

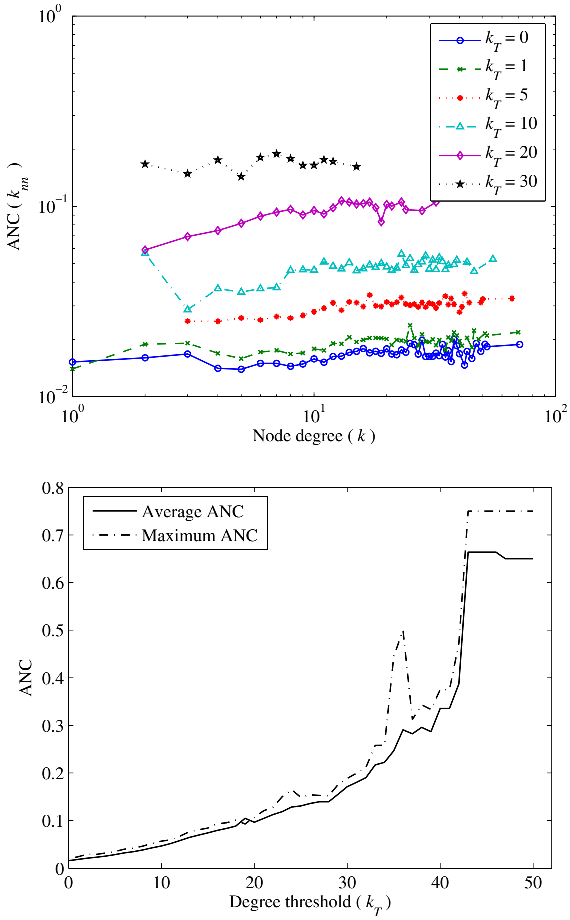


Fig. 5. *Top*: Log-log plots of the ANC $k_{nn}(k)$ versus degree k. *Bottom*: The average and maximum values of ANC versus threshold k_T.

when k_T reaches some value. It is shown that the ANC of the REMNs with low-threshold k_T is self-similar.

The other metrics method of mixing characteristic of networks is assortativity coefficient liking the following equation [6]: $r = (m^{-1}\sum_i j_i k_i - (m^{-1}\sum_i (j_i + k_i)/2)^2)/(m^{-1}\sum_i (j_i^2 + k_i^2)/2 - (m^{-1}\sum_i (j_i + k_i)/2)^2)$, where j_i and k_i is respectively the degree of two end points of the i_{th} link of the network, and m is the total links of the network. Figure 6 displays assortativity coefficients of the REMNs versus threshold. The result shows that the REMNs with $k_T \leq 29$ are assortative ($0 \leq r < 1$), and assortativity coefficient r changes little, so the REMN is self-similar. The REMNs with $k_T > 29$ are disassortative ($-1 \leq r < 0$), and assortativity coefficient r has big fluctuation, so it isn't self-similar. Research shows that the social networks are assortative, and the technological networks

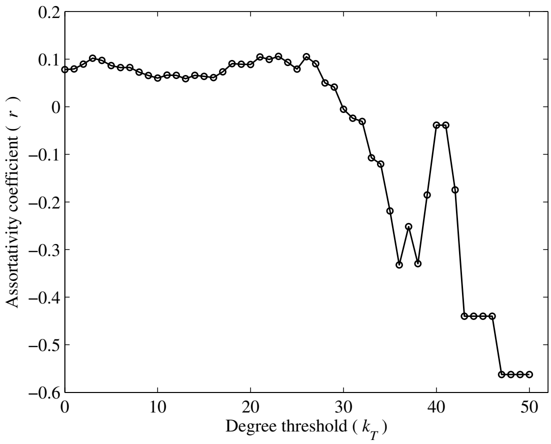


Fig. 6. The assortativity coefficients r versus threshold k_T

are disassortative [3]. And in consequence, the REMNs with low-threshold k_T are known as the social networks and the REMNs with high-threshold k_T are known as the technological networks.

The networks with high k_{nn} or small r have an excess radial links connecting email users of dissimilar degrees, and email viruses spread faster in these topologies [20]. The results show that the spread of the REMNs with low-threshold is self-similar.

3.5 Clustering Characteristic

The mixing characteristic of REMNs only tells us information about ANC and the degrees of neighbors for the nodes, but clustering characteristic does tell us how these neighbors interconnect. There have three main parameters of clustering characteristic: *local clustering coefficient* (LCC), *average clustering coefficient* (ACC), and *global clustering coefficient* (GCC). The LCC is the ratio of this number to the maximum possible such links: $C(k) = 2\overline{m_{nn}(k)}/k/(k-1)$, where $\overline{m_{nn}(k)}$ is the average number of links between the neighbors of nodes with degree k. The ACC is the average value of the LCC [21]: $C_{avg} = \sum_k P(k)C(k)$. The GCC is as follows: C = (3×number of triangles in the network)/(number of connected triples of vertices) [3].

LCCs of the REMNs versus degree k, and clique number and clustering coefficients versus threshold k_T were plotted in Figure 7.

Figure 7 (Top) shows that the LCCs can be approximated by a power-law distribution of degrees for all thresholds. The LCCs of the REMNs with different threshold are self-similar. The clique number of the REMNs is 3, 2 and 0, and one of the most REMNs is similar in Figure 7 (middle). The maximum values of the LCC are much larger than the ACC and the GCC in Figure 7 (bottom). Although the ACC and the GCC reflect clustering characteristic, there is difference between them. The value range of the ACC and GCC are 0.16 $\sim$ 0.35 and

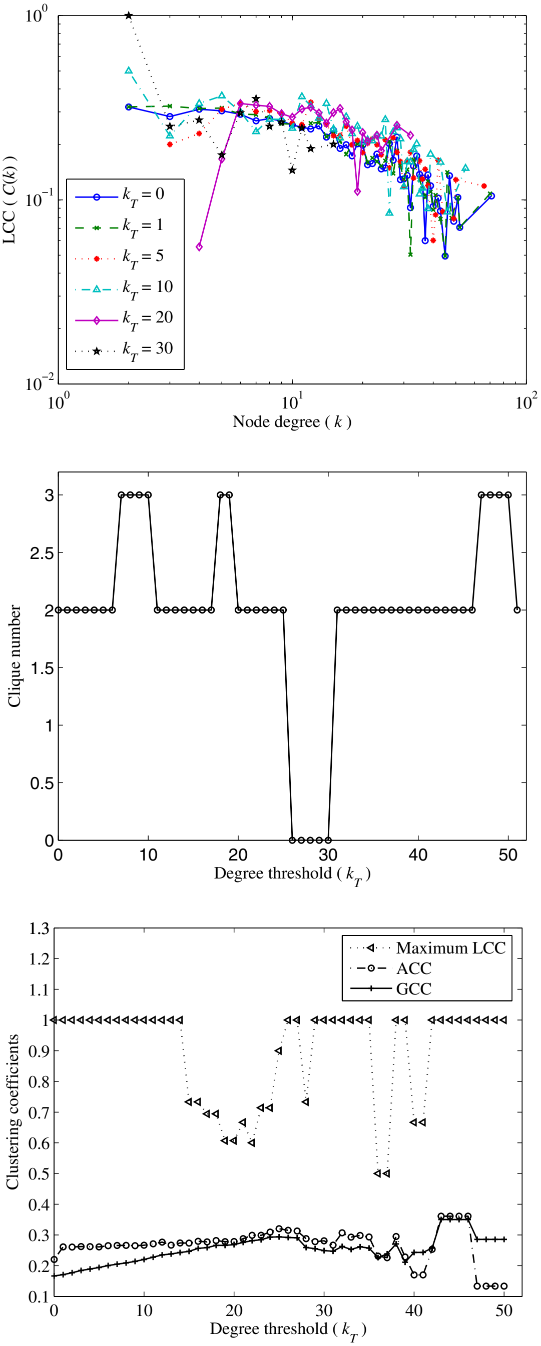


Fig. 7. *Top*: Log-log plots of the LCCs $C(k)$ versus degree k. *Middle*: The clique number versus threshold k_T. *Bottom*: The maximum value of LCC, ACCs C_{avg} and the GCCs C versus threshold k_T.

$0.13 \sim 0.36$, respectively. They with $k_T \leq 25$ increase with increasing threshold and those with $k_T > 25$ are frequent volatility.

The higher the clustering of an email address, the more interconnected are its neighbors, thus increasing the path diversity locally around the email address. It can been concluded that email virus outbreak spread faster in high-clustered REMNs according to the results in [22], and the spread behavior of email virus is similar in the REMNs.

3.6 Rich-Club Phenomena

The above properties cannot contain the connectivity of rich email addresses with high-degree, but the *rich-club connectivity* (RCC) can do. The RCC $\phi(r/n)$ is the ratio of number of links E_r in the sub-network induced by the largest-degree email addresses r to the maximum possible links $r(r-1)/2$: $\phi(r/n) = 2E_r/r/(r-1)$. Examples of these observations were provided in Figure 8, which show that the RCC versus normalized node rank r/n.

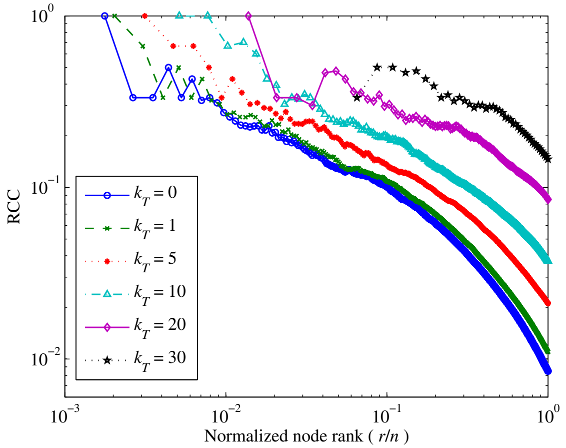


Fig. 8. Log-log plots of the RCCs $\phi(r/n)$ versus normalized node rank r/n

The results show the REMNs with the RCC more than 0 have rich-club phenomena and the RCC decreases approximately with node rank increasing, and increases with threshold increasing. The RCC exhibits power-laws for the REMNs in the area of medium and large normalized node rank r/n, and the exponents are similar.

In Figure 9 the ratio of rich email addresses of the REMNs to total email addresses versus degree threshold was plotted. Besides the REMNs with threshold $k_T = 26$, 27, 28, 29 and 30, the RCCs $\phi(r/n) = 1$, namely the rich email addresses of the REMNs is a complete connected graph. The ratio of the REMNs with threshold $k_T < 26$ is approximately similar. The ratio of the REMNs with

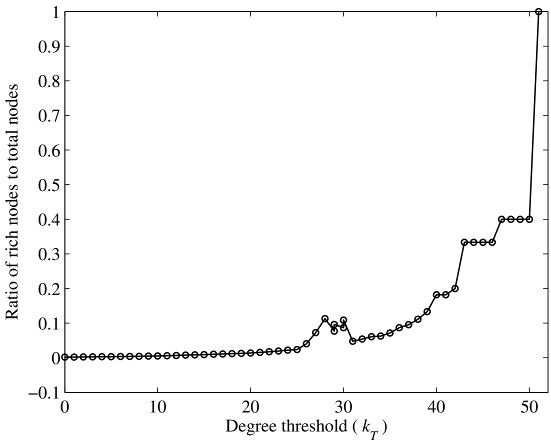


Fig. 9. The ratio of rich email addresses to total email addresses versus degree threshold k_T

threshold $k_T > 30$ appears in rapid increase stage, and it is not similar. There are two rich-clubs in the REMNs with threshold $k_T = 29$ and 30.

One can show that REMNs with the same JDDs have the same RCCs. The converse is not true, but one can fully describe all the JDDs having a given form of RCCs. Therefore, that the spread of the REMNs with low-ratio is self-similar.

4 Conclusions and Discussion

Various features of the REMNs based on a degree-thresholding renormalization method applied to the observed email network have been studied. Results show that the REMNs have various properties which differ from those in other many complex net-works as follows: positive and negative double exponential power-law distribution, assortative and disassortative hybrid mixing characteristic, and which is similar to those as follows: high-clustering coefficients and rich-club phenomena. It has been found that the properties of the REMNs with low-threshold k_T, where k_T is less than half of the maximum threshold of the REMNs, are self-similar, and those of the REMNs with high-threshold are self-dissimilar. Furthermore, for REMNs being self-similar, the spread behavior of email viruses in the REMNs is similar.

The self-similarity of statistical properties of the REMNs can be served as a general characteristic of email networks as well as giving a further insight into the understanding of evolutionary mechanism of the real email networks. Furthermore, it can used to validate existing models and also develop better topology models.

Acknowledgments. We have benefited from a discussion with Wenjun Xiao. This work was supported by the National Natural Science Foundation of China under Grant No.60773012, the China Postdoctoral Science Foundation under Grant No.20070420782, and the Natural Science Foundation of Hunan Province of China under Grant No. 07JJ6127.

References

1. Mahadevan, P., Krioukov, D., Huffaker, B., Dimitropoulos, X., Claffy, K., Vahdat, A.: Lessons from Three Views of the Internet Topology. Technical Report, Cooperative Association for Internet Data Analysis, CAIDA (2005)
2. Donnet, B., Friedman, T.: Internet Topology Discovery: a Survey. IEEE Communications Surveys and Tutorials 9(4), 56–69 (2007)
3. Newman, M.E.J.: The Structure and Function of Complex Networks. SIAM Review 45, 167–256 (2003)
4. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On Power-Law Relationship of the Internet Topology. ACM SIGCOMM Computer Communication Review 29(4), 251–262 (1999)
5. Albert, R., Jeong, H., Barabási, A.L.: Diameter of the World-Wide Web. Nature 401, 130–131 (1999)
6. Newman, M.E.J.: Assortative Mixing in Networks. Phys. Rev. Lett. 89, 208701 (2002)
7. Shi, Z., Mondragón, R.J.: The Rich-Club Phenomenon in the Internet. IEEE Communications Letters 8(3), 180–182 (2004)
8. Chaoming, S., Shlomo, H., Hernán, A.M.: Self-Similarity of Complex Networks. Nature 433, 392–395 (2005)
9. Kim, J.S., Goh, K.-I., Salvi, G., Oh, E., Kahng, B., Kim, D.: Fractality in Complex Networks: Critical and Supercritical Skeletons. Phys. Rev. E 75, 016110 (2007)
10. Doyle, J.C., Alderson, D.L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., Willinger, W.: The "Robust yet Fragile" Nature of the Internet. PNAS 102(41), 14497–14502 (2005)
11. Colizza, V., Flamini, A., Serano, M., Vespignani, A.: Detecting Rich-Club Ordering in Complex Networks. Nat. Phys. 2, 110–115 (2006)
12. Ebel, H., Mielsch, L., Bornholdt, S.: Scale-Free Topology of E-mail Networks. Phys. Rev. E 66, 035103 (2002)
13. Zou, C.C., Towsley, D., Gong, W.: Email Virus Propagation Modeling and Analysis. Technical Report TR-CSE-03-04, University of Massachusetts (2003)
14. Itzkovitz, S., Levitt, R., Kashtan, N., Milo, R., Itzkovitz, M., Alon, U.: Coarse-Graining and Self-Dissimilarity of Complex Networks. Phys. Rev. E 71, 016127 (2005)
15. Serrano, M.A., Krioukov, D., Boguñá, M.: Self-Similarity of Complex Networks and Hidden Metric Spaces. Phys. Rev. Lett. 100, 078701 (2008)
16. Network Data Sets, `http://deim.urv.cat/~aarenas/data/welcome.htm`
17. Guimerà, R., Danon, L., Diaz-Guilera, A., Giralt, F., Arenas, A.: Self-Similar Community Structure in a Network of Human Interactions. Phys. Rev. E 68, 065103 (2003)
18. Pastor-Satorras, R., Vázquez, A., Vespignani, A.: Dynamical and Correlation Properties of the Internet. Phys. Rev. Lett. 87, 258701 (2001)
19. Vázquez, A.V.: Degree Correlations and Clustering Hierarchy in Networks: Measures, Origin and Consequences. Ph.D. dissertation, Scuola Internazionale Superiore di Studi Avanzati International School for Advanced Studies (2002)

20. Newman, M.E.J., Forrest, S., Balthrop, J.: Email Networks and the Spread of Computer Viruses. Phys. Rev. E 66, 035101 (2002)
21. Dorogovtsev, S.N.: Clustering of Correlated Networks. Phys. Rev. E 69, 027104 (2004)
22. Newman, M.E.J.: Properties of Highly Clustered Networks. Phys. Rev. E 68, 026121 (2003)