Applied Mathematics & Information Sciences
*An International Journal*

http://dx.doi.org/10.12785/amis/090310

# A Network Classification Method by using Community Structure

*Xue-Guang Wang*[1,2,*]

[1] Department of Computer Science, East China University of Political Science and Law, 201620, Shanghai, China
[2] Department of Computer Science, University College London (UCL), WC1E 6BT, London, UK

Received: 11 Jul. 2014, Revised: 12 Oct. 2014, Accepted: 13 Oct. 2014
Published online: 1 May 2015

**Abstract:** The social networks are human-centered relationship networks and the research in this field will impact on people's daily life. The development of the internet and the emergence of the social network sites provide a perfect platform for our research. This paper builds the social networks by large-scale datasets. We confirm and observe that they have several properties and features. We classify them according to the role-to-role connectivity profiles and finds that they are subject to stringy-periphery class.

**Keywords:** Social Networks, Network classification, Network Structure, Network Evolution

## 1 Introduction

Social network is a complicated structure composed of the nodes and the links, where the nodes represent people and the links represent the relationship between them. In our daily life, there are many examples of social networks, such as the E-mail network, co-authorship network and so on. With the development of the Internet and the emergence of the social network sites, we can obtain the datasets of the large scale social networks from the telecom operators or the Internet. Based on the different functions, the social networks could be simply classified into friendship networks (like Facebook and MySpace), media sharing networks (like Youtube and Flickr), blog networks (like Livejournal and Twitter), IM networks (like MSN and QQ), and BBS. All of them own tens of millions registered users, making a perfect platform for our research on the structure and the nature of social networks.

In 1960s, Stanley Milgram, the social psychologist at Harvard University put forward the six degree of separation theory. It means that the average distance of any pair of people in the world is 6 [1], which is seen as theoretical basis of social networks. With the great improvement of the computing capacity, the development of the Internet and the interpenetration of different disciplines, we can handle and compare the data of large-scale social networks to find out their common structural features. Since Watts and Storagatz raised the small-world network model in 1998 [2], Barabsi and Albert raised the BA scale-free network model in 1999 [3], the research of complex networks has obtained great achievements, and attracted the attention of the researchers from all the subjects like Mathematics, Statistics, Life Science, Information Science, Social Science and Economics. Analyzing the social networks using the theory of complex networks focuses mainly on the correlations among the nodes and their topology structures, which are also the foundation for figuring out the features and functions of social networks. In terms of computer science, a lot of valuable papers have been published in VLDB, KDD, WWW, IMC, ICDM, WI, and other international conferences. And there are also some symposiums in this field like WOSN (ACM SIGCOMM Workshop on Online Social Networks) and SNS (ACM EuroSys Workshop on Social Network Systems).

The researches of complex networks mainly consider its static and dynamic features. The former contains the topology analysis [4,5], community mining[6,7], discovering key nodes [8,9] and so on. As time goes on, the number of the nodes in the network will increase and there will be new links between the nodes, which makes social networks the dynamic features. The research in this aspect mainly involves in the formation and evolution of social networks [10,11]. Another hot spot is the interaction between the topology and dynamics of social

* Corresponding author e-mail: wangxueguang@ecupl.edu.cn

networks, one example is the information spreading in the social networks [12, 13].

The contributions of our paper are as follows:

– The feature and structure of social networks. We do a full research on the datasets and confirm their degree power-law distribution, densification law, shrinking diameter properties. We find the scale of the communities in these networks obeys the power-law distribution and the medium-scale communities are of star structure. Apart from these, we also find that there is a core structure with high-connectivity and small-diameter in the social networks.
– The classification of social networks. Currently, according to the classification based on the assortativity coefficient, it is widely thought that social networks are assortative. However, we find that some real social networks are disassortative. We have applied the role-to-role connectivity profiles to classify social networks, which will help us to understand the internal structure of social networks.

The rest of this paper is organized as follows. Section 2 describes the dataset applied in this paper and the related important conceptions. Section 3 confirms the features of social networks and we present the structure characteristics in section 4. Section 5 introduces the research on the classification of social networks based on the role-to-role connectivity profiles and we get our conclusions in the last section.

## 2 Datasets and related conceptions

### 2.1 Datasets

The datasets applied in this paper are from DBLP computer science bibliography and Facebook social network site.

#### 2.1.1 DBLP dataset

DBLP is a tool for the researchers when they need to search the computer science papers published in international journals and conferences, and it enjoys high reputation in academic circles. Presently, more than 1.4 million journals have been collected by DBLP and are available in the XML format. We select the papers published in 80 important journals of computer science from 1986 to 2006 as basic data and treat the authors as nodes, an undirected edge will be added between two nodes if the two corresponding authors are in the same paper. In this way we build a DBLP co-authorship network that consists of 83943 nodes and 190342 edges.

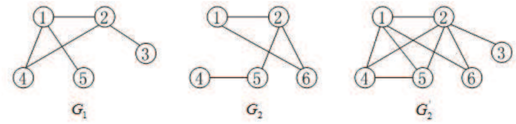


**Fig. 1:** An example of temporal and cumulative graphs

#### 2.1.2 Facebook dataset

Facebook is the most famous social network site owning more than 500 million registered users. The data applied in this paper is from [14], which collected the friendship relation in New Orleans. We build a Facebook network with 60567 nodes and 583766 edges using the data collected from 2007 to 2008.

In addition, the Enron dataset applied in this paper is collected by SRI and made public by FERC (U.S. Federal Energy Regulatory Commission); Cohen from CMU builds the website for it with the purpose of making it available for all researchers. It contains 500 thousand E-mails. The address is represented as a node and an undirected edge will be added if there is an E-mail between two addresses, the network built in this way consists of 36692 nodes and 367662 edges. Youtube dataset is from [5], and is selected from Jan 1, 2007 to Jan 15, 2007. The YouTube network consists of 35648 nodes and 261191 edges.

### 2.2 Related Conceptions

Generally speaking, for the social networks composed of the nodes and the edges, all the corresponding data are obtained during the whole collecting duration; we may lose the time information when the nodes interact with each other. So we give a dynamic model, which is a discrete time series of temporal graphs. The graph is formed by the nodes and edges obtained at specific time. We give its formal definition here.

Let $G_t = (V_t, E_t)$ denote the temporal graph at time $t$ ,$V_t$ represents a set of all the nodes at time $t$, and $E_t$ represents a set of all the edges appearing at $t$ . The temporal graph sequence $G^R$ consists of $G_1, G_2, \cdots, G_T$ ; if $\acute{G_i} = (\acute{,}\acute{)}$ such that $\acute{V_i} = \cup_{k=1}^{i} V_k$ and $\acute{E_i} = \cup_{k=1}^{i} E_k$, then $\acute{G_1}, \acute{G_2}, \cdots, \acute{G_T}$ is called cumulative graphsequence $G^A$. Let $G = \acute{G_T}$ denote final graph. (Anexample shown in Fig. 1)

The other related papers have given a lot of statistical properties of the network structure, so we just introduce three basic conceptions here, APL (Average Path Length), clustering coefficient, degree distribution [15] and assortativity coefficient [16].

In a network with $N$ nodes, the distance $d_{ij}$ between node $i$ and node $j$ is defined as the number of edges in the shortest path between them. The APL of the network is defined as

$$L = \frac{1}{\frac{1}{2}N(N+1)} \sum_{i \geq j} d_{ij}$$

Assume that node $i$ has $k_i$ neighbor nodes, i.e., node $i$ has degree $k_i$ . The ratio of the existed edge number $e_i$ among all these $k_i$ nodes to the theoretical maximum edge number $k_i(k_i-1)/2$, is defined as the clustering coefficient of node $i$, that is, $C_i = \frac{2e_i}{k_i(k_i-1)}$. So, the clustering coefficient of the network $C = \frac{1}{N}\sum_i C_i$. The node degree distribution $P(k)$ is the probability that a node chosen uniformly at random has degree $k$.

Degree correlation denotes that the average degree distributions of the neighborhood nodes for specific node with a given degree, which can be characterized by assortativity coefficient,

$$r = \frac{M^{-1}\sum_i j_i k_i - [M^{-1}\sum_i \frac{1}{2}(j_i+k_i)]^2}{M^{-1}\sum_i \frac{1}{2}(j_i^2+k_i^2) - [M^{-1}\sum_i \frac{1}{2}(j_i+k_i)]^2}, -1 \leq r \leq 1$$

Where $j_i$, $k_i$ are the degrees of the nodes at the ends of the $i$ th edge with $i = 1, \cdots, M$ and $M$ is the total number of all edges in the network. If $r > 0$, the network is assortative; if $r < 0$ , the network is disassortative; for $r = 0$, there are no correlation between node degrees.

## 3 The feature of social networks

In the theory of complex networks, one of its research contents is to reveal the statistical features of network architecture and find how to measure them. People have found that the networks have the small-world, free-scale features, and have developed some statistical methods such as degree distribution, APL, clustering coefficient and so on [17]. In this paper, we discuss the features of DBLP and Facebook from the standpoint of complex networks. Free-scale distribution widely exists in Physics, Geosciences, Computer Science, Biology, Demography and Economics [18]. Its formal definition is $P(k) = ck^{-\alpha}$ where $k$ denotes network feature variable, $c$ is a constant. For the degree distribution of nodes in the network, $\alpha$ generally lies in the range of 2 to 3 [19].

Firstly, we discuss the degree distribution of DBLP and Facebook datasets. Here it is necessary to point out that these networks are undirected and the influence of the ”missing past” on network features is relatively minor [20], From Fig. 2, we can see that the degree distributions of DBLP and Facebook are obviously scale-free, the power exponent of the former is 2.732 and the latter is 2.251, their fitting precisions (R-Square) are 0.9967 and 0.9874 respectively. In addition, using CNM algorithm we partition the communities of DBLP and Facebook and calculate their size, we find the community scale also obeys power-law distribution as shown in Fig. 3.

Secondly, we discuss how the APL changes with time and we characterize the changes by effective diameter [20]. If the distances of at least 90% node pairs are smaller than $d$, the minimum of $d$ is defined as effective diameter. As shown in Fig. 4, the effective diameter becomes smaller as time going. For DBLP, the journals are selected as data source artificially, and many fields of computer science are at beginning stage before 1990, so the effective diameter changes greatly and declines gradually afterwards. For Facebook, the curve declines obviously at the very start. The shrinking of the effective diameter is related to the density of the nodes and edges and their increment rate. So we discuss the changes of nodes and edges as time going. From Fig. 5 we can see that the edge number $E$ to the node number $N$ is nonlinear growth, which means the increase of average degree in the network. Formally, it obeys scale-free distribution ($E \propto N^{\beta}$). In this way, we confirm the densification laws and shrinking diameters of the DBLP and Facebook, for the former $\beta = 1.285$, $R^2 = 0.9998$ and the latter $\beta = 1.626$, $R^2 = 0.9987$.

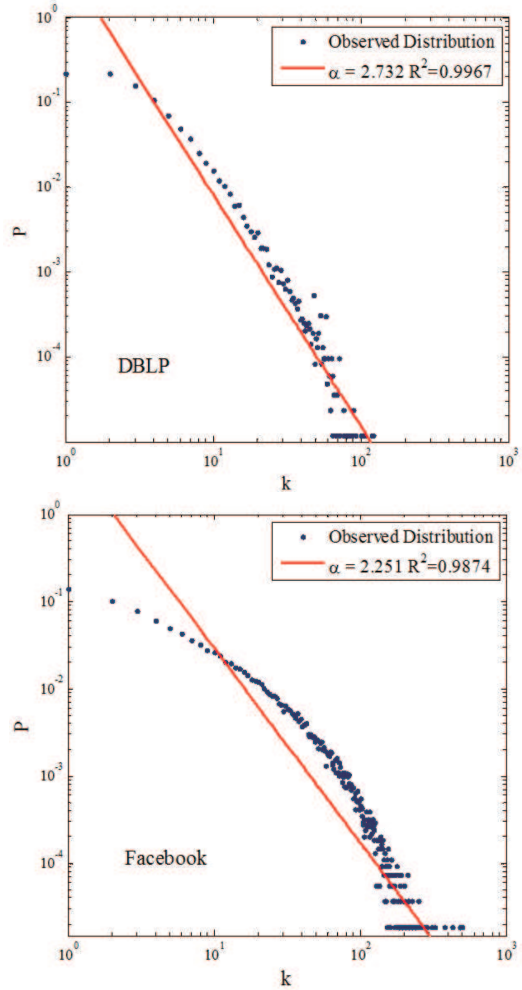


**Fig. 2:** Degree distribution

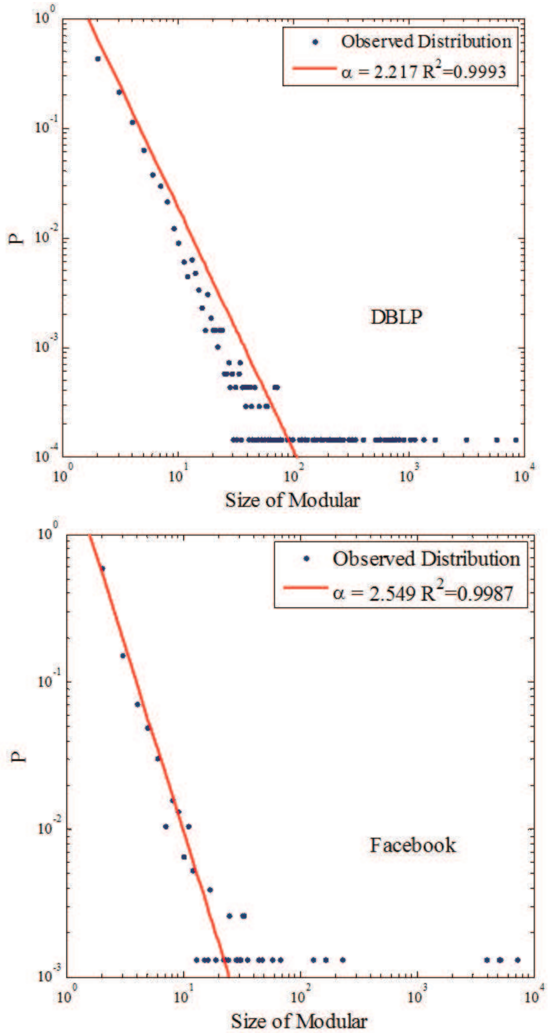


**Fig. 3:** Community size distribution

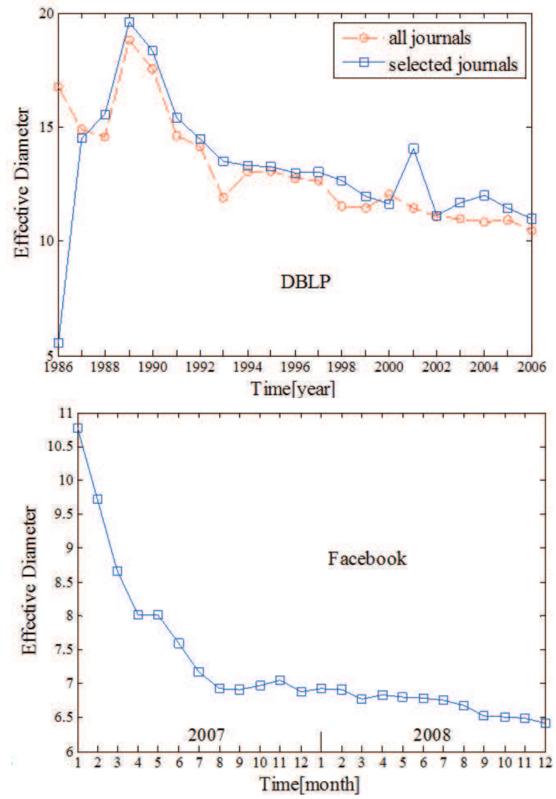


**Fig. 4:** Shrinking diameters of Networks

## 4 The structure of social networks

In section 3 we introduce the macro properties of social networks and in this section we will discuss the micro structure. Firstly, we explore how the community size of the network changes with time. Using CNM algorithm we partition the modules of the cumulative graphs of DBLP and Facebook and get the percentages of the communities with different size. The result is shown in Fig. 6. The bottom of the graph represents small-scale community (owns a few nodes), the middle part represents the medium-scale community (owns tens to hundreds of nodes), and the top represents the large-scale community (owns thousands of nodes). During the formation of the networks we find that the percentage of medium-scale community gradually decreases over time while the large-scale community increases and the small-scale community basically keeps unchanged. In addition, we find that the assortativity coefficient of medium-scale community is almost disassortative, which reveals that the nodes with the bigger degree are inclined to connect with those with less degree. Therefore, most of medium-scale communities in these two networks have obvious star structure.

Secondly, we discuss whether there are cores with small diameter and high connectivity in social networks. We remove the nodes from those of smaller degree to bigger degree gradually and record the changes of effective diameter and the ratio of edge to node. The result is shown in Fig. 7. Fig. 7(a) shows that the effective diameter of DBLP decreases by 28.94% and Facebook by 21.85%. On the contrary, their ratio of edge to node increases to 3.28 times and 1.88 times respectively, which is shown in Fig. 7(b). Based on these results we can draw the conclusion that there are core structures in both networks.

However, it cannot mean that the conductance of the core structure is the best. For a node set, the ratio of the number of the edges with one end inside the set and the other end outside the set to the total number of the edges in the set is defined as its conductance. But it is a hard problem for obtaining the subset with nodes and the weakest conductance. Using NCP method (Network Community Profile plot) [4], we consider the conductance of the communities with different in DBLP and

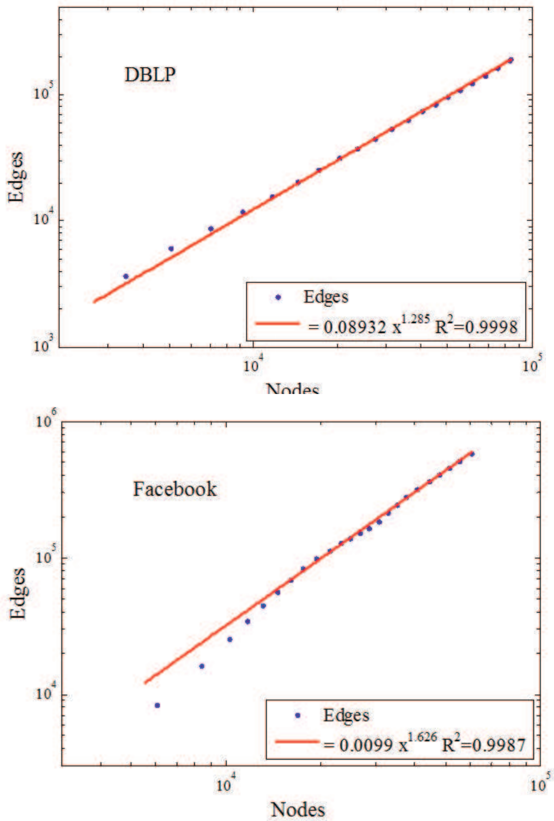


**Fig. 5:** Network densification

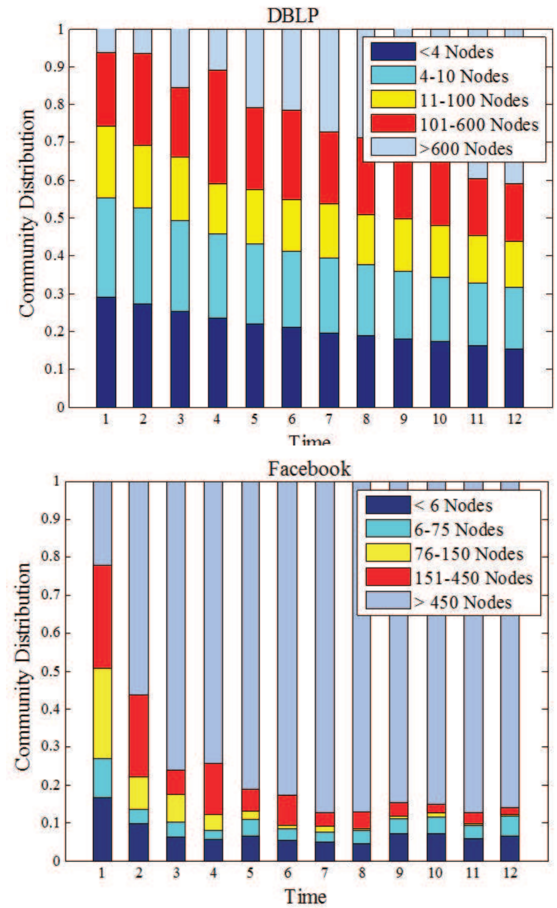


**Fig. 6:** The size of network community changes over time

Facebook. As shown in the Fig. 8, the node number of a community with good conductance ranges between 10 and 100. We can see that the conductance will become weak if the community size further increases. By seeking out the communities with weak conductance, we find that most of them lie in the boundary of the network and have few links to other communities.

## 5 The classification of social networks

Newman classified the networks in the real world into biological networks, technical networks, information networks and social networks according to their different features [21]. The related research shows that technical networks, biological networks and information networks are always disassortative, social networks are generally assortative. But we find that some on-line social networks like Pussokram in Sweden ($r = -0.048$) [22] and Cyworld in South Korea ($r = -0.13$) [23] are disassortative. We believe that the reason is that everyone in these on-line social networks is inclined to build relationship with those enjoying a high popularity. The assortativity coefficients of the datasets used by this paper respectively are 0.372 (DBLP), -0.033 (Youtube), 0.177 (Facebook) and -0.11 (Enron).

The degree correlation reflects the global feature of the network like degree distribution, APL and cluster coefficient, but only from these indicators we cannot approach its internal structure. It's a popular opinion that there are module structure (in other papers, it may be called community or group; in this paper, they share the same meaning) inside the social networks [24,25,26], in other words, all of its nodes in the module have more connections between them than with the rest of the network. Guimera et al. [27,28] assign the nodes different roles according to their connectivity pattern inside and outside the module, and then, they classify the networks with module structure into stringy-periphery class and multi-star class by applying RCP (role-to-role connectivity profiles). Using this method, they have analyzed and classified the protein-interaction networks, global and regional air transportation networks and the Internet at autonomous system level. However, we are not clear the internal characters of social networks, so the classification for social networks will help us understand their functional features and topological structures.

There are 3 steps to obtain RCP of the network.

–Partition the modules of the network

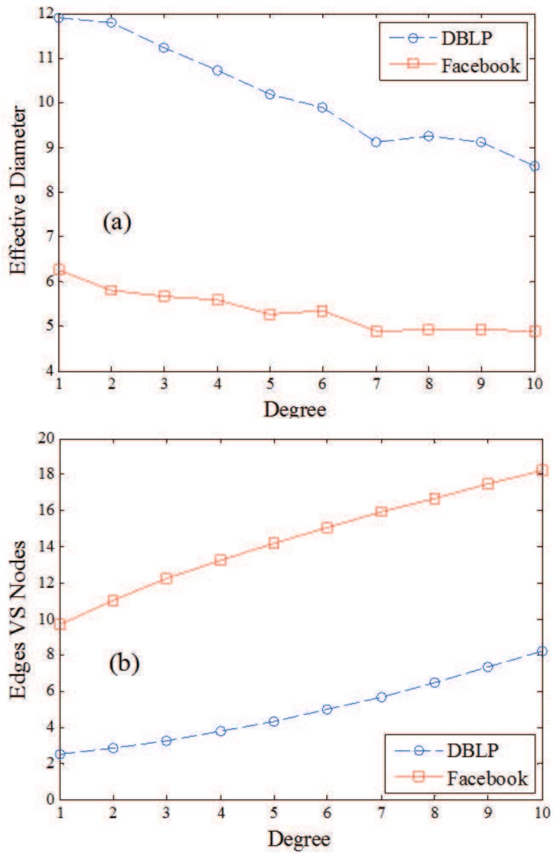


**Fig. 7:** The effective diameter and the ratio of edge to node change when nodes are removed from the network

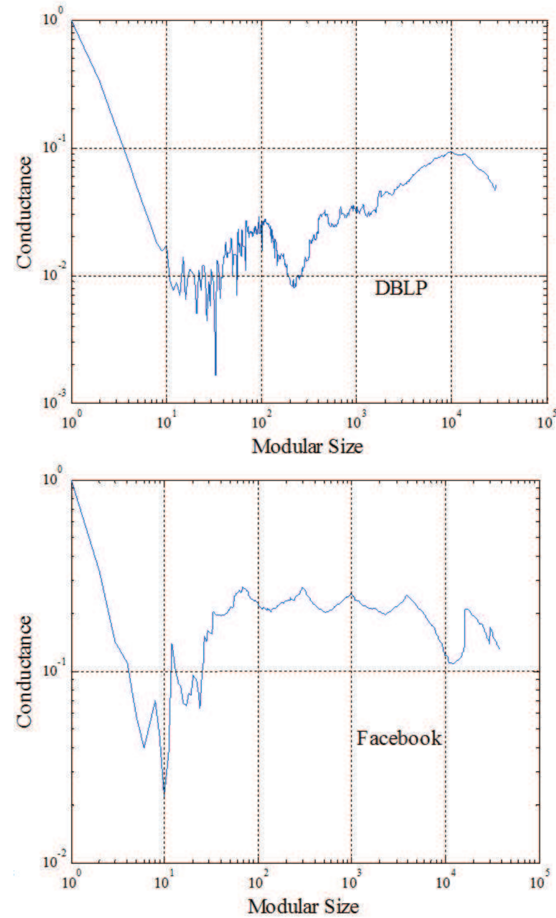


**Fig. 8:** NCP charts for DBLP and Facebook

Currently, there are many methods to partition the modules [29]. Guimera uses the simulated annealing (SA for short) to find the partition with the largest modularity [27]. But SA is sensitive to the initial parameters and the convergence rate is low, so we use the CNM algorithm [30] here raised by Clauset, Newman and Moore. CNM is a new greedy algorithm with linear time complexity $O(n\log^2 n)$ based on Fast Newman Algorithm and heap data structure. Using this algorithm, they analyzed the Amazon book-recommend network with 400000 nodes and 2000000 edges. Let $N_M$ denote the number of modules in the network.

–Assign each node in a module a role

Guimera et al. assign a role for each node in a module based on two parameters z and $P$. Parameter z for each node $i$

$$P_i = 1 - \sum_{S=1}^{N_M} (\frac{k_S^i}{K_i})^2$$

where $K_i = \sum_S k_S^i$ is the number of links of node $i$ to the other nodes in the same module $S$, $S_i$ is the module to which node $i$ belongs, and the averages $< \cdots >_{j \in S}$ are taken over all nodes in module $S$. The parameter $P$ for node $i$ is

$$P_i = 1 - \sum_{S=1}^{N_M} (\frac{k_S^i}{K_i})^2$$

where $K_i = \sum_S k_S^i$ is the degree of node $i$. Parameter z denotes the connectivity of the node to the other nodes in the same module while $P$ denotes the connectivity of the node to the other modules. The role is assigned according to the range of z and $P$ shown in Fig. 9 and Fig. 10. We have to point out that R4 and R7 rarely exist in real world.

–Calculate the Z-Score

Z-Score is the number of links between nodes with different roles. In this paper we use the random networks to get Z-Score, to make the result more precise, all the following parameters of the random networks we used here are the same as the real networks: the number of nodes and edges, degree of each node, number of modules and the connectivity between different modules. Here we apply the Markov-chain Monte Carlo switching algorithm [31] in the same module and between different modules of real networks to generate random network set $\mathfrak{A}$. The algorithm selects the random pairs (a, b) and (c, d)

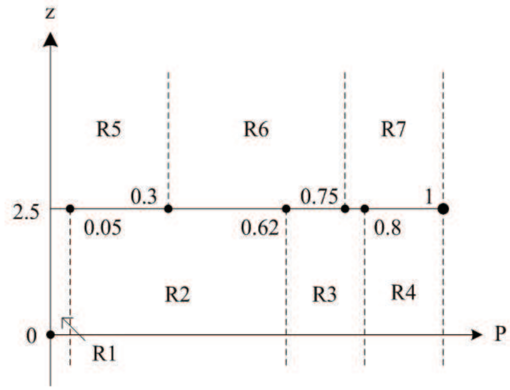


**Fig. 9:** Role partition in $z$ - $P$ parameters space

| | Role ID | Role Name | Notes |
|---|---|---|---|
| Non-hub nodes $z<2.5$ | R1 ($P\leq0.05$) | Ultra-peripheral | nodes with all their links within their module |
| | R2 ($0.05<P\leq0.62$) | Peripheral | nodes with most links within their module |
| | R3 ($0.62<P\leq0.80$) | Satellite connectors | nodes with many links to other modules |
| | R4 ($P>0.80$) | Kinless | nodes with links homogeneously distributed among all modules |
| Hub nodes $z\geq2.5$ | R5 ($P\leq0.30$) | Provincial hubs | hub nodes with the vast majority of links within their module |
| | R6 ($0.30<P\leq0.75$) | Connector hubs | hubs with many links to most of the other modules |
| | R7 ($P>0.75$) | Global hubs | hubs with links homogeneously distributed among all modules |

**Fig. 10:** The meanings of roles in Fig. 9

repeatedly and transform them into (a, d) and (c, b) for keeping their roles unchanged in the random network. We use $r_{ij}$ and $R_{ij}$ to denote the number of links between role $i$ and role $j$ in real networks and random networks respectively. The Z-Score for role $i$ and $j$,

$$z_{ij} = \frac{r_{ij} - < R_{ij} >_{\mathfrak{A}}}{\sqrt{< R_{ij}^2 >_{\mathfrak{A}} - < R_{ij} >_{\mathfrak{A}}^2}}$$

Therefore, we can get RCP chart of the network according to the Z-Scores from all role pairs.

In this way we get RCP charts of DBLP, Youtube, Facebook and Enron as displayed in Fig. 11.

The key to distinguish the types of different networks lies in their role-to-role connectivity patterns which mainly relates to the connection pairs R1-R1 and R6-R6. The air transportation network Fig. 11(a) and Internet at autonomous system level Fig. 11(b) are technical networks and their degree correlations are disassortative, but there are great differences if judged by their role-to-role connectivity. In Fig. 11(a), the Z-Score of R1-R1($Z^a_{R1-R1}$) is greater than the average level ($Z^a_{R1-R1} >> 0$), while the Z-Score of R1-R2($Z^a_{R1-R2}$) is smaller than the average level ($Z^a_{R1-R2} << 0$). This situation results in long node chains. However, the Z-Scores of R1-R1 and R1-R2 in Fig. 11(b) are opposite, $Z^b_{R1-R1} << 0$ while $Z^b_{R1-R2} >> 0$. It means that there are many star structures in the network. At the same time, there are $Z^a_{R6-R6} >> 0$ and $Z^b_{R6-R6} << 0$ for the Z-Scores of R6-R6 in Fig. 11(a) and 11(b). This is in correspondence with the real networks. The hierarchical structure of Internet leads the links from an autonomous system to the other systems are not too many. For the air transportation network, there may be many air lines from a transportation hub in a country or region to transportation hubs in other countries or regions. So we classify the former as multi-star structure, while the latter stringy-periphery structure.

There are similar characteristics like Fig. 11(a) when analyzing R1-R1, R1-R2 and R6-R6 in Fig. 11 (c), (d), (e) and (f), that is, $Z_{R1-R1} >> 0$, $Z_{R1-R2} << 0$, $Z_{R6-R6} > 0$ where $Z^c_{R6-R6} = 0.96$, $Z^e_{R6-R6} = 2.01$. There, we conclude that all these networks are subject to stringy-periphery class. (We have also made research on Flickr, Livejournal, Orkut and Wikipedia and drawn similar conclusions but we don't present here for space constraints) At the same time, we can also look inside the internal structure of social networks from RCP. For examples, in co-authorship network DBLP, $Z^c_{R5-R5} > 0$ and $Z^c_{R5-R6} > 0$ mean that the influential authors are more willing to cooperate with those at the same level with them. In Youtube, $Z^f_{R5-R5} < 0$ and $Z^f_{R5-R6} < 0$ indicate that the users prefer to upload videos rather than share them.

## 6 Conclusions and future works

In this paper we introduce our research on the feature, structure and classification of social networks. The social networks are human-centered relationship networks and the research in this field will impact on people's daily life. The development of the internet and the emergence of the social network sites provide a perfect platform for our research. How are the natures of social networks formed? Is there a theory model which can be able to explain the natures appearing during the interaction of the nodes? What influence exists between different network topologies and individual behavior? How to characterize and control the information spreading on social networks? These are the problems for us to further research and resolve.

## Acknowledgement

This work is supported by National Social Science Foundation of China (No.11BFX125).
The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

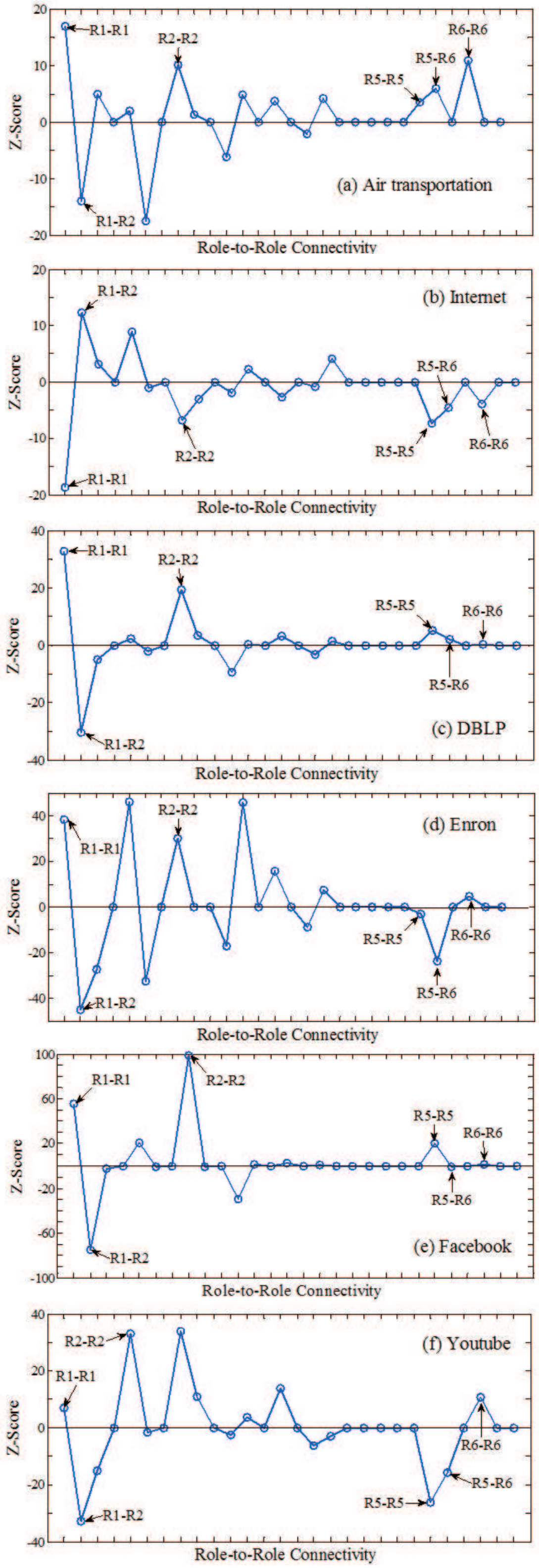


**Fig. 11:** (a), (b) from [28] respectively belong to stringy-periphery class and multi-star class. (c), (d), (e), (f) belong to stringy-periphery class

## References


[1] S. Milgram, The small world problem, Psychology today **2**, no.1, 60-67 (1967).
[2] D. Watts and S. Strogatz, Nature **393**, 440-442 (1998).
[3] A. Barabsi and R. Albert, Science **286**, 509 (1999).
[4] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney, Proceedings of the 17th international conference on World Wide Web, ACM, 695-704 (2008).
[5] A. Mislove, M. Marcon, K. Gummadi, P. Druschel, and B. Bhattacharjee, Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, ACM, 29-42 (2007).
[6] K. Wakita and T. Tsurumi, Proceedings of the 16th international conference on World Wide Web, ACM, 1275-1276 (2007).
[7] H. Kwak, Y. Choi, Y. Eom, H. Jeong, and S. Moon, Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference, ACM, 301-314 (2009).
[8] S. White and P. Smyth, Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data mining, ACM, 266-275 (2003).
[9] Y. Li, C. Lai, and C. Chen, Proceedings of the 11th International Conference on Electronic Commerce, ACM, 335-340 (2009).
[10] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins, Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 462-470 (2008).
[11] Y. Lin, Y. Chi, S. Zhu, H. Sundaram, and B. Tseng, ACM Transactions on Knowledge Discovery from Data (TKDD) **3**, 24-31 (2009).
[12] J. Tang, J. Sun, C. Wang, and Z. Yang, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 807-816 (2009).
[13] M. Kimura, K. Saito, and H. Motoda, ACM Transactions on Knowledge Discovery from Data (TKDD) **3**, 1-23 (2009).
[14] B. Viswanath, A. Mislove, M. Cha, and K. Gummadi, Proceedings of the 2nd ACM workshop on Online social networks, ACM, 37-42 (2009).
[15] X. F. Wang, X. Li, and G. R. Chen, Theory and application of complex networks, Beijing: Tsinghua University Press (2006).
[16] M. Newman, Physical Review Letters **89**, 208701 (2002).
[17] F. Rodrigues, G. Travieso, and V. Boas, Advances in Physics **56**, 167-242 (2007).
[18] H. B. Hu and L. Wang, PHYSICS **34**, 889-896 (2005).
[19] R. Albert and A. Barabsi, Reviews of modern physics **74**, 47-97 (2002).
[20] J. Leskovec, J. Kleinberg, and C. Faloutsos, Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, ACM, 177-187 (2005).
[21] M. Newman, SIAM Review **45**, 167 (2003).
[22] P. Holme, C. Edling, and F. Liljeros, Social Networks **26**, 155-174 (2004).
[23] Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong, Proceedings of the 16th international conference on World Wide Web, ACM, 835-844 (2007).
[24] S. Wasserman and K. Faust, Social network analysis, New York: Cambridge (1994).
[25] M. Girvan and M. Newman, Proceedings of the National Academy of Sciences of the United States of America **99**, 7821 (2002).

[26] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, Proceedings of the National Academy of Sciences of the United States of America **101**, 2658 (2004).
[27] R. Guimera and L. Amaral, Nature **433**, 895-900 (2005).
[28] R. Guimera, M. Sales-Pardo, and L. Amaral, Nature physics **3**, 63-69 (2006).
[29] B. Yang, D. Y. Liu, J. Liu, D. Jin, and H. B. Ma, Journal of Software **20**, 54-66 (2009).
[30] A. Clauset, M. Newman, and C. Moore, Physical Review E **70**, 66111 (2004).
[31] S. Maslov and K. Sneppen, Science **296**, 910 (2002).

**Xue-Guang Wang** is professor of East China University of Political Science and Law. Visiting Academic at University College London (UCL, UK). He received the Phd degree in Computer Networks at East China University of Science and Technology. His main research interests are: computer networks, information security and electronic evidence. He has published research articles in scientific world journal.