Physics 563 Term Paper

# Renormalization Group Analysis of Complex Networks: Applications in Scientific Networks

Po-Yao Chang*

May 5, 2010

Abstract:
Complex networks have relevance to many real systems in biological and social fields [1]. Applications of renormalization group (RG) theory technique give us the full pictures of the configurations in networks topologies (small-world or fractal types). Moreover using RG technique we can find the shortcuts structures hidden in the complex networks and even understand the growth of the networks. The mechanism of scientific credits diffusion and the ranking of scientists in academic fields is based on the structure of the publication networks and the citation system. Here I propose the RG technique to comprehend the scientific network structure and understand the evolution of scientific credits.

*chang153@uiuc.edu Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green St., Urbana, IL 61801, USA

# 1 Introduction

Complex networks exist everywhere. From real world human society to virtual online networks, or even the biological systems such as food web can be described as a complex network system. The standard way to study network is applying graph theory. By using nodes and edges connecting to each nodes, we can simply construct a complex networks. The inner structure in the networks are quite different type by type. There are two main types of networks, small-world one and fractal one. In small-world networks, the structure is more compact, which means almost every nots are connected by the edges. In other word, each two nods only need small numbers of edges to connect. We can describe the size of the networks by the average diameter $\bar{r}$, where $r$ is the shortest distance in each two nodes. The diameter of small-world networks have the relation, $\bar{r} \sim \ln N_0$. Where $N_0$ is the total number of nodes in the networks.

For fractal networks, the networks structures are not so compact and have the scale free property. The size of the networks have power-law relation to the number of nodes. That is $\bar{r} \sim N_0^{1/d_B}$, where $d_B$ is the fractal dimension [2] related to the length in the networks [3]. In recent studies of complex system, the renormalization group (RG) analysis provide a very powerful technique to classify the networks topologies by finding the fixed points in the RG flow [3] [4] [5] [6]. Moreover, by understanding the small-world-fractal transition, we can extract and find out the distribution of the shortcut structures in the networks which can help us having more clear view to the inside structure of the networks.

Another application of RG approach is to analyze the evolution of the networks. The idea is by inverting the RG procedure, then the networks can have a growth mechanism [7]. There are two different types of growth mechanism, hub-hub attraction or hub-hub repulsion which corresponding to small-world networks or fractal one.

Here I try to propose the above RG method to analyze the current scientific publications networks structures and study how the scientific credits spreading. Based on the enormous growth of internet data base and citation system in the academia, we can systematically rank the scientists through the publications networks by graph-based ranking algorithms [8]. By applying RG technique, we can have the deep look of the shortcuts structures hidden in the authors publication networks and try to find the evolution of this networks. My ultimate goal is after having clear view of the scientific authors networks, we can refine our ranking algorithms in more accurate way and have more reliable credits in the scientific field.

## 2 Renormalization group analysis of small-world-fractal phase transition

Renormalization group technique provides a very powerful tool to analyze the structure of networks. We can course-grain the networks in real space by box-covering technique [4]. We divide the networks with boxes and in each box the largest distance between each nods is smaller than $l_B$. After box-covering, boxes can be connected if there are nodes linked between different boxes. That means we can rescale the boxes into nodes and reconstruct a network (as shown in Figure 1.)

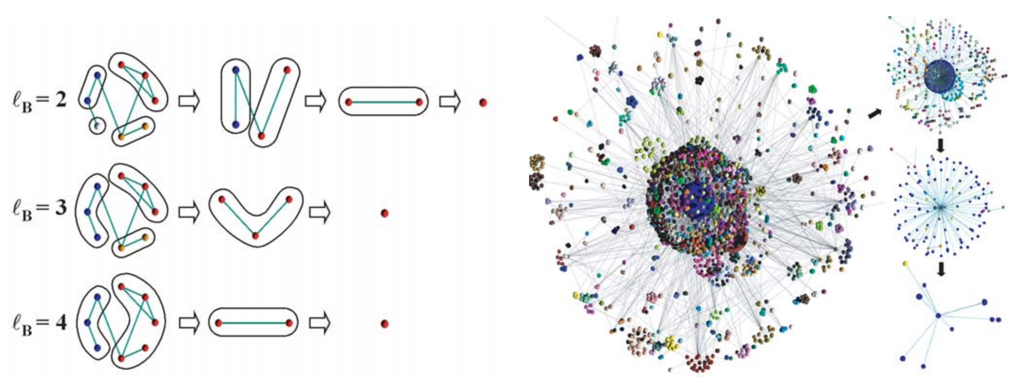


Figure 1: (from ref. [4])Left: Box-covering with different $l_B$. Right: The renormalization procedure of world-wide-web (WWW) with box size, $l_B = 3$.

## 2.1 RG flow and shortcuts structure in the network

The original network is $G_0$ and the box-covering RG transformation is $R_b$, with the box size $b$. It transfers the original network $G_0$ into a new network $G_b$. For a pure fractal network, the power law property of the network size $\bar{r}$ to the nods $N_0$ shows the self-similar trait in the network. The RG transformation transfers the original network to the same structure. In RG language, that is for a fractal network we have $R_b(G_0) = G_0$, where $G_0$ is the original network and also a fixed point of the RG flow.

To understand other types network structure, we can add the shortcuts to the fractal network. The new networks with shortcuts is $G'$ and the shortcuts structure can be described as adding the link in a distance $r$ between two nodes with the probability $P(r) = Ar^{-\alpha}$. In figure 2 (a), we can see the shortcuts structure in the networks will break the self-similar property in the fractal network. After RG transformation, it leads $R_b(G') \neq G'$. We can see for a fractal network, it is an unstable fixed point since the network goes away from it after the RG transformation. One interesting thing is there exists a trivial fix point in the RG flow. If all the nods are linked to each other (complete graph), it is a stable fixed point in the RF transformation.

We use the RG method to find the nontrivial fixed point hidden in the RG transformation. The shortcuts can be tuned by the probability $P(r) = Ar^{-\alpha}$ with the variable $\alpha$. Define a critical value $S = \alpha/d_B$, for $s < 2$ the RG flow goes to a trivial fixed point as a complete graph. If $S > 2$, the RG flow turns to a unstable fixed point as a fractal network. The nontrivial fixed point occurs at $S = 2$ with a specific shortcuts structure overlaying the fractal networks (as shown in figure 2 (c)). Moreover, we can use the above RG approach to analyze the shortcuts structure. It can be determined by the probability of the shortcuts $P(r) = Ar^{-\alpha}$ with different $\alpha$. We can extract the shortcuts structure in the networks between the fractal type and the small-world type and have a more clear view for information flow in the networks.

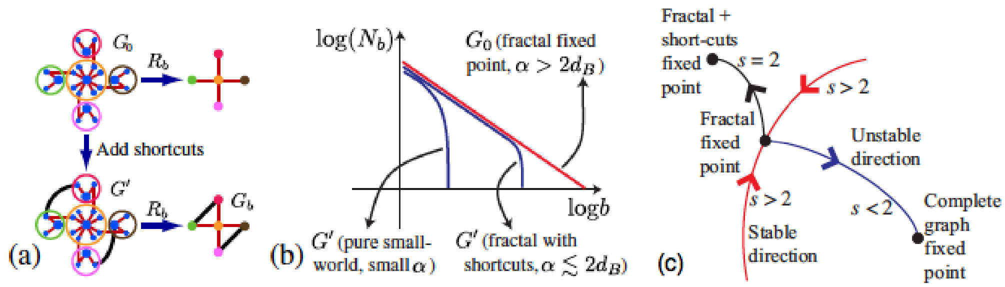


Figure 2: (from ref. [3])(a) Scheme of adding the shortcuts on the fractal networks with the box- covering RG transformation. (b) Sketch of the number of boxes versus the size of box with the probability of shortcuts structure $P(r) = Ar^{-\alpha}$. (c) The fixed points and the RG flows under RG transformation.

## 2.2 Evolution of the networks

How do the networks growth is strongly related to the structure inside the networks. The inverse procedure of the RG approach provide a systematical way to understanding the evolution of the networks [7]. The growth mechanism of the networks can be classified into two different types, hub-hub attraction type and hub-hub repulsion type. The hubs are the most connected nods in the networks. As figure 3 shown, the hubs are connected between each boxes for mode 1 (hub-hub attraction) and the edges of the boxes are linked to each others for mode 2 (hub-hub repulsion).

The growth mechanism of mode 2 is correlated to the fractal structure which are more robust networks and have the self-similar property in the complex systems. On contrary, the small-world networks are more like the complete graph with more connections between each nodes.

The growth mechanism of mode 1 with the strong hubs-hubs attraction provides the more compact structure of networks corresponded to the small-world networks. For the real complex neworks such as WWW, the growth mechanism is the combination of mode 1 and mode 2 as figure 3 (a) shown. The percentage of mode 1 and mode 2 in the network can be modified by the

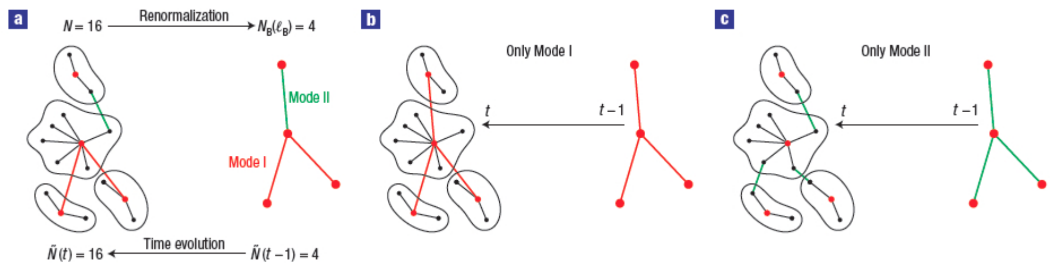


Figure 3: (from ref. [7]) (a) The inverse procedure of RG approach with two types of growth mechanism in the networks (b)Scheme of mode 1 growth mechanism (hub-hub attraction).(c) Scheme of mode 2 growth mechanism (hub-hub repulsion).

same method as overlaying the shortcuts structures on the fractal networks with adding the probability distributions $P(r) = Ar^{-\alpha}$ between two nods. Then we can apply the growth mechanism to trace the evolution of the networks.

In this section, we can see the RG technique provides a powerful analysis for the detail structure of the networks. By understanding the shortcuts structures inside the networks, we can use the inverse procedure of RG approach to predict the evolution of the networks. It is important to propose this RG method to analyze the publication networks in the academia and also necessary to provide some further applications for the improvements of our citation system.

# 3 Scientific citation network and ranking algothm

The citation system had been well used in the academia for a long time. Due to the revolution of technology, publications and citations can be recorded electronically. These data allow us to construct the network of citations between all the publications. The scientific credit can be spread due to this citation system and the ranking algorithm of the scientists becomes an important issue now. To construct an integrated network and a convinced ranking algorithm is necessary in the academia.

There are various of ranking algorithm such as PageRank [9] and CiteRank [10] based on the author-to-author network. Recently, Filippo Radicchi [8] proposed the science author rank algorithm (SARA) by simulating the diffusion of credits through the weighted author citation network (WACN).

In order to trace the credit flow for each scientists and use the diffusion algorithm for the ranking system, we need to construct weighted author citation network (WACN) from the paper citation network (PCN). The WACN [8] is the network based on the authors with weighted citation links. Suppose a paper $i$, written by coauthors $i_1, i_2, ..., i_n$ cites a paper $j$ with coauthors $j_1, j_2, ..., j_m$. Then the weighted link between two authors is $w = 1/(nm)$ as shown in figure 4.

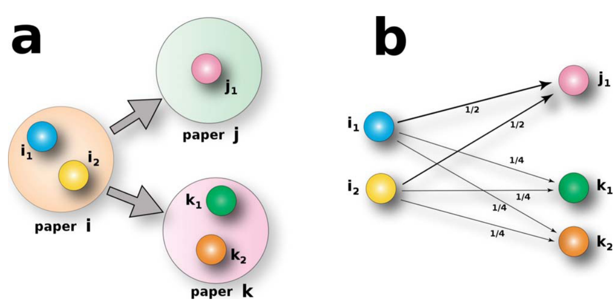


Figure 4: (from ref. [8]) Mapping PCN to WACN (a) The network based on the PCN.(b) The WACN is constructed by the authors with weighted links.

Based on the weighted links between each authors, their ranking algorithm (SARA) provides more reliable ranking results than the PageRank and CiteRank algorithm. The year that Nobel laureates received Nobel price is qualitatively related to the best ranking evaluated by SARA as shown in figure 5. In the next section, I'll briefly demonstrate how can we improve the ranking algorithm by RG approach.

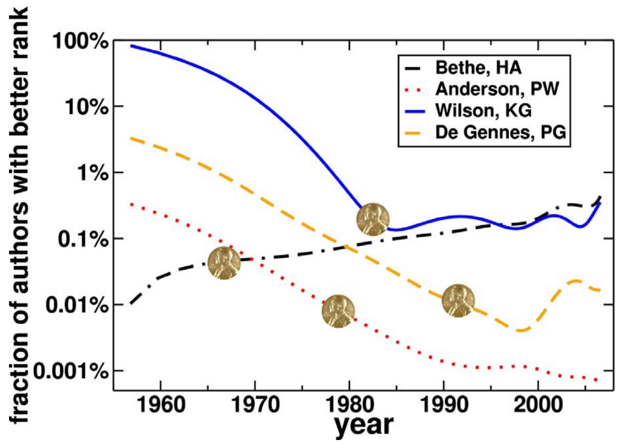


Figure 5: (from ref. [8]) SARA method showing the trend of better ranking of the Nobel laureates is highly related to the year they received the Nobel price.

# 4 Construct a evolving author network and the prediction of diffusion in scientific credits

The improved scientific network analysis procedure as shown in figure 6. First we use all the publications data to construct the author to author citation networks without the weighted links. Then we apply the RG technique to find out the shortcut structure and the evolution of the citation network. After the above work, we put the weighted links on the network with the arrows pointed to the cited authors. Finally we can use SARA to analyze the trend of scientists ranking with the growing network.

Before we do the refined procedure, we may ask what kind of authors network is it? Fractal or small-world? In figure 7 we can see the network based on the coauthor-ship has the fractal structure [11]. For this quick glance, we can naive think this subfields structure should provide the fractal property in the author citation network. Due to the human social interaction,

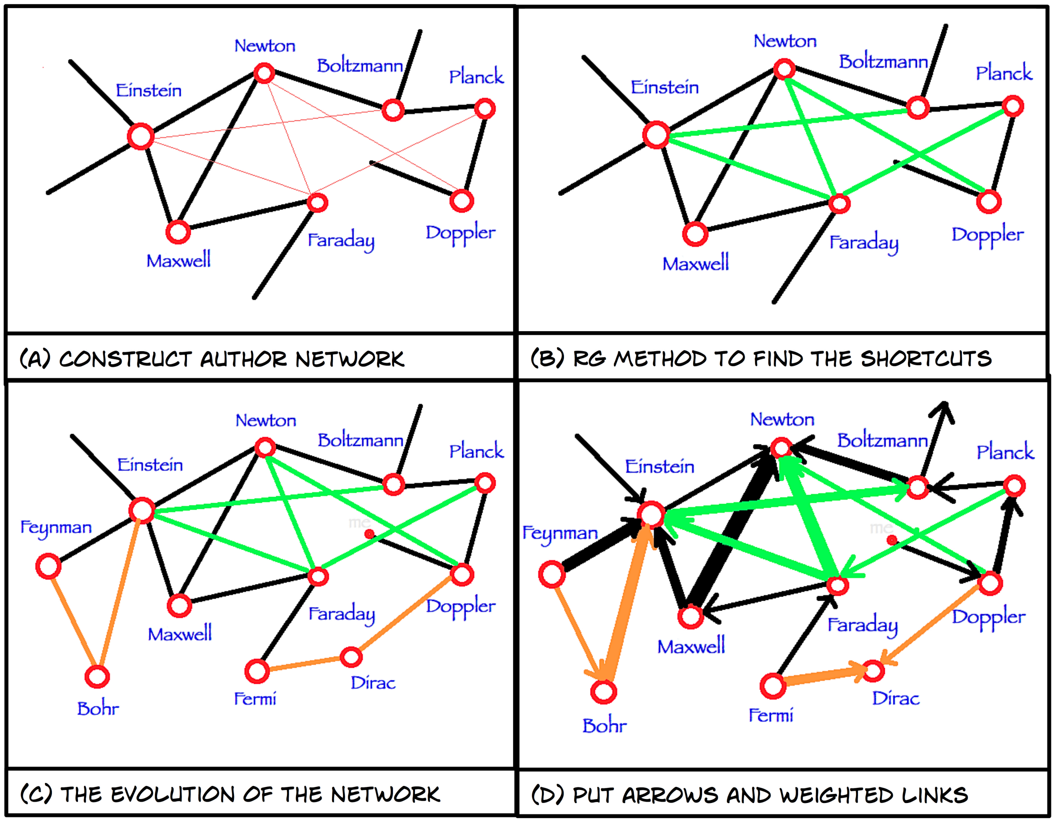


Figure 6: Scheme of the procedure to improve the scientific network ranking algorithm. (a) Using the publications data to construct the author-to-author citation network. Black lines are the links for the fractal network and light red lines are the hidden shortcuts structure.(b) Applying RG approach to find the shortcuts structure (green lines) overlaying the network. (c) Using inverse RG process to grow the network (orange lines). (d) After adding the arrows pointed to the cited authors and weighted links between the authors, we can apply the SARA method to find the trend of the spreading credits and the ranking of scientists among the evolving networks

the shortcuts will be overlaid on this fractal network and we can extract it by the RG approach.

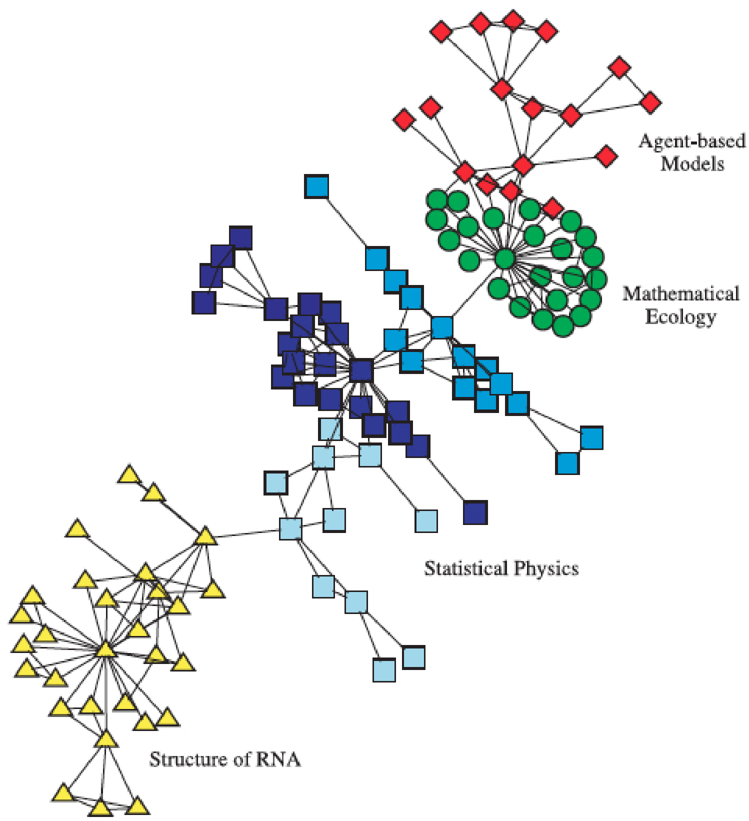


Figure 7: (from ref. [11]) The coauthor-ship network with different fields in a private institution.

Another interesting issue is the evolution of this network. Does it grow like hub-hub attraction type or hub-hub repulsion type? For hub-hub attraction type growth mechanism, the networks structure is more fragile. On the contrary, hub-hub repulsion type growth mechanism protect the fractal property with the robust network structure. In the author citation network, it more like a hybrid growth mechanism contained with both type. We can trace back the growth of the scientific network before and observe the decadence (fragile) and thrifty (robust) of the subfields networks. Does this phenomena relate to the defferent types of growth mechanism? The author network of high temperature superconductors publications may be the candidate model for hub-hub attraction growth mechanism with the property of highly correlation between two nodes [7]. We can have a more clear view after collecting the author citation data from all the high temperature superconductors publications and analyzing the evolution of the network.

# 5 Conclution

The renormalization group technique is a very powerful tool to study the complex networks. By understanding small-world-fractal phase transition [3], we can have a clear view of the shortcuts structure in the networks. Moreover, applying the inverse RG procedure [7], we can understand the evolution mechanism of the network. In this term paper, I try to propose the RG approach to find the improvements of the current ranking algorithm in the academia. Moreover, since for a fractal network, the growth mechanism is hub-hub repulsion type with a more robust structure, by understanding the scientific citation network we can improve our citation system and construct a more firm network such as fractal type for scientific society.

The further application of this RG method to the complex networks is quite huge. For example, the food web is a complex network as shown in figure 8 (a). If we add the predator-prey model into the food web network, the links between each species should be weighted and with arrows. The arrows are pointed to the predators and the weighted link depended on the percentage of preys source to the predators. This network is very similar to the weighted author citation network (WACN) (figure 8 (b)) and can be analyzed the dynamic evolution of the populations as the procedure proposed in this term paper.

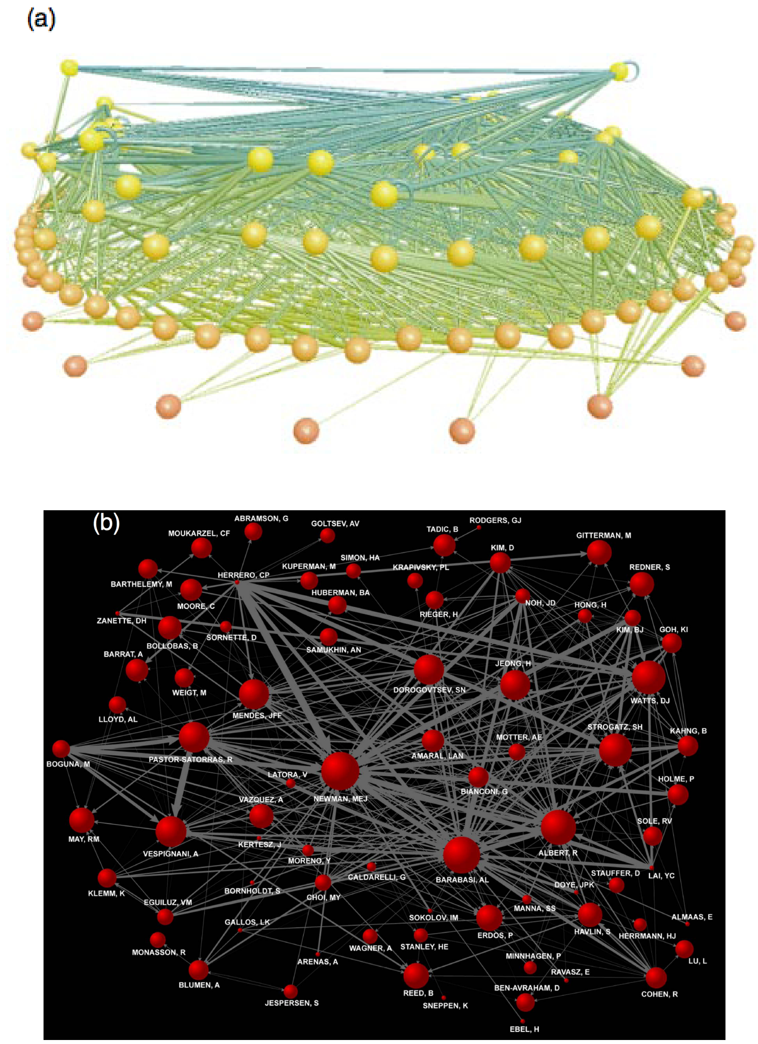


Figure 8: (a) Food web of Little Rock Lake, the nodes are the species. The nodes from top to the bottom are the trophic level from fishes to phytoplankton [1]. (b) Weighted author citation network (WACN) comes from the publications related to "complex networks" [8] .

## References


[1] S. H. Strogatz, Nature 410, 6825 (2001)

[2] Feder, J. Fractals (Plenum, New York, 1988).

[3] H. D. Rozenfeld, C. Song, and H. A. Makse, Phys. Rev. Lett. 104, 025701 (2010).

[4] C. Song, S. Havlin and H. A. Makse, Nature 433, 392 (2005).

[5] D. M. Auto, A A. Moreira, H. J. Herrmann, and J S. Andrade, Phys. Rev. E 78, 066112 (2008).

[6] M. E. J. Newman and D. J. Watts, Phys. Rev. A, 263, (1999)

[7] C. Song, S. Havlin and H. A. Makse, Nature Physics 2, 275 (2006).

[8] F. Radicchi, S. Fortunato, B. Markines, and A. Vespignani, Phys. Rev. E 80, 056103 (2009).

[9] S. Brin and L. Page, Comput. Netw. ISDN Syst. 30, 107 (1998).

[10] D. Walker, H. Xie, K. K. Yan, and S. Maslov, J. Stat. Mech.: Theory Exp. (2007), P06010.

[11] M. E. J. Newman, PNAS vol. 101 no. Suppl. 1 5200-5205, April 6, (2004)