Experimental Study on Relationship between Indices of Network Structure and Spectral Distribution of Graphs

Ryosuke Sawada*, Yusuke Sakumoto*, Chisa Takano†, and Masaki Aida*
*Graduate School of System Design Tokyo Metropolitan University, Hino-shi, Tokyo, 191–0065, Japan
†Graduate School of Information Sciences Hiroshima City University, Hiroshima, 731–3194, Japan

***Abstract*—Spectral graph theory uses the Laplacian matrix of a network to analyze its network structure. The Laplacian matrix is defined by node degrees and adjacency relationships between nodes. A previous study clarified that the eigenvalue distribution of the Laplacian matrix has a high similarity to the distribution of node degrees used in the Laplacian matrix. Degree distribution does not contain the information of adjacency relationships between nodes, so the previous study did not clarify the effect of the adjacency relationship information on the eigenvalue distribution of a Laplacian matrix. In this paper, for understanding such an effect, we investigate the relationship between the eigenvalues of the Laplacian matrix and the familiar indices (i.e., average path length and clustering coefficient), which are affected with adjacency relationship information.**

I. Introduction

In the actual world, there are a lot of types of networks including food webs, friendship networks, power networks, and world wide web. Such actual networks gather attention in various fields such as ecology, engineering, sociology, and physics, and many people have strived to investigate properties of their network structures [1]–[5]. In the investigation, the studies have used indices (e.g. average path length, clustering coefficient and many kinds of node centralities [6]–[9]) of network structures.

Spectral graph theory is a well-known approach to analyze network structures [10]. In this theory, network structures are represented algebraically with Laplacian matrix. Laplacian matrix is defined by using node degrees and adjacency relationships between nodes in a network. Its eigenvalues and eigenvectors play a significant role for analyzing network structures. For example, the second smallest eigenvalue of a Laplacian matrix represents the strength of the connection among nodes in a network [11]. The eigenvector associated with the second smallest eigenvalue is useful to extract a weak connection between a node pair in the network.

In [12], the authors clarified that the eigenvalue distribution of a Laplacian matrix has a high similarity to the degree distribution when a network is generated with well-known network models (i.e., Erdös-Rény model [13], Watts-Strogatz model [14], Newman-Watts model [15], and Barabási-Albert model [16]). Degree distribution does not contain the information of adjacency relationships between nodes, so the previous study did not clarify the effect of the adjacency relationship information on the eigenvalue distribution of a Laplacian matrix. The indices of network structures are affected with adjacency relationship information, and so would provide a connection between adjacency relationship information and the eigenvalue distribution of a Laplacian matrix.

In this paper, we experimentally investigate the relationship between eigenvalues of the Laplacian matrix and familiar indices (i.e., average path length and clustering coefficient) that represent the information of adjacency relationships between nodes. First, we generate networks with different adjacency relationships between nodes and identical degree distribution. To generate such networks, we make a network generation model, which is an extension of the Watt and Strogatz model (WS model). Then, we compare the eigenvalue distributions of the Laplacian matrices of the generated networks. In addition, we focus on the second smallest eigenvalue of the Laplacian matrix and investigate the relationship between the second smallest eigenvalue of the Laplacian matrix and the indices about adjacency relationships between nodes (i.e., average path length and clustering coefficient).

The rest of this paper is organized as follows. Section 2 outlines the definition and basic properties of the Laplacian matrix. In Section 3, we explain average path length and clustering coefficient. In addition, we make a network model to generate networks with different adjacency relationships between nodes and the identical degree distribution. Section 4 details numerical experiments conducted to investigate the relationship between the eigenvalues of the Laplacian matrix and the indices about adjacency relationships between nodes. Finally, in Section 5, we conclude this paper.

II. Laplacian Matrix

In this section, we outline the definition and basic properties of a Laplacian matrix. In addition, we explain that eigenvalues and eigenvectors of a Laplacian matrix are useful to investigate network structures.

A. Definition of the Laplacian Matrix

In network analysis, network structure is frequently expressed with a matrix. Let us consider an undirected graph $G = (V, E)$ where V is the set of nodes (i.e., $V =$

$\{1, \ldots, n\}$), and E is the set of links. We define the following $n \times n$ adjacency matrix $A = [A_{ij}]$ as

$$A_{ij} := \begin{cases} 1 & (i,j) \in E, \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

Next, we define the degree matrix D as

$$D := \mathrm{diag}(d_1, d_2, \ldots, d_n), \tag{2}$$

where d_i is the degree of a node i ($i = 1, 2, \ldots, n$).

According to [10], the Laplacian matrix of graph G is defined as

$$L := D - A. \tag{3}$$

Figure 1 shows an example of degree matrix D, adjacency matrix A, and Laplacian matrix L of a simple network ($n = 4$).

B. Basic Properties of the Laplacian Matrix

Since Laplacian matrix is a real symmetric matrix, Laplacian matrix has the following properties.

- L has n real eigenvalues λ_μ ($\mu = 1, \ldots, n$). λ_1 is always zero. λ_μ are arranged in non-decreasing order. Namely, $\lambda_1 = 0 \leq \lambda_2 \leq \cdots \leq \lambda_n$.
- The eigenvectors of L associated with eigenvalue λ_μ, $\boldsymbol{v}_\mu$ ($\mu = 1, \ldots, n$), can be chosen to be orthogonal to each other. Even if some eigenvectors are associated with the same eigenvalue, we can choose orthogonal eigenvectors.
- The normalized n orthogonal eigenvectors span the eigenbasis of the n-dimensional eigenspace of L.

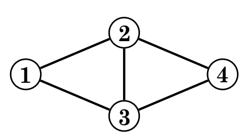
The other properties of a Laplacian matrix are as follows.

- The smallest eigenvalue of L is 0, the corresponding eigenvector is the constant one vector $\mathbf{1}$, that is $L\,\mathbf{1} = 0$.
- The multiplicity of the eigenvalue 0 of L is the number of connected components of G.
- The second smallest eigenvalue of L represents the strength of the connection among nodes in a network, and is called *algebraic connectivity* [11].

We describe the significance to analyze the eigenvalues and the eigenvectors of a Laplacian matrix. Since the Laplacian matrix is a real symmetric matrix, it can be diagonalized by the orthogonal matrix $P := (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ constructed by the orthogonal eigenvectors of L, as

$$P^{-1}LP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \Lambda, \tag{4}$$

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

Fig. 1. An example of the Laplacian matrix of a simple network ($n = 4$).

where Λ is a diagonal matrix, and its diagonal components are eigenvalues of L.

By using the Λ and P, we rebuild L as follows

$$P\Lambda P^{-1} = PP^{-1}LPP^{-1} = L. \tag{5}$$

That is, the information of L is completely included in P and Λ. Since Laplacian matrix has all the information of network structure, Λ and P includes the properties of the network structure, so they should be useful to analyze networks.

III. Indices of Network Structures and Network Generation Model

In this section, we explain the average path length and the clustering coefficient. In addition, in order to investigate the relationship between eigenvalue distribution of the Laplacian matrix and indices of network structures, we make a network generation model, which can construct various networks with different adjacency relationships between nodes while keep the identical degree distribution.

A. Indices of Network Structures

1) Average Path Length: The shortest path between nodes i and j is defined as a path having the smallest number of links between them. The average path length is the mean of shortest paths between all the pairs of distinct nodes. For the length ℓ_{ij} of a given shortest path between node i and j, the average path length ℓ is defined as

$$\ell = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \ell_{ij}. \tag{6}$$

2) Clustering Coefficient: The clustering coefficient measures the degree to which nodes tend to cluster together. This index has become popular because Watts and Strogatz use it to investigate the small-world property of complex networks [14]. Clustering coefficient C_i of node i is defined as

$$C_i = \frac{2M_i}{d_i(d_i - 1)}, \tag{7}$$

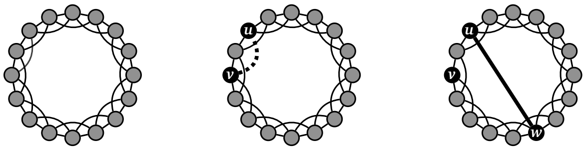
where M_i is the number of links between neighbors of node i.

Clustering coefficient C of a network is the mean of clustering coefficients C_i of its nodes, and is defined as

$$C = \frac{1}{n} \sum_{i=1}^{n} C_i. \tag{8}$$

B. Network Generation Model

In this subsection, on the basis of the WS model, we make a network generation model to generate networks with different adjacency relationships between nodes and identical degree distribution to investigate the relationship between eigenvalues of the Laplacian matrix and indices about information of adjacency relationships between nodes (i.e., average path length and clustering coefficient).



(a) 4–regular network like a ring (b) Cut off a link (u, v) (c) Replace the original link (u, v) with a new link (u, w)

Fig. 2. The algorithm of the WS model.

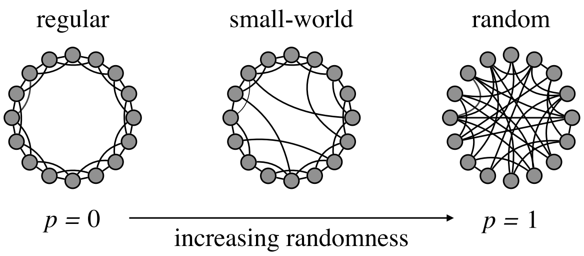


Fig. 3. Schematic diagram of the WS model.

1) Watts and Strogatz Model: WS model is a random network generation model that generates networks with small-world properties, which have small average path length and high clustering coefficient [14].

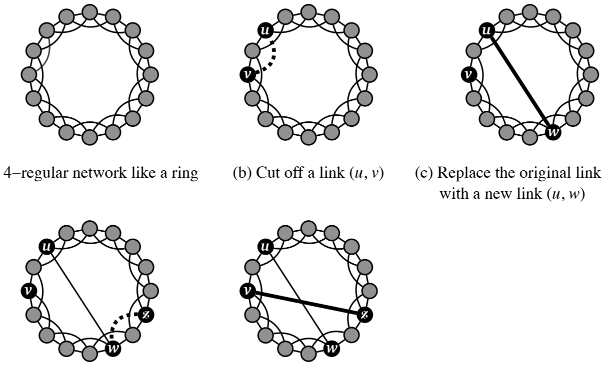
The WS model generates a network by the following steps (Fig. 2).

(a) Construct a k-regular network, where n nodes are connected to k neighbor like a ring.
(b) Cut off a link (u, v) that chosen at random.
(c) Replace the original link (u, v) with a new link (u, w) where w $(w \in V \setminus \{u,\, v,\, \partial u\})$ is chosen at random. ∂u denotes the set of adjacent nodes of u. If A and B are sets, $A \setminus B$ is the relative complement of B in A. Note that we avoid a self-loop and multiple links.
(d) Repeat steps (b) and (c) until the number of rewiring links reaches $p\, n\, k/2$, where p denotes the percentage of the links which are going to change.

The rewiring probability p in the WS model controls the interpolation between the regular network and the random network (Fig. 3). When $p = 0$, the original regular network is not changed. When $p = 1$, all links are rewired randomly, so the generated network is equivalent to a random network. Otherwise, we obtain the network intermingled properties of both regular network and random network.

2) Extended WS model with Identical Degree Distribution: On the basis of the algorithm of the WS model, let us consider network generation model which randomly rewires links while retaining the same degree distribution of the original regular network. In this paper, we call it as the extended WS model (exWS model) with the identical degree distribution. The algorithm of exWS model is as follows (Fig. 4).

(a) Construct a k-regular network, where n nodes are connected to k neighbor like a ring.
(b) Cut off a link (u, v) that is chosen at random.



(a) 4–regular network like a ring (b) Cut off a link (u, v) (c) Replace the original link (u, v) with a new link (u, w) (d) Cut off a link (w, z) (e) Replace the original link (w, z) with a new link (v, z)

Fig. 4. The algorithm of the extended WS model with the identical degree distribution.

(c) Replace the original link (u, v) with a new link (u, w) where w $(w \in V \setminus \{u,\, v,\, \partial u\})$ is chosen at random.
(d) Cut off a link (w, z) where z $(\neq v)$ is a adjacent node of w chosen at random.
(e) Replace the original link (w, z) with a new link (v, z).
(f) Repeat step (b)–step (e) until the number of rewiring links reaches $p\, n\, k/2$.

In steps (b) and (d), we keep all the components of the network connected with connectivity at least 1. In addition, in steps (c) and (e), we avoid a self-loop and multiple links. This network generation model keeps all node degrees to k unlike the WS model.

IV. Numerical Experiment

In this section, we investigate the relationship between the eigenvalues of the Laplacian matrix and the indices of network structures (i.e., average path length and clustering coefficient). First, we generate networks with the exWS model. Then, we evaluate the indices of the generated networks, and compare the eigenvalue distributions of the Laplacian matrices of them. Next, we investigate the relationship between the second smallest eigenvalue of the Laplacian matrix and the indices about adjacency relationships between nodes.

This experiment uses the parameter configuration of the exWS model shown in Table I. We conducted this experiment until the width of 95% confidence interval becomes sufficiently small.

TABLE I
Parameter configuration of the exWS model

number of nodes	$n = 400$
k-regular network	$k = 30$
rewiring probability	$p = 0,\, 0.01,\, \ldots,\, 0.99,\, 1.00$

A. Average Path Length and Clustering Coefficient

Figure 5 shows average path length ℓ for differrent values of rewiring probability p of the exWS model. In this figure, average path length ℓ dramatically changes when p is small.

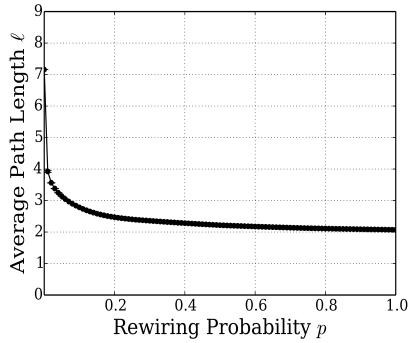


Fig. 5. Average path length ℓ for different values of rewiring probability p.

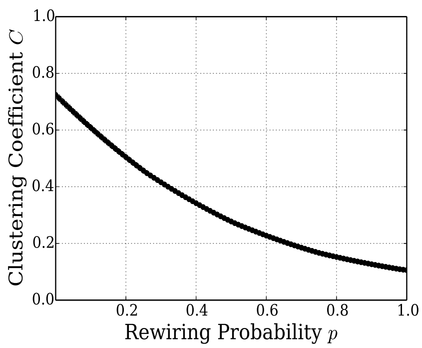


Fig. 6. Clustering coefficient C for different values of rewiring probability p.

This is because shortcuts between distant nodes are created by rewiring links.

Figure 6 shows clustering coefficient C for differrent values of rewiring probability p of the exWS model. In this figure, clustering coefficient C decreases gently as rewiring probability p increases.

B. Eigenvalue Distribution of the Laplacian Matrix

Figure 7 shows eigenvalue distributions of the Laplacian matrices for rewiring probability $p = 0.1$, 0.3, 0.5, and 1.0 of the exWS model. According to figure, we show eigenvalues of the Laplacian matrices in ascending order. In this figure, the rewiring probability p affects the eigenvalues of the Laplacian matrices. Since networks with different values of p have identical degree distribution, the effect of p on the eigenvalues is caused by the information of adjacency relationship between nodes. In particular, the second smallest eigenvalue has a significant influence of rewiring probability p. Hence, we focus on second smallest eigenvalue λ_2, and show λ_2's for different values of rewiring probability p of the exWS model in Fig. 8. In this figure, second smallest eigenvalue λ_2 increases monotonically as rewiring probability p increases. In addition, the second smallest eigenvalue is almost linear function of p.

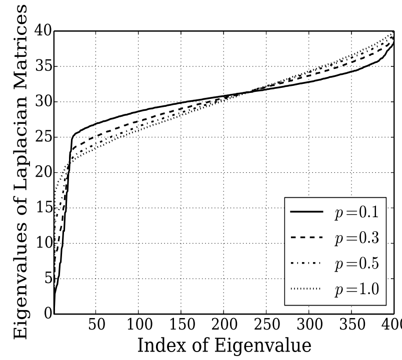


Fig. 7. Eigenvalue disributions of the Laplacian matrices for rewiring probability p = 0.1, 0.3, 0.5, and 1.0.

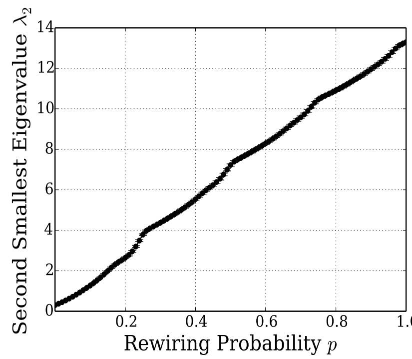


Fig. 8. Second smallest eigenvalue λ_2 of the Laplacian matrix for different values of rewiring probability p.

From Figs. 5, 6, and 7, we can confirm that the eigenvalues of the Laplacian matrix indirectly affects the average path length and the clustering coefficient through rewiring probability p. In particular, the second smallest eigenvalue of the Laplacian matrix indirectly affects the indices. In the following subsections, we investigate the relationship between the second smallest eigenvalue and the average path length, and the second smallest eigenvalue and the clustering coefficient.

C. Relationship between the Second Smallest Eigenvalue of the Laplacian Matrix and the Average Path Length

Figure 9 shows scatter diagrams of $\log_{10} \ell$ and $\log_{10}(\lambda_2 + c)$ $(c = -0.31)$, and its fitted curve $(y = -8.39x + 3.73)$. In this figure, the second smallest eigenvalue and the average path length have negative correlation. In addition, ℓ and $\lambda_2 + c$ have a relationship of power-law function $(\lambda_2 + c = a\,\ell^b)$. According to the gradient and the intercept of the fitted curve, $a = 5.4 \times 10^3$ and $b = -8.39$. From the above result, we obtain the following relationship:

$$\lambda_2 - 0.31 = (5.4 \times 10^3)\,\ell^{-8.39}. \tag{9}$$

Note that $|c|$ is the minimum value of the second smallest eigenvalues for $0 \leq p \leq 1.0$. In the algorithm of the exWS model, we keep all the components of the network connected

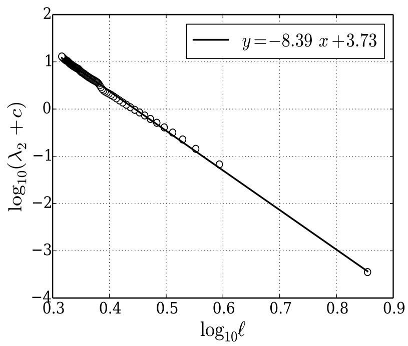


Fig. 9. Relationship between $\log_{10} \ell$ and $\log_{10}(\lambda_2 - 0.31)$.

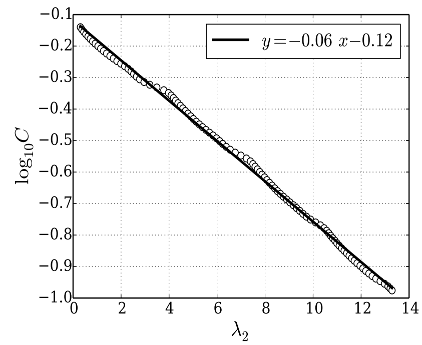


Fig. 11. Relationship between λ_2 and $\log_{10} C$.

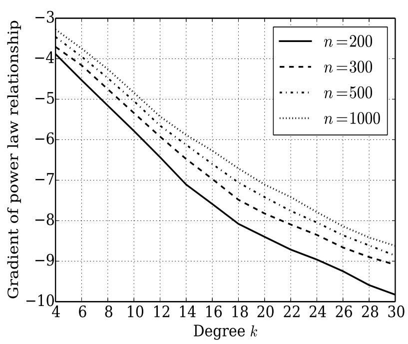


Fig. 10. Gradient of power law relationship between $\log_{10} \ell$ and $\log_{10}(\lambda_2 + c)$ for n = 200, 300, 500, and 1000, and with respect to degree k.

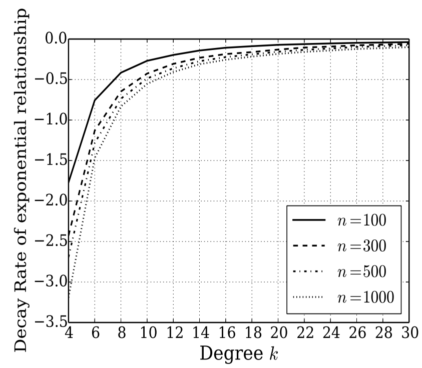


Fig. 12. Decay rate of exponential relationship between λ_2 and $\log_{10} C$ for n = 100, 300, 500, and 1000, and with respect to degree k.

with connectivity at least 1. In addition, the multiplicity of the eigenvalue 0 of the Laplacian matrix is the number of connected components of the network. That is, no matter how the links are rewired, the second smallest eigenvalue remains non-zero.

Figure 10 shows the gradient of the fitted curve for the relationship between $\log_{10} \ell$ and $\log_{10}(\lambda_2 + c)$ with respect to n = 200, 300, 500, and 1000, and different values of k of the exWS model. In this figure, there is a regularity in the parameter n and k of the exWS model.

D. Relationship between the Second Smallest Eigenvalue of the Laplacian Matrix and the Clustering Coefficient

Figure 11 shows scatter diagrams of λ_2 and $\log_{10} C$, and its fitted curve ($y = -0.06\,x - 0.12$). In this figure, the second smallest eigenvalue and the clustering coefficient have negative correlation. In addition, the second smallest eigenvalue and the clustering coefficient have a relationship of exponential function ($C = a\,e^{b\lambda_2}$) where $a = 0.76$ and $b = -0.14$. From the above result, we obtain the following relationship

$$C = 0.76\,e^{-0.14\,\lambda_2}. \quad (10)$$

Figure 12 shows the decay rate b in the exponential relationship between λ_2 and $\log_{10} C$ for n = 100, 300, 500, and 1000, and with respect to degree k of the exWS model. In this figure, there is a regularity in the parameter n and k of the exWS model.

V. Conclusions

In this paper, we investigated the relationship between the eigenvalues of the Laplacian matrix and the indices (i.e., average path length and clustering coefficient) of network structures. In order to investigate the relationship, we developed the extended WS model a network generation model that ganerates networks with different adjacency relationships between nodes and identical degree distribution. Through our numerical experiments, we first showed that the adjacency relationship information affects the eigenvalues of Laplacian matrices. We specifically found that the second smallest eigenvalues is mostly affected by the information. Hence, we focused on the second smallest eigenvalue, and showed the relationship between the second smallest eigenvalue and the indices of network structures. In particular, we showed that ℓ and $\lambda_2 + c$ have a relationship of power-law function ($\lambda_2 + c = a\,\ell^b$), and λ_2 and C have a relationship of exponential function ($C = a\,e^{b\lambda_2}$).

Acknowledgment

This research was supported by Grant-in-Aid for Scientific Research (B) No. 26280032 (2014–2016) from JSPS.

REFERENCES

[1] Jennifer A. Dunne, Richard J. Williams, and Neo D. Martinez, "Food-web structure and network theory: The role of connectance and size," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 20, pp. 12917–12922, October 2002.

[2] V. Rosato, S. Bologna, F. Tiriticco, "Topological properties of high-voltage electrical transmission networks," *Electric Power Systems Research*, vol. 77, no. 2, pp. 99–105, February 2007.

[3] H. Jeong, S. P. Mason, A. L. Barabási and Z. N. Oltvai, "Lethality and centrality in protein networks," *Nature*, vol. 411, no. 6833, pp. 41–42, May 2001.

[4] M. Faloutsos, P. Faloutsos, C. Faloutsos, "On power-law relationships of the Internet topology," *ACM SIGCOMM Computer Communication Review*, vol. 29, no. 4, pp. 251–262, October 1999.

[5] N. Eaglea, A. S. Pentland and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 36, pp. 15274–15278, September 2009.

[6] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks," *Review of Modern Physics*, vol. 74, no. 1, pp. 47–97, January 2002.

[7] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.

[8] P. J. Carrington, J. Scott and S. Wasserman, *Models and Methods in Social Network Analysis*, Cambridge University Press, 2005.

[9] A. Mislove, M. Marcon, K.P. Gummadi, P. Drushel and B. Bhattacharjee, "Measurement and analysis of online social networks," *ACM SIGCOMM conference on Internet measurement*, pp. 29–42, October, 2007.

[10] D. Spielman, "Chapter 16: Spectral graph theory," *Combinatorial Scientific Computing*, ed. by Uwe Naumann and Olaf Schenk, Chapman and Hall/CRC, pp. 495–524, 2012.

[11] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, October 1973.

[12] C. Zhan, G. Chen and L. F. Yeung, "On the distributions of Laplacian eigenvalues versus node degrees in complex networks," *Physica A*, vol. 389, no. 8, pp. 1779–1788, April 2010.

[13] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, October 1960.

[14] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small–world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, June 1998.

[15] M. E. J. Newman and D. J. Watts, "Renormalization group analysis of the small-world network model," *Physics Letters A*, vol. 263, no. 4, pp. 341–346, March 1999.

[16] A. L. Barabási, R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, October 1999.