

Scientific paper

Randić Index, Irregularity and Complex Biomolecular Networks

Ernesto Estrada

Department of Mathematics & Statistics, Department of Physics, and Institute of Complex Systems at Strathclyde, University of Strathclyde, Glasgow G1 1XH, UK

* Corresponding author: E-mail: Ernesto.estrada@strath.ac.uk

Received: 21-06-2010

This paper is dedicated to Professor Milan Randić on the occasion of his 80th birthday

Abstract

A new formulation of Randić index is carried out as a minimization of a quadratic form which involves the Laplacian matrix of a graph. Using this formulation it is easy to realize that Randić index is useful in defining a new index of irregularity, which is then formulated here. A new context for the study of irregularity of graphs is advanced as a necessity for studying complex (biological) networks. We analyze both Randić and irregularity indices for random networks with Poisson and power-law degree distributions. Then, we analyze the irregularity of 10 protein-protein interaction networks in different organisms ranging from 50 to 3000 nodes. Finally, some ‘ruminations’ about the elegance and applicability of Randić index are remarked.

Keywords: Complex networks; scale-free networks; protein-protein interaction networks; graph irregularity; quadratic forms; Laplacian matrix

1. Introduction

Chemistry is defined as the science that studies the “composition, structure, properties, and reactions of matter, especially of atomic and molecular systems”.¹ A great deal of efforts was put forward in the 20th century for understanding the structure and properties of isolated molecular species. Quantum chemistry, molecular mechanics and molecular graph theory are examples of these fruitful attempts to express molecular structure in an abstract way. Such (mathematical) abstractions allow the description, and more importantly the prediction, of many molecular properties, ranging from physico-chemical to complex biological behaviours. However, 21th century has presented a new and more difficult challenge to chemistry. The availability of new experimental techniques allows mapping the interactions of thousands of molecules in cellular environments. As a consequence systems biology was born.² Systems biology is an interdisciplinary approach to study complex interactions in biological systems, resting basically on the study of emergent properties in complex biological networks. Among their targets we can mention genetic networks, transcription net-

works, protein-protein interaction networks, and metabolic networks. More recently, the term system chemistry³ has been coined, which goes beyond biological networks to cover the whole spectrum of emergent properties in chemical networks.

Among the approaches developed in the 20th century to study properties of ‘isolated’ molecules only a tiny fraction is useful to study the new problems posted by systems biology and chemistry. It goes beyond any imagination to ask how quantum chemistry or even molecular mechanics can deal with more than 6,000 interactions between more than 2,700 proteins in the main connected component of the human protein-protein interaction network. In contrast, methods based on network theory, which rest basically on the study of topological properties of these interaction networks, has become the standard analysis in systems biology and chemistry.

In 1975 Milan Randić proposed one of such topological invariants to study properties of small isolated organic molecules.⁴ This index was proposed as a branching index and it is nowadays known as the *Randić index* of a graph.⁴ The Randić index ${}^1R_{-1/2}$ of a graph having $|E| = m$ undirected links is given by⁴

$${}^1R_{-1/2} = \sum_{i,j \in E}^m (k_i k_j)^{-1/2}. \quad (1)$$

Here we designate the degree of a node, i.e., number of links attached to it, by k_i . This index was then generalized by Bollobás and Erdős by changing the exponent $-1/2$ to any real value exponent α .⁵ The Randić index has been intensively studied in mathematics and widely applied in chemistry and biomolecular sciences.^{6–10} Here we propose a new interpretation for this index and a new context for applications in systems biology and chemistry.

2. Network Irregularity and Randić Index

The study of network irregularity can be traced back to the seminal paper of Collatz and Sinogowitz in 1957.¹¹ These authors raised the question of finding the most irregular graphs of a given size by using the following measure

$$CS(G) = \lambda_1 - \bar{k}, \quad (2)$$

where λ_1 is the principal eigenvalue of the adjacency matrix and \bar{k} is the average degree. Clearly this index is zero for regular graphs and Collatz and Sinogowitz conjectured that stars maximize this index.¹¹ They showed that in fact this is the case for graphs with up to 5 nodes. However, Cvetković and Rowlinson¹² rejected the conjecture of Collatz and Sinogowitz¹¹ by constructing several families of graphs having larger values of $CS(G)$ than the corresponding stars of the same size.

Another irregularity index was proposed in 1992 by Bell as the variance of node degrees k_i :¹³

$$VAR^D = \frac{1}{n} \sum_{i=1}^n (k_i - \bar{k})^2. \quad (3)$$

This index was proposed for the first time as a measure of centralization by Sniijders¹⁴ based on the intuition that “centralization is synonymous with the dispersion or heterogeneity” of a node. It takes the minimum value for regular graphs as there is no heterogeneity at all in their degrees. However, the maximum value depends on the number of nodes in the network.

Another measure of graph irregularity was proposed by Albertson as:¹⁵

$$A(G) = \sum_{(i,j) \in E} |k_i - k_j|. \quad (4)$$

Recently, Hansen and Mélot¹⁶ have shown that the Albertson index is not always maximized for star graphs but it is in a particular class of split graphs consisting of a clique, an independent set, and some links joining a node in the clique to another one in the independent set.

Although the three indices are minimized for regular graphs they identify different classes of graphs as the most

irregular ones. A comparison between these indices has been published by Gutman *et al.*¹⁷ Consequently, we are urged to define a new irregularity index from some sort of ‘first-principles’. We need an index minimized by regular graphs and maximized for starlike ones. The reason for this necessity will become clear in the next section.

Let us restate the problem of finding an irregularity index as that of minimizing the quantity

$$\sum_{i,j \in E} \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2. \quad (5)$$

The intuition behind this definition of irregularity is that for regular networks this quantity is equal to zero as in the case of the other irregularity indices previously defined in the literature. However, as the difference in the degrees of adjacent nodes increases, the index increases. For instance, let us consider a node of degree one connected to a node of degree k . Then, as k increases the term inside the bracket tends to one. If we consider the general case

$$\sum_{i,j \in E} (k_i^\alpha - k_j^\alpha)^2, \quad (6)$$

we can see these indices as a generalization of the Albertson index of irregularity¹⁵ in which instead of using a modulus function we use the square one. Now, let us define the Laplacian matrix \mathbf{L} of a network whose entries are given by¹⁸

$$L_{ij} = \begin{cases} k_i & \text{for } i = j, \\ -1 & \text{for } i \sim j, \\ 0 & \text{otherwise,} \end{cases}$$

where $i \sim j$ stands for pairs of adjacent nodes. The matrix \mathbf{L} can be obtained as $\mathbf{L} = \mathbf{K} - \mathbf{A}$, where \mathbf{K} is a diagonal matrix of degrees and \mathbf{A} is the adjacency matrix of the network, whose A_{ij} is one if, and only if, the corresponding nodes are joined by a link or zero otherwise. This discrete Laplacian has been intensively studied in algebraic graph theory and the reader is referred to the excellent reviews for details.^{19, 20} Let $|\mathbf{k}^\alpha\rangle = (k_1^\alpha, k_2^\alpha, \dots, k_n^\alpha)$ represents a column vector where k_i is the degree of the node i . Then, it is easy to realize that expression (6) can be stated as a quadratic form of the Laplacian matrix:

$$\begin{aligned} \sum_{i,j \in E} (k_i^\alpha - k_j^\alpha)^2 &= \frac{1}{2} \langle \mathbf{k}^\alpha | \mathbf{L} | \mathbf{k}^\alpha \rangle \\ &= \sum_{j=1}^n (k_j^\alpha)^{2\alpha+1} - 2 \sum_{i,j \in E} (k_i k_j)^\alpha \end{aligned} \quad (7)$$

The second term in the right hand side of eq. (7) is two times the *generalized Randić index* R_α .^{8–10} Let ${}^0R_\alpha = \sum_{i,j \in E} (k_i k_j)^\alpha$ be designated as the *generalized zeroth-order Randić index*.^{8–10} It is straightforward to realize that

the quantity to be minimized is the difference between the $R_{2\alpha+1}$ index and twice the Randić index ${}^1R_{\alpha} R_{2\alpha+1} - 2{}^1R_{\alpha}$.

$$\sum_{i,j \in E} \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2 = n - 2{}^1R_{-1/2}. \quad (8)$$

Now, let us consider some known facts about the ${}^1R_{-1/2}$ index. Among connected networks with n nodes, the star S_n attains the minimum ${}^1R_{-1/2}$ index and the regular ones have maximum ${}^1R_{-1/2}$. Then, among connected networks of size n , the original Randić index is bounded as:^{8–10}

$$\sqrt{n-1} \leq {}^1R_{-1/2} \leq \frac{n}{2}, \quad (9)$$

where the lower bound is reached for the star S_n and the upper bound is attained for any regular network with n nodes irrespective of its degree. Then, we can normalize the Randić index as follows

$$\rho(G) = \frac{n - 2{}^1R_{-1/2}}{n - 2\sqrt{n-1}}, \quad (10)$$

which gives a zero value for regular graphs and one for the star. It is straightforward to realize that

$$\rho(G) = \frac{1}{n - 2\sqrt{n-1}} \sum_{i,j \in E} \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2, \quad (11)$$

which clearly identifies $\rho(G)$ as a normalized irregularity index for a network.

We would like to close this section by developing a little bit further the connection between the Randić index and the graph Laplacian. Let $\vec{\varphi}_j$ be an orthonormal eigenvector of the Laplacian matrix associated with the μ_j eigenvalue. We recall that the Laplacian matrix is positive semidefinite. Let

$$\cos \theta_j = \frac{\mathbf{k}^{-1/2} \cdot \vec{\varphi}_j}{\|\mathbf{k}^{-1/2}\|}, \quad (12)$$

be the angle formed between the orthonormal eigenvector $\vec{\varphi}_j$ and the vector $\mathbf{k}^{-1/2}$ previously defined. The Euclidean norm $\|\mathbf{k}^{\alpha}\|$ can be written as $\|\mathbf{k}^{\alpha}\| = \sqrt{{}^0R_{2\alpha}}$. Then, the generalized Randić index can be written as

$${}^1R_{\alpha} = \frac{1}{2} \left[{}^0R_{2\alpha+1} - {}^0R_{2\alpha} \sum_{j=2}^n \mu_j \cos^2 \theta_j \right]. \quad (13)$$

In particular, the original Randić index is expressed as follows

$${}^1R_{-1/2} = \frac{1}{2} \left[n - {}^0R_{-1} \sum_{j=2}^n \mu_j \cos^2 \theta_j \right]. \quad (14)$$

The term $\cos^2 \theta_j$ represents the ‘‘contribution’’ of the normalized degree to the corresponding eigenvector (or vice versa). For instance, $\cos^2 \theta_j = 0$ means that the vector \mathbf{k}^{α} is perpendicular to the corresponding eigenvector, and no ‘‘duplicated’’ information is contained in both vectors.

3. Why Does Irregularity Matter?

The study of complex networks representing systems in disparate real-world contexts ranging from biological to technological systems has become an important area of interdisciplinary research.^{21–23} These complex networks share many universal topological characteristics such as *small-worldness*,²⁴ *scale-freeness*,²⁵ the existence of network motifs²⁶ and self-similarity characteristics.²⁷ In particular, the power-law degree distribution of many real-world networks contrasts with the regularity observed in random models like the one proposed by Erdős and Rényi.²⁸

Let us consider the property of scale-freeness in more detail. Let $p(k) = n(k)/n$ be the probability of randomly selecting a node of degree k in a network, where $n(k)$ is the number of nodes having degree k in a network of size n . Then, a plot of $p(k)$ versus k represents the degree distribution for the network.^{22,23} A random network displays a Poisson degree distribution of the form

$$p(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!}, \quad (15)$$

where \bar{k} is the average degree of the network.

Most biological networks behave in a completely different way. Instead of having such kind of ‘democratic’ distribution of node degrees, they display a very ‘egoistic’ one, in which very few nodes have very large degrees and most of the nodes has relatively low degrees. In mathematical terms, the degree distribution of these networks follows a power-law of the form:^{22, 23}

$$p(k) \sim k^{-\gamma}. \quad (16)$$

In Fig. 1 we illustrate the plots of Poisson and power-law degree distributions for a hypothetical network.

The term ‘scale-free’ describes the fact that by scaling the degree by a constant factor c , only produces a proportionate scaling of the function:

$$p(k, c) = A(ck)^{-\gamma} = Ac^{-\gamma} \cdot p(k). \quad (17)$$

Power-law relations are usually represented in a logarithmic scale, where we obtain a straight line, $\ln p(k) = -\gamma \ln k + \ln A$, where $-\gamma$ is the slope and $\ln A$ the intercept of the function.

However, strictly speaking there are not very many networks displaying a ‘perfect’ power-law degree distri-

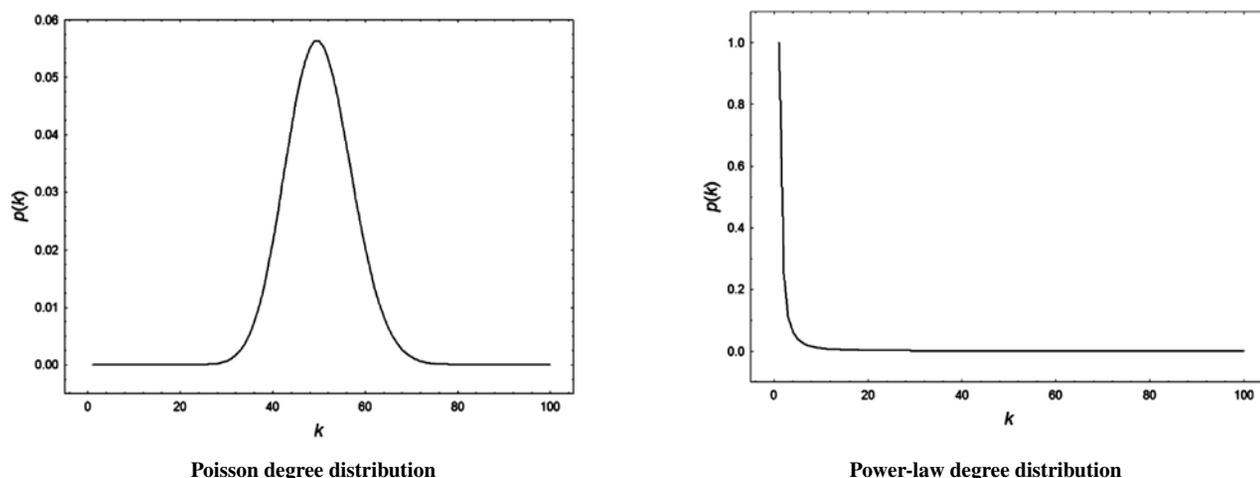


Figure 1. Illustration of Poisson and power-law degree distributions for hypothetical networks.

bution. Instead, a complete zoo of ‘fat-tail’ distributions, such as lognormal, Burr, logGamma, Pareto, etc., can be found in complex (biological) networks. Then, the more generic term of ‘fat-tail’ degree distribution has been proposed to globalize those distributions. The idea is that networks with ‘fat-tail’ distributions display a large heterogeneity, or in our terms a large irregularity.

The identification of fat-tail degree distribution in networks has become an intense area of research in system biology and other interdisciplinary areas of research. The consequences that a network displays such kind of degree distributions are remarkable.^{21–23} For instance, let us consider a protein-protein interaction network, in which nodes represent proteins and links represent interactions between pairs of proteins in a cell. If such network displays a fat-tail degree distribution it will be robust to random failures of their nodes. That is, due to the fact that about 80% of proteins have low number of interactions, their failure does not produce the collapse of the whole system. However, a targeted attack to the most connected proteins, the very few with very large degrees, disconnects the network into isolated chunks, which are unable to function as a system. In other words, if a network is very irregular, i.e., it is *starlike*, it will be robust to random failures but very vulnerable to intentional attacks.

The consequences of the previously described relationship between vulnerability and irregularity are tremendous. Irregular biological networks can be very vulnerable to the failure of some specific nodes indicating possible sources of diseases as well as possible targets for drug design. The question that remains is whether we need an index of irregularity for such networks instead of analyzing their degree distributions.

Identifying the degree distribution of a network is not a trivial question. There are hundreds of possible distributions to test. Sometimes the differences in the statistical fittings between several distributions are very small.

More difficult is the situation for relatively small networks, where the number of data points is not enough for having a good fit of any of the candidate distributions. Consequently, having an index such as $\rho(G)$, which has its minimum for regular networks and its maximum for star-like ones, is an urgent necessity.

4. Irregularity in Random Networks

Now, we are going to investigate random graph models used to model real-world complex networks. The Erdős-Rényi graphs²⁸ $G_{n,p}$ are generated by considering n nodes which are linked by pairs according to a probability p , $0 \leq p \leq 1$. It is well-known that when $p \sim \omega(n) \log n/n$, where $\omega(n) \rightarrow \infty$, $G_{n,p}$ is almost surely connected, and the degrees of almost all nodes are asymptotically equal. Then, in this regime ${}^1R_{-1/2}(ER) \approx n/2$ and $\rho(G) \approx 0$.

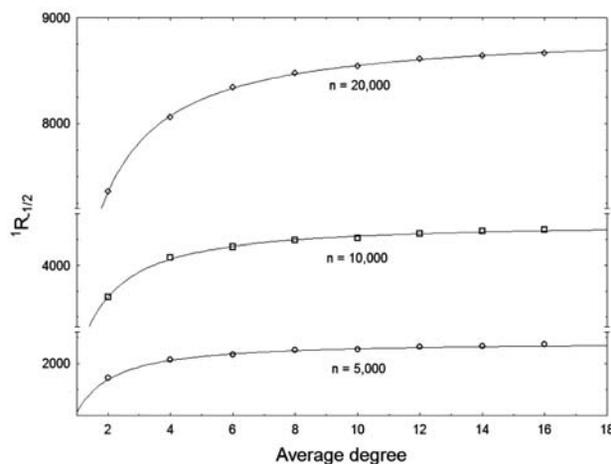


Figure 2. Change of Randić index as a function of average degree for networks generated with Barabási-Albert model of preferential attachment.

Another popular model for generating random graphs is the Barabási-Albert (BA) model,²⁵ which produces random graphs with power-law degree distributions. In this model the graphs are generated from an initial set of $m_0 \geq 2$ nodes. The new nodes are added one at a time by connecting them to m of the existing nodes according to the probability $p_i = k_i / \sum_j k_j$. We have generated BA graphs with 5,000; 10,000 and 20,000 nodes and average degrees ranging from 2 to 16. Then, we have averaged the values of ${}^1R_{-1/2}$ for 100 realizations of each graph. In Fig. 2 we illustrate the plots of ${}^1R_{-1/2}$ versus $\langle k \rangle$. The plots can be approximated by using the following power-law expression,

$${}^1R_{-1/2}(BA) \approx \frac{n}{2.27 + \bar{k}^{-1.18}}. \quad (18)$$

Accordingly, the irregularity index for BA networks is approximated by the following expression:

$$\rho(BA) \approx \left(\frac{0.27 + \bar{k}^{-1.18}}{2.27 + \bar{k}^{-1.18}} \right) \left(\frac{n}{n - 2\sqrt{n-1}} \right), \quad (19)$$

which for very large networks, $n \rightarrow \infty$, with very large average degree approximates to a constant value, $\rho(BA) \approx 0.1189$. The first conclusion that can be extracted from these results is that the scale-free networks are more irregular than random networks with homogeneous degree distributions, which is a trivial conclusion based on the proper definition of both kinds of networks. However, in particular the BA model is a poor generator of such irregularity in graphs, which is able to produce only about 12% of the irregularity of a star-like graph. In addition, it can be clearly seen that the increase of the average degree in scale-free graphs reduces the irregularity. Thus, both degree distribution and average degree play important roles in understanding the irregularity of a network.

5. Irregularity in Protein-protein Interaction Networks

Here we study 10 protein-protein interaction (PPI) networks. In these networks, nodes represent proteins and undirected links represent the interaction between two proteins determined experimentally. The networks studied correspond to following organisms: *D. melanogaster* (fruit fly),²⁹ Kaposi sarcoma herpes virus (KSHV),³⁰ *P. falciparum* (malaria parasite),³¹ varicella zoster virus (VZV),³⁰ human,³² *S. cerevisiae* (yeast),³³ *A. fulgidus*,³⁴ *H. pylori*,³⁵ *C. elegans*³⁶ and *B. subtilis*.³⁷ We study only the largest (main) connected component of each of these networks, which ranges from 50 to 3039 proteins.

In Table 1 we give the values of irregularities of all these PPIs together with the values calculated for random networks with preferential attachment using BA model.

All networks display significantly larger irregularity than the regular random networks. That is, all networks have $\rho(G) \gg 0$, which indicates that they have some kind of fat-tail degree distributions characterized by very few hubs that keep the whole network connected. In comparing these networks with those generated at random using BA preferential attachment we see that most of them display larger irregularity than the BA ones. In particular the PPI of yeast, *S. cerevisiae*, displays a ratio $\rho(G) / \rho(BA) = 1.74$, followed by the human PPI with $\rho(G) / \rho(BA) = 1.49$.

Table 1. Irregularity of complex protein-protein interaction networks of several organisms.

No.	Network	n	m	$\rho(G)$	$\rho(BA)$	ratio
1	<i>D. melanogaster</i>	3039	3687	0.148	0.246	0.602
2	KSHV	50	118	0.151	0.246	0.614
3	<i>P. falciparum</i>	229	604	0.174	0.196	0.888
4	VZV	53	154	0.278	0.227	1.225
5	Human	2783	6007	0.283	0.190	1.489
6	<i>S. cerevisiae</i>	2224	6829	0.294	0.169	1.740
7	<i>A. fulgidus</i>	32	37	0.308	0.372	0.828
8	<i>H. pylori</i>	710	1396	0.323	0.205	1.576
9	<i>C. elegans</i>	314	363	0.383	0.273	1.403
10	<i>B. subtilis</i>	84	98	0.386	0.309	1.249

Stumpf and Ingram⁴⁰ have carried out a detailed statistical analysis of the empirical data existing for the yeast PPI showing that the best fit for the degree distributions are the stretched exponential followed by a lognormal distribution. The same distribution is observed for *H. pylori*.⁴⁰ The stretched exponential distribution is also known as the Weibull distribution and has the following form:

$$p(k) = \frac{\gamma}{k} \left(\frac{k}{\bar{k}} \right)^{(\gamma-1)} e^{-(k/\bar{k})^\gamma}. \quad (20)$$

The PPI of *C. elegans* displayed a power-law degree distribution and that of *D. melanogaster* was fit to a gamma distribution.⁴⁰ Not all networks analyzed here were previously discussed by Stumpf and Ingram. In addition these authors do not study an exhaustive list of all possible distributions for the networks analyzed. Then, it is difficult to obtain a parallel between both approaches. According to our irregularity measure, the PPI of *D. melanogaster* appears as the least heterogeneous one while that of *C. elegans* is the most irregular among those considered by Stumpf and Ingram. What is clear here is that the current approach allows a classification of networks according to their irregularity, which is not possible by using degree distributions when a variety of fits are considered. In Fig. 3 we illustrate two of the smallest networks with close to extreme values of irregularity, i.e., that of KSHV and that of *B. subtilis*.

12. D. Cvetković, P. Rowlinson, *Publ. Inst. Math. (Beograd)*, **1988**, *44*, 29–34.
13. F. K. Bell, *Linear Algebra Appl.* **1992**, *161*, 45–54.
14. T. A. B. Snijders, *Social Networks* **1981**, *3*, 163–174.
15. M. O. Albertson, *Ars Combinatoria* **1997**, *46*, 219–225.
16. P. Hansen, H. Mélot, in S. Fajtlowicz (Eds.), *Graphs and Discovery, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, Providence, American Mathematical Society, **2005**, *69*, 253–264.
17. I. Gutman, P. Hansen, H. Mélot, *J. Chem. Inf. Model.* **2005**, *45*, 222–230.
18. N. L. Briggs, *Algebraic Graph Theory*, Cambridge University Press, 1974.
19. B. Mohar, in G. Hahns, G. Sabidussi (Eds.), *Graph Symmetry: Algebraic Methods and Applications*, NATO ASI Ser. C 497, Kluwer, **1997**, 225–275.
20. T. Bitikošlu, J. Leydold, P. F. Stadler, *Laplacian Eigenvector of Graphs*, Springer-Verlag, Berlin, **2007**.
21. S. H. Strogatz, *Nature* **2001**, *410*, 268–276.
22. R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **2002**, *74*, 47–97.
23. M. E. J. Newman, *SIAM Rev.* **2003**, *45*, 167–256.
24. D. J. Watts, H. Strogatz, *Nature* **1998**, *393*, 440–442.
25. A.-L. Barabási, R. Albert, *Science* **1999**, *286*, 509–512.
26. R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, *Science* **2002**, *298*, 824–827.
27. C. Song, S. Havlin, H. A. Makse, *Nature* **2005**, *433*, 392–395.
28. P. Erdős, A. Rényi, *Publ. Math.* **1959**, *6*, 290–297.
29. L. Giot, J. S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y. L. Hao, C. E. Ooi, B. Godwin, E. Vitols, G. Vijayadarmodar, P. Pochart, H. Machineni, M. Welsh, Y. Kong, B. Zerkhusen, R. Malcolm, Z. Varrone, A. Collis, M. Minto, S. Burgess, L. McDaniel, E. Stimpson, F. Spriggs, J. Williams, K. Neurath, N. Ioime, M. Agee, E. Voss, K. Furtak, R. Renzulli, N. Aanensen, S. Carrolla, E. Bickelhaupt, Y. Lazovatsky, A. DaSilva, J. Zhong, C. A. Stanyon, R. L. Finley Jr., K. P. White, M. Braverman, T. Jarvie, S. Gold, M. Leach, J. Knight, R. A. Shimkets, M. P. McKenna, J. Chant, J. M. Rothberg, *Science* **2003**, *302*, 1727–1736.
30. P. Uetz, Y.-A. Dong, C. Zeretzke, C. Atzler, A. Baiker, B. Berger, S. Rajagopala, M. Roupelieva, D. Rose, E. Fossum, J. Haas, *Science* **2006**, *311*, 239–242.
31. D. J. LaCount, M. Vignali, R. Chettier, A. Phansalkar, R. Bell, J. R. Hesselberth, L. W. Schoenfeld, I. Ota, S. Sahasrabudhe, C. Kurschner, S. Fields, R. E. Hughes, *Nature* **2005**, *438*, 103–107.
32. J.-F. Rual, K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G. F. Berriz, F. D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou, N. Klitgord, C. Simon, M. Boxem, S. Milstein, J. Rosenberg, D. S. Goldberg, L. V. Zhang, S. L. Wong, G. Franklin, S. Li, J. S. Alcala, J. Lim, C. Fraughton, E. Llamas, S. Cevik, C. Bex, P. Lamesch, R. S. Sikorski, J. Vandenhaute, H. Y. Zoghbi, A. Smolyar, S. Bosak, R. Sequerra, L. Doucette-Stamm, M. E. Cusick, D. E. Hill, F. P. Roth, M. Vidal, *Nature* **2005**, *437*, 1173–1178.
33. C. von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, P. Bork, *Nature* **2002**, *417*, 399–403.
34. M. Motz, I. Kober, C. Girardot, E. Loeser, U. Bauer, M. Albers, G. Moeckel, E. Minch, H. Voss, C. Kilger, M. Koegl, *J. Biol. Chem.* **2002**, *277*, 16179–16188.
35. C.-Y. Lin, C.-L. Chen, C.-S. Cho, L.-M. Wang, C.-M. Chang, P.-Y. Chen, C.-Z. Lo, C. A. Hsiung, *Bioinformatics* **2005**, *21*, 1288–1290.
36. A. David, P. Bello, N. Thierry-Mieg, J. Vaglio, J. Hitti, L. Doucette-Stamm, D. Thierry-Mieg, J. Reboul, S. Boulton, A. J. Walhout, O. Coux, M. Vidal, *EMBO Rep.* **2001**, *2*, 821–828.
37. P. Noirot, N. F. Noirot-Gross, *Curr. Op. Microb.* **2004**, *7*, 505–512.
38. E. Estrada, M. Fox, D. Higham, G.-L. Oppo (Eds.) *Network Science. Complexity in Nature and Technology*, Springer, **2010** in press.
39. G. Farmelo (Ed.), *It Must be Beautiful: Great Equations of Modern Science*. Granta, U.K. **2003**.
40. M. P. H. Stumpf, P. J. Ingram, *Europhys. Lett.* **2005**, *71*, 152–158.

Povzetek

Nova formulacija Randićevega indeksa je izpeljana kot minimizacija kvadratne oblike, ki vključuje Laplaceovo matriko grafa. Iz te formulacije je enostavno razvidno, da Randićev indeks lahko uporabimo za definicijo novega indeksa »nepravilnosti«, ki je tukaj predstavljen. Potreba po novem načinu študija nepravilnosti grafov se je pokazala pri obravnavi kompleksnih bioloških sistemov, predstavljenih v obliki mrež. Analizirali smo Randićeve in »nepravilnostne« indekse za ključne mreže s Poissonovo in potenčno porazdelitvijo. Nadalje smo analizirali nepravilnost desetih mrež, ki ponazarjajo interakcije med proteini v različnih organizmih, v območju med 50 in 3000 vozlišči. V zaključnem razmisleku je poudarjena eleganca in uporabnosti Randićevega indeksa.