31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Structure of Brain Functional Networks

Oleksii Kuchaiev, Po T. Wang, Zoran Nenadic, and Nataša Pržulj

Abstract—**Brain is a complex network optimized both for segregated and distributed information processing. To perform cognitive tasks, different areas of the brain must "cooperate," thereby forming complex networks of interactions also known as brain functional networks. Previous studies have shown that these networks exhibit "small-world" characteristics. Small-world topology, however, is a general property of all brain functional networks and does not capture structural changes in these networks in response to different stimuli or cognitive tasks. Here we show how novel graph theoretic techniques can be utilized for precise analysis of brain functional networks. These techniques allow us to detect structural changes in brain functional networks in response to different stimuli or cognitive tasks. For certain types of cognitive tasks we have found that these networks exhibit geometric structure in addition to the small-world topology. The method has been applied to the electrocorticographic signals of six epileptic patients.**

I. Introduction

In response to external stimuli and/or cognitive processes, cortical neurons generate a series of electrical pulses, commonly known as action potentials, which propagate through various cortical layers and areas. Direct measuring of action potentials requires invasive recording procedures, such as intracranial implantation of recording microelectrodes. A far more common recording modality in human studies is to record gross potentials, normally representing the average activity of many neurons. These gross potentials can be measured either non-invasively, by means of electroencephalography (EEG), or invasively, by means of electrocorticography (ECoG). These measurements then can be used to construct *brain functional networks* (BFNs). In these networks (also called *graphs* throughout this paper), nodes are abstractions of the regions of the brain (represented by sensors) and two nodes are connected by an edge if the signals at corresponding sensors are dependent.

There are many empirical and theoretical reasons why it is useful to model brain activity using graph theory. First, the brain is a complex network itself, consisting of millions of neurons and connections between them. Second, to successfully perform any cognitive task, different areas of the brain need to work together. Finally, during the brain's evolution, the structure of brain functional networks have been optimized for efficient information processing [2]. Hence, understanding the structural properties of brain functional networks will help us gain new insight into the ways in which the brain performs its functions.

Recent studies of brain functional networks [1], [2] showed that they exhibit small-word properties. The small-world property was introduced by Watts and Strogatz [3] and observed in many social and biological networks. Intuitively, this property quantifies the efficiency of connectivity structure in a network. That is, small-world networks have higher clustering coefficients compared to random networks of the same size and small average pathlengths. It is believed that such a network structure has a lower cost of information processing [2] and that therefore, BFNs have evolved to have this structure. Note however, that the small-world topology is a very general property. Many topologically different networks can exhibit this property. Hence, it is important to investigate the fit of more restrictive network models to BFN data. Furthermore, whether different cognitive tasks performed by a subject result in changes in BFN structure and what random graph models provide the best fit for those BFNs is in itself an interesting research question. Using novel graph-theoretic methods [4], we were able to show how the topology of BFNs corresponding to different cognitive tasks changes throughout the course of a single experiment. We further show that in addition to their small-world topology these networks are consistent with *geometric networks* [5] (defined in section IV.C below).

II. Experimental Setup

The analyzed data represents ECoGs of six epileptic patients, measured invasively by subdural or intracortical implantation of recording electrodes. All the patients suffered from intractable epilepsies (i.e. did not respond well to a drug therapy) and elected to have a brain surgery, with the purpose of removing, or simply resecting (disconnecting), epileptogenic zones—the areas of the brain where seizures originate from. In preparation to these types of surgeries, the role of ECoG measurement is two-fold (in both cases, the technique is considered to be the "gold standard"): (i) Epileptogenic zone localization and (ii) Eloquent brain mapping. Table 1 lists the type and number of ECoG electrodes that were analyzed for each subject. For the original reference and related follow-up studies, the

Manuscript received April 21, 2009. This work was supported by the Opportunity Award from the Center for Complex Biological Systems at University of California Irvine, and University of California Irvine Academic Senate Council on Research, Computing and Libraries.

O. Kuchaiev is with the Computer Science Department, University of California, Irvine (e-mail: oleksii.kuchaiev@uci.edu).

P. Wang is with the Department of Biomedical Engineering , University of California, Irvine (e-mail: ptwang@uci.edu).

Z. Nenadic is with Department of Biomedical Engineering , University of California, Irvine (phone: 949-824-1548 e-mail: znenadic@uci.edu).

N. Przulj is with the Computer Science Department, University of California, Irvine (e-mail: natasha@ics.uci.edu).

TABLE I
TYPES OF ECOG ELECTRODES

Subject	Electrode type	Number of electrodes
S001	Bilateral depth	75
S002	R grids	123
S003	R depth/grids	76
S004	L grids	126
S005	Bilateral depth	61
S006	Bilateral depth	91

Types of the ECoG data that were analyzed. The major differences between ECoG grids and depth electrodes are: (i) depth electrodes are inserted (perpendicular to the cortical surface) into the brain tissue, and (ii) depth electrode contacts are much smaller (~1mm) with inter-electrode spacing of 8 mm compared to grid electrodes with ~4mm and ~1cm, respectively.

reader is referred to [6], [7], [8]. These patients were implanted with electrodes usually for a week, and sometimes longer. During this time, they are typically waiting for seizures to happen, otherwise doing nothing. Since this is a rather uninteresting procedure, some of them sign up for scientific studies, such as the ones presented here. It should be noted that these experiments have nothing to do with clinical procedures and are purely voluntary. It should also be noted that the electrode placement was based solely on the requirements of the clinical evaluation without any regard for the scientific study.

A brief account of the experiment is presented next (Fig. 1). For the original reference the reader is referred to [6]. At

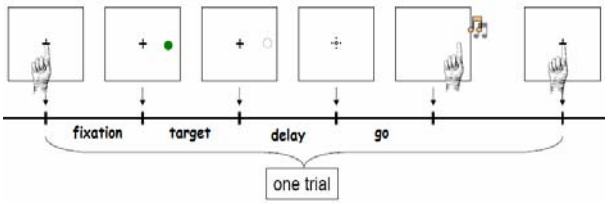

Fig. 1. Experimental protocol. Each trail has four stages: fixation, target, delay and go. During each stage, the brain is supposed to control and perform different tasks.

the beginning of each trial, a fixation stimulus is presented in the center of a touch screen and the participant initiates the trial by placing his right hand on the stimulus. This marks the onset of the *fixation* stage. After a short period, a peripheral target is flashed on the screen, at one of the 6 or 8 locations (depending on the subject), which marks the onset of the *target* stage. The subjects were instructed not to move their eyes, but no quantitative assessment of their eye movements was available. The target stage is followed by the *delay* stage, whose onset is defined by extinguishing the peripheral target. The fixation is extinguished after the *delay* stage, and it acts as a cue for the participants to reach to the memorized location and marks the onset of the *go* stage. Upon completion of the *go* stage, a beep indicates whether it was correct or incorrect and a new trial is initiated after a brief *delay* period. The position of the peripheral target was randomly interleaved on a trial-by-trial basis. The number of trials varied between 69 and 82 per target location, yielding a total number of trials between 1,532 and 1,948. To prevent phase-locked response, the durations of *fixation, target* and *go* stages were randomized (uniform distribution between 1 and 1.3 s).

III. METHODS

We model the *brain functional network* as an undirected graph (network), $G = (N, E)$, where N is a set of nodes and E is a set of edges. Each node, n, from the node set N corresponds to one of the ECoG electrodes. Since electrodes measure signals in different areas of the brain, functionally, nodes represent these regions. We put an edge between two nodes u and v if the signals on the corresponding sensors are "dependent enough" (see below).

A. Constructing BFNs

For each subject, stage and target location we take the median of the signals at each sensor calculated over trials and we work further with these data. Note that this procedure significantly reduces the noise in ECoG signals. From these data BFNs are constructed for each subject, stage and target location, as explained below.

There are many ways to decide whether signals at two sensors are co-dependent, and the simplest one is to calculate the correlation coefficient between two signals. Then, if this coefficient is higher than some chosen threshold, the nodes corresponding to these sensors are connected by an edge. While this is the most straightforward approach that has been used frequently [2], there are some serious drawbacks. Before reaching the sensor, a signal in the brain may undergo nonlinear transformations. Generally, this may significantly alter the correlation between two signals, and even change the sign of the

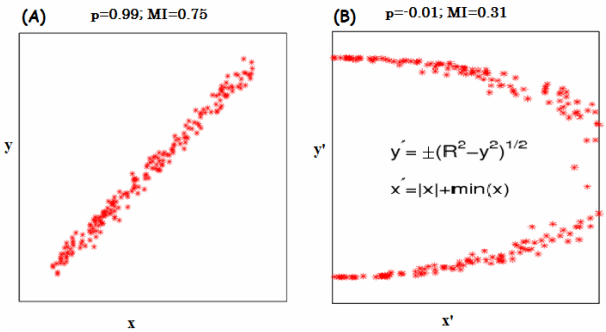

Fig. 2. (A): Two signals x and y are highly correlated with Pearson correlation coefficient p=0.99 and mutual information MI=0.75 between them. (B): Two signals x' and y' are non-linear transformations of signals x and y. The correlation between signals has been destroyed by this transformation (p=-0.01). However, as it can be seen from the picture, there is still a strong dependency between x' and y' and about half of the mutual information (MI=0.31) between signals was preserved.

correlation (see Fig. 2). Hence, a more robust connectivity metric needs to be chosen that is more tolerant to non-linear transformations in the data. Furthermore, the correlation coefficient measures the strength of linear interactions between two random variables, however it ignores nonlinear ones. For this purpose, we chose *mutual information*, which is more robust to nonlinear transformations and can measure

nonlinear interactions in the data [9]. *Mutual information* between two continuous random variables X and Y is defined as

$$\mu I(X,Y) = h(X) - h(X \mid Y) \quad (1)$$

where $h(X)$ is differential or Shannon's *entropy* of the random variable X, defined as

$$h(X) = -\int p_X(x) \log_2 p_X(x) dx \cdot \quad (2)$$

Here, $p_X(x)$ is the probability density function of the random variable X. Since we do not know $p_X(x)$, we estimate it from the data by using histograms. Thus, we calculate the discretized entropy and mutual information. It can be shown that if the discretization step is Δ, for discretized entropy, $H(X)$, and differential entropy, $h(X)$, the following is true [9]:

$$\lim_{\Delta \to 0} (H(X) + \log_2 \Delta) = h(X) \cdot \quad (3)$$

Therefore, mutual information between two continuous random variables is

$$\lim_{\Delta \to 0} (H(X) - H(X \mid Y)) = h(X) - h(X \mid Y) = \mu I(X,Y) \quad (4)$$

Hence, we estimate the amount of mutual information between two signals using discretized versions of the entropies, $H(X)$ and $H(X|Y)$ (this is the conditional entropy of X given that the value of Y is known). The number of bins for the histograms was chosen using Sturges' formula [9]. After calculating $\mu I(X,Y)$, we normalize it by $\max\{H(X), H(Y)\}$ and get the connectivity matrix M:

$$M(i,j) = \frac{\mu I(i,j)}{\max\{H(i), H(j)\}}, \quad i, j = 1 \ldots n \cdot \quad (5)$$

Next, we need to choose an appropriate threshold for $M(i,j)$ to decide whether to put an edge between nodes i and j. To avoid arbitrariness in setting the threshold, we perform further statistical analysis. If two sensors i and j have some amount of mutual information $\mu I(X,Y) = x$, we want to know whether this amount of information is statistically significant. Thus, we calculate the following p-values:

$$P(\mu I(i,j) \geq x \mid i \text{ and } j \text{ are independent}) \quad (6)$$

Empirically, we found that the threshold of 0.2 for the normalized mutual information always corresponds to p-values lower than 10^{-3}.

B. Graph theoretic analysis

1) Network parameters

First, we calculate the standard graph theoretic parameters of the networks, such as the *average clustering coefficient (C)* and the *average pathlength (L)*. Based on these two parameters, it is possible to classify a network as being small-world. The clustering coefficient of a node v with degree k (number of its neighbors) is defined as

$$C(v) = \frac{2 * E_v}{k(k-1)} \quad (7)$$

Here, E_v is the number of edges amongst the neighbors of node v. For nodes with degree $k=0$ or 1 (nodes without neighbors or only with one neighbor), the clustering coefficient is defined to be zero. Then, the average clustering coefficient of the network is simply the average of clustering coefficients over all of its nodes.

The pathlength between two nodes v and w is the number of edges on the shortest path between them. The average pathlength of the graph is the average of pairwise pathlengths between all pairs of its nodes. Small-world networks are defined to be those networks that have clustering coefficients significantly larger than the clustering coefficients of random networks of the same size and average pathlengths approximately the same as random networks of the same size [3]. Note that random networks have small average pathlengths [10]. As in [2], we define $\lambda = \frac{L}{L_{rand}}$ and $\gamma = \frac{C}{C_{rand}}$, where L_{rand} and C_{rand} are averaged values of pathlengths and clustering coefficients of random networks with the same number of nodes and edges as the BFNs synthesized from the data. For small-world networks, λ should be approximately equal to 1 and γ should be greater than 1. Hence, if the quantity $\sigma = \frac{\gamma}{\lambda}$ is greater than 1, the network is considered to have the small-world topology [2].

2) Network comparison

To capture the structural changes in BFNs as the brain engages in the 4 cognitive tasks (corresponding to the 4 experimental stages), we need to compare different BFNs and quantify their differences and similarities.

The first metric we use is called *edge correctness* (EC). Since all brain functional networks that we construct for the same subject have the same set of nodes, we can calculate the fraction of edges which were preserved between these networks. More specifically, if we have two BFNs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$, where V is the set of nodes (vertices) and E is the set of edges, the *edge correctness* between them is

$$EC = \frac{|\{(u,v) : (u,v) \in E_1 \text{ and } (u,v) \in E_2\}|}{\max\{|E_1|, |E_2|\}} * 100\% \quad (9)$$

Here $|E_1|$ is the number of elements in E_1. Clearly, if $G_1 = G_2$, than their edge correctness is 100%. On the other hand, low EC means that the networks are different.

To provide a detailed evaluation of *local* structural similarities between two networks, we use a highly constraining measure of local structural similarity between two networks, *Graphlet Degree Distribution* (GDD) agreement, introduced by Przulj [4]. GDD can be understood as a generalization of commonly used graph degree distribution. To briefly explain this similarity measure, we first need to give a few definitions.

A subgraph H of G is *induced,* if its node set is a subset of the node set of G and H contains *all* edges between its nodes that are present in G. A *graphlet* is a small induced subgraph of a network [4]. Because of the small-world nature of the networks we are dealing with, to capture the structural properties of the network, it is sufficient to consider graphlets containing between 2 and 5 nodes. There are 30 possible non-isomorphic graphlets on 2, 3, 4 and 5 nodes. From a topological point of view, it is relevant to distinguish between some nodes in the graphlets. For example, for the 3-node graphlet G_1, the node in the middle of the path is different from its end-nodes (it is of degree 2), whereas the

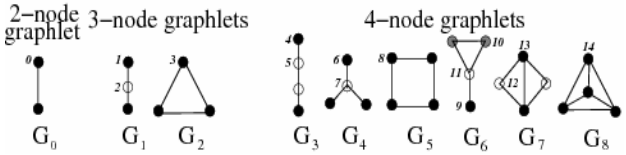

Fig. 3. All possible graphlets on 2 to 4 nodes and their automorphism orbits. Numbers near the nodes correspond to the automorphism orbit's number, which equals to the corresponding component number of GDD vector.

two end-nodes are topologically identical (they are of degree 1 and adjacent to the middle node). This is mathematically defined by the notion of *automorphism orbits* (see [4] for details). Hence, the 3-node path has 2 automorphism orbits, one containing its end-nodes and the other its middle node (see Fig. 3 for examples). There are 73 automorphism orbits for graphlets on 2 to 5 nodes (see [4]). To calculate the GDD of a network corresponding to orbit i, for each of its nodes we calculate how many times it touches orbit i. We do this for each of the 73 orbits. This results in 73-dimensional distribution. Since the only 2-node graphlet is an edge, the first component of this multi-dimensional distribution is the degree distribution. *GDD agreement* of two networks is a similarity measure between the i^{th} graphlet degree distributions over all orbits i for the two networks. If it is close to 1, the distributions are similar and thus the networks have a similar local structure; if it is close to 0, the networks have substantial structural differences [4].

IV. Results and Discussion

A. Small-world property

As described in the Methods, for each subject, each experiment and each stage of the experiment (see Fig. 1), we construct a brain functional network. To determine if these BFNs exhibit small-world properties, we calculate their clustering coefficients and average pathlengths. To calculate λ, γ and σ for each BFN that we analyze, we generate 20 instances of Erdos-Renyi random graphs [10] with the same number of nodes and edges as the BFN and calculate C_{rand} and L_{rand}. We find that all of the BFNs have small world characteristics. The average value of σ did vary between different stages of experiments, but in most cases these variations were not sufficient to distinguish between the stages. While we found that the values of σ depend on the electrode type (e.g. depth vs. grids), they are uniformly greater than 1, which is consistent with previous studies showing that BFNs have small-world properties [1], [2]. Note that $\sigma > 1$ does not guarantee the "small-world" property, since λ (the average pathlength) can be too high. However, in our analysis, for all networks λ was less than 1.5 meaning that they all are undoubtedly "small-world."

B. Differentiation between stages

The experiments that we analyzed are naturally divided into four stages: *fixation, target, delay* and *go* (Fig. 1). It is reasonable to assume that during these stages the brain

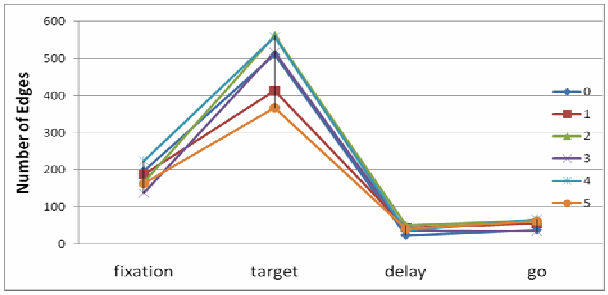

Fig. 4. Subject 005: bilateral depth electrodes. x-axis corresponds to the four different stages of the experiment; y-axis corresponds to the number of edges in the BFN for a given stage. Lines of different colors correspond to the experiments where the peripheral target appeared at one of the six locations.

functions differently and therefore the corresponding brain functional networks should also be different. First, if we examine how the number of edges changes depending on the stage, we can see a clear difference. As Figure 4 shows, the *target* stage has the maximum number of edges while the delay stage has the least number of edges. For the *target* stage this observation holds in 94% of the bilateral depth experiments (subjects S1, S5 and S6). The *delay* stage has the least number of edges in 77% of the bilateral depth experiments. For the grid type of electrodes these dependencies hold only for the subject number 2.

TABLE 2
TARGET AND DELAY STAGES ARE TOPOLOGICALLY THE MOST DISSIMILAR

Electrode type	% of experiments
Bilateral depths	83% (94%)
All experiments	72% (75%)
All except bilateral depths	61% (55%)
Can be expected at random	33% (33%)

First column represents subjects with particular electrode type. Second column shows % of experiments in which, according to GDD agreement (Edge Correctness) "target" and "delay" stages are most dissimilar.

Next, for each subject, we calculate a pairwise edge correctness and GDD agreement between BFNs corresponding to different stages and experiments. Based on these scores, we find that although target and delay stages are contiguous in time, they are the most dissimilar, not only by the number of edges of their BFNs, but also topologically as measured by GDD agreements (Edge Correctness), respectively. This is true for 72% (75%) of experiments we analyzed. For bilateral depth electrodes this rate is 83% (94%), whereas for grid type of electrodes it is 61% (55%). Table 2 summarizes these findings.

The better discrimination between target and delay stages for BFNs constructed from the bilateral depth electrodes may be explained by their superior recording properties. Namely, bilateral depth electrodes are: (i) placed in the cortex and subcortical areas (as opposed to grids which are placed on the cortical surface), (ii) have smaller sensors (1 mm in bilateral depths *vs* 4 mm in grids), which presumably yield betters signal resolution, and (iii) placed bilaterally, which may provide a more complete description of the brain states than single hemisphere grid electrodes.

C. Network Models

To further describe the structure of these networks, we test how well several most commonly used random graph models fit these data. Using Graphcrunch software package [11], we compare each BFN with 30 instances of each of the following random graph models: (i) Erdos-Renyi [10], (ii) Erdos-Renyi with the same degree distribution as in data, (iii) Scale-free Barabasi-Albert [12], (iv) Geometric graphs [5], and (v) Stickiness-index-based model [13]. A *geometric random graph* is constructed as follows. We place nodes uniformly at random in a metric space (in our case, a 3-dimensional Euclidean unit cube); then we choose a parameter ε and connect two nodes by an edge if the Euclidean distance between the nodes is less than ε.

We find that in 67% of the cases, brain functional networks corresponding to the fixation and go stages are best modeled by geometric random graphs (see Fig. 5). In

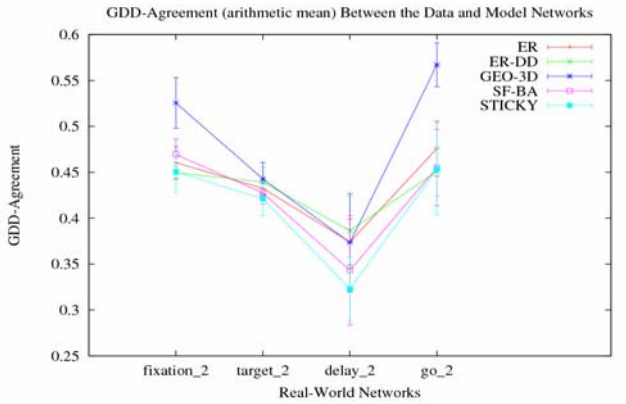

Fig. 5. An example of GDD agreement between brain functional networks corresponding to different stages and model networks.

all other cases, no model fits any stage well; in particular, the popular scale-free model does not fit well these data. Hence, when brain performs specific cognitive tasks, in addition to their small-world properties, brain functional networks also exhibit a geometric structure.

V. Conclusion

Insights into the structure of brain functional networks can potentially improve our understanding of the brain's function. Previously, it was shown, and also confirmed by our experiments, that BFNs have small-world properties [1], [2]. However, the small-world property is not a discriminating factor between BFNs that underlie different cognitive tasks. For this reason, we take the graph-theoretic analysis further and show that even though all BFNs are small-world, it is possible to further distinguish between them if network parameters such as Edge Correctness and GDD Agreement are used. In particular, we were able to show that BFNs corresponding to different cognitive tasks can be separated based these metrics. Furthermore, in terms of a more precise characterization of the network structure, we show that BFNs corresponding to certain cognitive tasks have a geometric graph structure.

References

[1] S. Micheloyannis, E. Pachou, C. J. Stam, M. Vourkas, S. Erimaki, V. Tsirka, (2006) "Using graph theoretical analysis of multi channel EEG to evaluate the neural efficiency hypothesis." *Neuroscience Letters.* 402, pp. 273-277

[2] D.S. Basset, E. Bullmore, (2006) "Small-World Brain Networks." *The Neuroscientist*. 12:6, pp. 512-523

[3] D. J. Watts, S.H. Strogatz, (1998). "Collective dynamics of 'small-world' networks". *Nature* 393: 440–442.

[4] N. Przulj, (2007) "Biological network comparison using graphlet degree distribution." *Bioinformatics.* 23(2), pp. e177-83.

[5] M. Penrose, *"Random Geometric Graphs."* Oxford Studies in Probability, 2003.

[6] Rizzuto D.S., Mamelak A.N., Sutherling W.W., Fineman I., and Andersen R.A. "Spatial selectivity in human ventrolateral prefrontal cortex." Nat. Neurosci., 8:415–417, 2005.

[7] Nenadic Z., Rizzuto D.S., Andersen R.A., and Burdick J.W. "Advances in cognitive neural prosthesis: Recognition of neural data with an information-theoretic objective." In G. Dornhege, J.d.R. Mill´an, T. Hinterberger, D.J. McFarland, and K.-R. M¨uller, editors, Toward Brain Computer Interfacing, chapter 11, pages 175–190. The MIT Press, 2007.

[8] Das K., Rizzuto D.S. and Nenadic Z., "Mental State Estimation for Brain-Computer Interface," IEEE T. Bio-med. Eng. (accepted for publication).

[9] Cover T.M., Thomas, J.A. *"Elements of Information Theory"* New York: Wiley, 1991.

[10] Edos P., Renyi A. "On random graphs." Publicationes Mathematicae 6, 1959.

[11] Milenkovic T., Lai J., Przulj N. (2008) "Graphcrunch: A Tool for Large Network Analyses," *BMC Bioinformatics*. 9:70.

[12] Barabási A.-L., Albert R. (1999). "Emergence of scaling in random networks". Science 286: 509-512.

[13] Przulj N., Higham D. (2006). "Modelling Protein-Protein Interaction Networks via a Stickiness Index," *Journal of the Royal Society Interface.* 3:10, pp. 711 - 716