Proceedings of the 2007 IEEE Symposium on
Foundations of Computational Intelligence (FOCI 2007)

Emergence of Scale-free Spike Flow Graphs in Recurrent Neural Networks

Filip Piękniewski *Member, IEEE*
Faculty of Mathematics and Computer Science
Nicolaus Copernicus University
Toruń, Poland
Email: philip@mat.uni.torun.pl

Tomasz Schreiber
Faculty of Mathematics and Computer Science
Nicolaus Copernicus University
Toruń, Poland
Email: tomeks@mat.uni.torun.pl

***Abstract*— In recent two decades neuroscience and computational intelligence has experienced a large progress, with some new important concepts being developed like spiking [1] and dynamical neurons [2], [3]. Due to increase in computational power available for researchers, many impressive simulations have been carried out and some other ones are yet to come. These developments in CI have occurred simultaneously with some very interesting theoretical and empirical results in random graph theory - the introduction of small-world [4] and scale-free [5], [6] models, and investigation of their properties. Based on results already published by other authors [7], [8] we believe that the edge of these two dynamically growing disciplines might be an interesting field for research. In this paper we introduce a simplified model of spike flow network which in some details resembles a recurrent neural network with stochastic dynamics. We argue that within this setup a scale-free network structure emerges as a natural consequence of model structuring principles and we provide numerical evidence supporting this claim. Further in the paper we investigate a number of interesting properties of this network and discuss some consequences of this result.**

I. Introduction

The concept of scale-free networks has recently emerged as a response to an increasing demand for theoretical tools providing a unified description of a wide variety of complex network topologies displaying the evidence of strong structuring principles co-existent with a considerable degree of randomness, see [9] for a comprehensive survey. A distinctive feature of a *scale-free* network is that the degree distribution of its nodes follows a power law, thus lacking a characteristic scale in the language of statistical mechanics, whence the name. The presence of such power laws has been observed for a broad class of networks, prominent examples including the World Wide Web [10], science collaboration networks [11], citation networks [12], ecological networks [13], linguistic networks [14] as well as cellular metabolic networks [15] and many other ones, see [9]. An important property enjoyed by scale-free networks, although non-distinctive as shared also with other classes of networks including the classical Erdős-Rényi random graphs, is the so-called small-world property stating that the average path length between two nodes in the graph is *small*, which in formal terms usually means logarithmic dependency on the size of the system, see [9]. The small-world property co-present with strong structuring principles has been reported in the topology of certain natural neural networks as well as other networks of biological character, [4]. On the other hand, many artificial recurrent neural networks are built on complete graphs with no scale-free property. In this context we found it natural to ask whether in some cases the scale-free property may emerge for the corresponding *spike flow graph*. While postponing the precise formulation of this question until our model is introduced below, we re-phrase it here rather informally by asking if the scale-free property can be recovered by neglecting or ascribing small usage frequencies to those network connections which are seldom used to transmit neuronal spikes and multiply counting those used often enough, with the resulting graph with weighed edges referred to as the *spike flow graph* in the sequel.

To put the above questions in formal terms we consider a stochastic recurrent neural network consisting of N neurons assuming states labeled by natural numbers $\sigma_i \in \mathbb{N} = \{0, 1, \ldots\},\ i = 1, \ldots, N$, interpreted as *neuronal potentials* below. The network is built on a complete graph in that there is a connection between each pair of neurons $\sigma_i, \sigma_j,\ i \neq j$, carrying a real-valued weight $w_{ij} \in \mathbb{R}$ satisfying the usual symmetry condition $w_{ij} = w_{ji}$. The values of w_{ij} are drawn indepedently from the standard Gaussian distribution $\mathcal{N}(0, 1)$ and are assumed to remain fixed in the course of the network dynamics. A configuration $\bar{\sigma} = (\sigma_i)_{i \leq N}$ of the network is assigned its Hamiltonian given by

$$\mathcal{H}(\bar{\sigma}) := \frac{1}{2} \sum_{i \neq j} w_{ij} |\sigma_i - \sigma_j|. \tag{1}$$

The dynamics of the network is defined as follows: at each step we randomly choose a pair of neurons $(\sigma_i, \sigma_j),\ i \neq j$, and denote by $\bar{\sigma}^*$ the network configuration resulting from the original configuration $\bar{\sigma}$ by decreasing σ_i by one and increasing σ_j by one, that is to say by *letting a unit of potential transfer from σ_i to σ_j*, whenever $\sigma_i > 0$. Next, if $\mathcal{H}(\bar{\sigma}^*) \leq \mathcal{H}(\bar{\sigma})$ we accept $\bar{\sigma}^*$ as the new configuration of the network whereas if $\mathcal{H}(\bar{\sigma}*) > \mathcal{H}(\bar{\sigma})$ we accept the new configuration $\bar{\sigma}^*$ with probability $\exp(-\beta[\mathcal{H}(\bar{\sigma}^*) - \mathcal{H}(\bar{\sigma})]),\ \beta > 0$, and reject it keeping the original configuration $\bar{\sigma}$ otherwise, with $\beta > 0$ standing for an extra parameter of the dynamics, in the sequel referred to as the inverse temperature conforming to the

usual language of statistical mechanics. Below we shall write $\pi_{i\to j}(\bar\sigma)$ for the above specified probability that for chosen i and j the potential transfer proposal from σ_i to σ_j gets accepted in configuration $\bar\sigma$. Observe that the sum $\sum_i \sigma_i$ of neuronal potentials is preserved by the dynamics and that, in the course of dynamics with some initial configuration $\bar\sigma^0$, any other $\bar\sigma$ with $\sum_i \sigma_i^0 = \sum_i \sigma_i$ is eventually reached with positive probability. Consequently, upon standard verification of the usual detailed balance conditions, we readily see that the stationary states of the above dynamics are the distributions

$$\mathbb{P}_n(\bar\sigma) = \begin{cases} \frac{\exp(-\beta\mathcal{H}(\bar\sigma))}{\sum_{\bar\sigma',\ \sum_i \sigma'_i = n} \exp(-\beta\mathcal{H}(\bar\sigma'))}, & \text{if } \sum_i \sigma_i = n, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

and their convex combinations. In particular, our model bears some resemblance to the usual stochastic Boltzmann machines [16], with the weights w_{ij} indicating the extent to which the system favours the agreement (for positive w_{ij}) or disagreement (for negative w_{ij}) of the neuronal states σ_i and σ_j. There are important differences though, one of them being the unbounded state space, the other one that precisely two neurons are affected in each update with clearly determined source and destination of the potential flow. This was in fact one of the features aimed at when constructing the model above. For the network dynamics running during a period $[0, T]$ we are now in a position to define the *spike flow graph* to be the complete directed graph with vertices corresponding to the neurons $\sigma_i,\ i = 1, \ldots, N$ and whose edges carry numbers (edge multiplicities) $F_{i\to j}$ indicating how many times in the course of the dynamics the potential flow occurred from σ_i to σ_j. Note that for T large enough to ensure that the system gets close to its equilibrium and that each edge is processed a statistically significant number of times we have

$$F_{i\to j}/T \approx \frac{1}{N(N-1)} \mathbb{E}_n \pi_{i\to j}(\bar\sigma) \tag{3}$$

with $\mathbb{E}_n$ standing for the expectation operator under $\mathbb{P}_n$ as given by (2) and where n is the total potential preserved in the course of the dynamics. In particular, the randomness of $F_{i\to j}/T$ vanishes for large T where it gets close to a deterministic function of the network and of n, the only remaining randomness coming from the weights w_{ij}. The out-degree of a neuron σ_i is now defined as $d_{\text{out}}(i) := \sum_j F_{i\to j}$ and, likewise, the in-degree $d_{\text{in}}(j) := \sum_i F_{i\to j}$. The main question considered in this paper is whether the so-defined spike flow graph is scale-free in that its in- and out-degree distributions follow power laws, that is to say $\mathbb{P}(d_{in/out}(i) \approx x) \sim c_{in/out} x^{-\gamma_{in/out}}$ for a randomly picked node i. This conjecture is strongly motivated by the fact that the dynamics of our system seems to exhibit a version of preferential attachment mechanism, as discussed in Sections VII and VIII of [9], in the sense that a node with high out-degree d_{out} is very likely to often reach high potentials in the course of the dynamics (because it can afford to spend a lot of potential transfering it to other nodes) hence if a next potential transfer is observed in the system to a given destination node, it is more likely to come from a node of high out-degree than from one with low out-degree which is in its turn rather unlikely to store high potential. Likewise, it is natural to expect that large in-degree of a node should usually be associated with frequent reaching of high potentials and consequently with large out-degree as well. Our numerical experiments seem to strongly confirm the above argument, see Figures 3, 4, 6 and 5 below, and it is known that preferential attachment very often yields the scale-free property, see [9]. The fact that the number of nodes is fixed in our network as opposed to the classical Albert-Barabási network growth model does not seem to lead to a problem because, although the number of nodes does indeed stay fixed, the way in which the spike flow graph is constructed amounts to allowing for multiple edges and unlimited addition of new edges.

In the sections below we present simulations and numerical evidence supporting the above scale-freeness conjecture for spike flow graphs.

II. Simulation setup

The simulations were carried out for different system sizes varying from n=100 to 2100 nodes (simulating bigger networks takes considerably more time: since every step requires recomputing the energy of the system, assuming that the required number of steps is $\mathcal{O}(\text{number of edges}) = \mathcal{O}(n^2)$[1] implies that the simulation complexity goes up to $\mathcal{O}(n^3)$) although instances of about 2000 nodes produce graphs whose properties are quite consistent with those of about 500 or less, and hence simulating bigger samples does not seem to provide any qualitatively different results. The precise length of a simulation run was set to $10 \cdot |E|$ where $|E|$ stands for the number of edges, i.e. $|E| = |V|^2$. Weights were randomly sampled from standard normal distribution, initial value of potential in every node was fixed to a small natural number (in most cases 5, but in general changing this number within a reasonable range did not affect the simulation too much[2]). We fixed $\beta = 1$ (inverse temperature) which corresponds to predominantly energy-driven transitions and rather negligible entropic factor, and we look forward to studying this model for other (higher) temperatures in the future. In our present setup the system seems to converge rapidly to a relative equilibrium whereupon the potential transfers and energy changes become considerably less frequent. This results in a rather sparse spike flow graph obtained in simulation and hence the chosen number of simulation steps can be claimed to yield statistically significant outcomes for medium and large in- and out-degrees only whereas for the nodes of very small degree (0,1,2) a positive bias (oversaturation) is likely to be present due to the insufficient number of iterations (see e.g. Figure 1 below). This is not a problem from the viewpoint of our purposes though, since anyway the crucial information about the scale-

[1] It is necessary to give each edge a chance to be chosen at least a couple of times in the course of the simulation.

[2] Some subtle impact can be observed though as we discuss in Section III below.

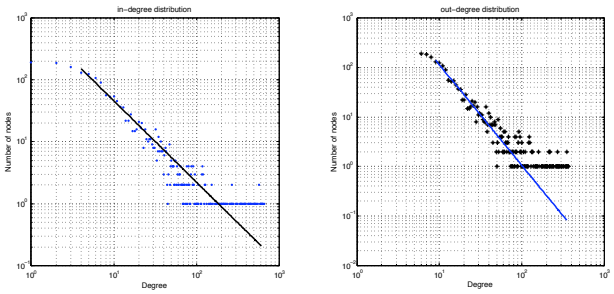

Fig. 1. Log-plot of the in-degree distribution (left) and out-degree distribution (right) for the original spike flow network obtained from the simulation (dropping multiple edges affects this image only slightly). Clearly these distributions follow a power law (linear fit marked).

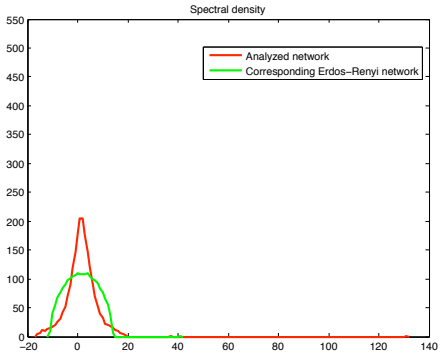

Fig. 2. Spectral density for the analyzed network (after dropping edge directions) and for the corresponding Erdős-Rényi random graph with the same edge density (exhibiting the usual semicircular law) for comparison. This triangle-like spectral density has been found characteristic for scale-free networks, see Section VII.D.4 in [9].

free nature of the spike flow graph is extracted from the properties of higher order nodes.

In Section III below, apart from verifying the power laws for in- and out-degrees of the nodes and analyzing the dependencies between the in- an out-degrees and mean potentials, whose nature seems to be fairly independent of the simulation length, we also study the properties of the spike flow graph which clearly do depend on the particular length of simulation run: these include the average path length between two nodes (decreases as the spike flow graph saturates), mean clustering coefficient (increases as the graph saturates) etc. This does make sense as aimed at showing that the onset of the small-world and high clustering properties is observed even in relatively early stages of simulation, thus providing a further support of our scale-freeness conjecture for spike flow graphs. Note also that when determining below certain characteristics of the spike flow graph (average path length, clustering coefficient, spectral properties) we drop the multiplicities $F_{i\to j}$ ascribed to edges, simply by declaring an edge $i \to j$ present in the graph if $F_{i\to j} > 0$ and absent otherwise.

III. Simulation results

The most interesting plot obtained in our simulations, on which the main conjecture stated in this paper depends, is the degree distribution. The spike flow network arising in our setting is directed, and so the in-degree and out-degree distributions are depicted separately (Fig. 1). The obtained logarithmic plot is close to linear for higher order nodes, thus confirming the power law hypothesis $\mathbb{P}(d_{in/out}(i) \approx x) \sim c_{in/out}x^{-\gamma_{in/out}}$ for a randomly picked node i. Simple linear fit on the log-plot was used to estimate the exponents in these power laws. It came out that in-degree exponent $\gamma_{in} \approx 1.3$, whereas out-degree exponent $\gamma_{out} \approx 2.0$. These approximations are rather rough, since the tail of the distribution is scattered thus biasing the outcome exponents, but one can say with a high degree of confidence that $\gamma_{in} < \gamma_{out}$ and that $\gamma_{in} < 2$. Scale-free networks with exponents within range of $(1, 2)$ were investigated in [17]. It is worth mentioning that in all instances of the simulation the output spike flow network had only one connected component, which was not obvious *a priori* since the obtained networks are rather sparse and the dynamics does not immediately imply that a single connected component should rapidly emerge (although the theory guarantees that it does happen at some point since, with the simulation length tending to ∞, the spike flow graph becomes eventually fully connected).

Another interesting tool for exploiting the properties of a graph is its spectrum, defined as the set of eigenvalues of its adjacency matrix or Laplacian matrix. The more advanced spectral graph theory is far beyond the scope of this article (see e.g. [18] for comprehensive introduction), we only mention here the spectral density which is, very roughly speaking, the histogram of eigenvalues (taking their multiplicities into account). It is well defined for undirected graphs (symmetric matrices) so, in the case of the investigated model, multiple edges and their directions have been neglected. Figure 2 depicts the spectral density of the considered spike flow network and the spectral density of the corresponding random Erdős-Rényi graph (the same edge density) for comparison. The latter exhibits semicircular (Wigner's) law [19], [20], [21], whereas triangle shaped spectral density (red plot on Fig. 2) is expected for scale-free networks (see [22]).

Next interesting feature of these networks is the dependency between in and out degree distributions. In the case of the considered network there is a very clear, nearly linear dependency (Fig. 3), which is only slightly disturbed after dropping multiple edges[3] (Fig. 4). This gives some insight into the structure of spike flow networks and suggests that nodes with high in-degree have usually high out-degree as well, as conjectured in the discussion of the preferential attachment phenomenon preceding Section II. This might have been expected, since as we argued in the introduction above, nodes with high in-degree should possess more potential, and therefore should be more likely to give that potential away, yet the striking

[3]We have noticed that in general the presence or absence of multiple edges does not seem to considerably affect the studied qualitative properties of spike flow graphs.

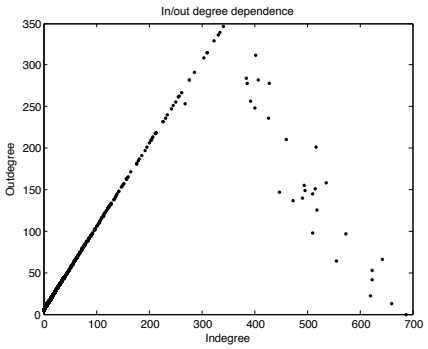

Fig. 3. Dependency between in- and out-degree in the original network obtained from the model (with multiple edges between pairs of nodes allowed). The relation is almost linear, with a constant below 1. This explains a group of outliers on the top of the graph - the sum of outgoing and incoming edges in a graph must be equal.

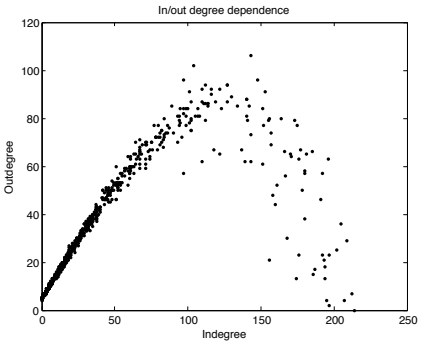

Fig. 4. Dependency between in- and out- degree after dropping multiple edges.

regularity of this dependency was a pleasant surprise and gives us motivation to some further analytic research. Another support of the conjecture that high potential translates into high in/out-degrees is given by figures 5 and 6. Depicted dependencies are again very regular and seem to follow power laws, which is of its own intrinsic interest and gives some further motivation for more detailed investigations.

The clustering coefficient (see Appendix) reveals the typical local structure of a graph. For the considered networks this coefficient is approximately a magnitude bigger than for the corresponding Erdős-Rényi random graphs (see Fig 9), therefore, taking into account the relative sparseness of the spike flow networks, they might be described as well clustered. Even more interesting was the dependency between node degree and average clustering coefficient (Fig. 7). This dependency is constant for random and most scale-free networks (see box 2 in [23]), follows a power law with a negative exponent for hierarchical networks whereas in our case it is an increasing function. This is curious as it suggests that vertices with high degree in spike flow graphs have neighbors that "contact each other" more frequently than neighbors of vertices with low degree.

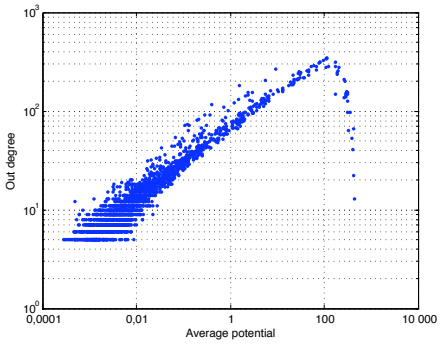

Fig. 5. Dependency between the average potential during the simulation and the out-degree of a vertex (log-plot). There is a clearly marked power-like dependency.

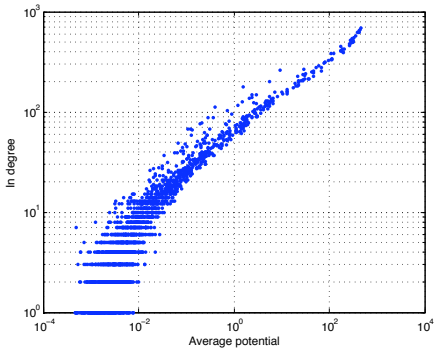

Fig. 6. Dependency between the average potential and the in-degree of a vertex (log-plot). Again this dependency is quite well pronounced and power-like with some unavoidable exceptions on the right side of the plot.

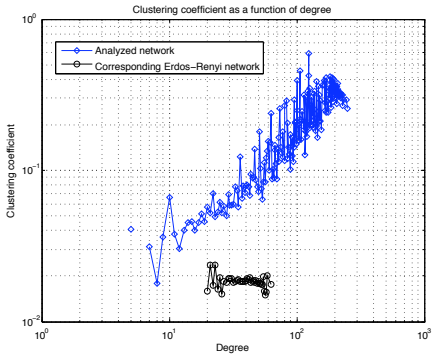

Fig. 7. This figure displays a plot of the clustering coefficient against vertex degree. The blue plot depicts the dependency for the analyzed network (after dropping edge directions) and the black one shows corresponding relation in a random Erdős-Rényi graph with the same connectivity density for comparison.

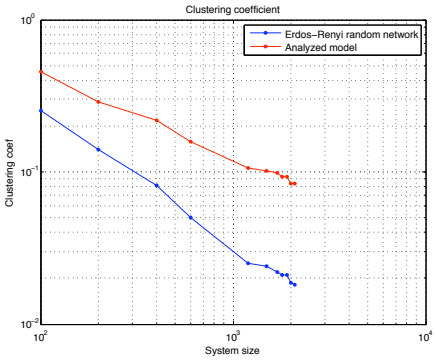

Fig. 8. Clustering coefficient of the network (after dropping multiple edges and directions) as a function of the system size. Note that the clustering coefficient of spike flow networks obtained from our model is considerably larger than for the corresponding random networks. The general decreasing tendency can be explained by growing sparseness of larger spike flow graphs.

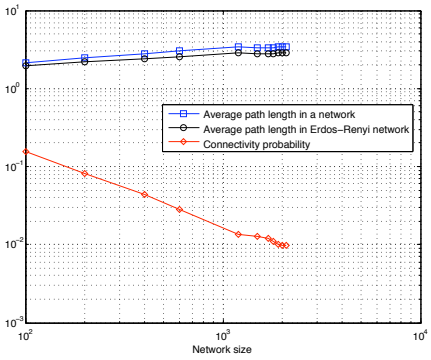

Fig. 9. Average path length as a function of the system size. One can notice that the average path length in the analyzed network only slightly deviates from the average path length in Erdős-Rényi random graph with the same connectivity density. This shows that the considered spike flow networks have a small-world character. The red line displays the connectivity density, which shows that the considered graph becomes more and more sparse as the systems grows, still however having one connected component only.

As already mentioned before, the results coming from different network sizes were quite consistent. This is displayed in figures 8 and 9 which show the dependency of clustering coefficient, average path length and connectivity density on the system size. Figure 9 shows that the small-world property for the spike flow graph emerges quickly in the course of the simulation: indeed, in the examined system the average path length between two vertices is only slightly larger than in the corresponding Erdős-Rényi graph. Recall that the connectivity density is simply the number of edges (after dropping multiple edges) divided by the number of edges in the corresponding fully connected graph (this is exactly the probability parameter in the Erdős-Rényi model). The monotone decrease of this quantity in Figure 9 is due to the fixed amount of potential available for every node. We plan to carry out some further computations in our future work, varying the amount of potential to achieve and explore different relations of connectivity density and size (assuming the number of steps in simulation is proportional to the number of edges).

IV. Conclusion

In this paper we show that in certain recurrent neural networks the scale-free structure emerges spontaneously for the corresponding graphs constituted by their most often used connections. Even more, evoking the notion of preferential attachment apparently present in some such systems, we argue that this may happen as a natural consequence of the usual stochastic dynamics. To achieve these goals we construct a particular potential flow neural network – while sharing many features with usual Boltzmann machines our model allows for storing arbitrarily large potential in a single neuron, thus endowing it with a kind of 'memory' which seems crucial for the emergence of preferential attachment-like mechanism. We believe that similar scale-free behavior for (appropriately defined) spike flow graphs should be present in many other (more usual) types of recurrent networks as well, yet we expect that often it can arise only at the level of appropriately chosen neuronal groups[4] rather than single neurons unless the architecture of a single unit is complicated enough to exhibit some memory/knowledge of its prior neural activity. These and related issues are the subject of our present work in progress.

Appendix

A. Spectral density

Any graph G with n nodes can be represented by its adjacency matrix A of $n \times n$ elements A_{ij} whose values are $A_{ij} = 1$ if there is an edge from node i to j and $A_{ij} = 0$ otherwise. For undirected graphs adjacency matrices are symmetric and therefore have real eigenvalues. The set of eigenvalues of the adjacency matrix A corresponding to a graph G is called the spectrum of the graph G. Eigenvalues contain a lot of interesting information (for example the multiplicity of 0 as eigenvalue coincides with the number of connected components etc.). To get a general idea of spectral properties of a graph it is useful to define its spectral density:

$$p(\lambda) = \frac{1}{n}\sum_{j=1}^{n}\delta(\lambda - \lambda_j) \tag{4}$$

which may approach a continuous function as $n \to \infty$. Spectral density has interesting links with graphs properties, for example $k-th$ moment of spectral density is the number of paths of length k that return to the their origin (possibly visiting other nodes multiple times). Alternatively (but not equivalently) graph spectra can be defined as sets of eigenvalues of (normalized) Laplacian matrix. Laplacian matrix is constructed as follows:

$$L = D - A \tag{5}$$

[4]This needs to be emphasized, since neural network of a worm C. elegans proved not to follow a power law on the level of single neurons (axons), see [24], [25] for details.

where A is the adjacency matrix and D is a diagonal matrix with $(i,i)-th$ entry containing the degree of node i. Matrix L can be normalized by

$$\mathcal{L} = D^{-1/2} L D^{-1/2} \tag{6}$$

so that diagonal values of $\mathcal{L}$ are all equal 1. This definition relates better to graph invariants (see [18] for more details).

B. Clustering coefficient

Clustering coefficient was introduced in [4] to reveal general clustering properties of a graph. For a node i its clustering coefficient is defined as follows:

$$C_i = \frac{2|\{e_{j,k}\}|}{d_i(d_i-1)} \tag{7}$$

where nodes j and k are adjacent to i, and d_i denotes degree of i. Roughly speaking it is the number of connections between the neighbors of the node i divided by the number of all possible connections between the neighbors of i. Clustering coefficient can be alternatively defined as

$$C_i = \frac{2\lambda_G(i)}{d_i(d_i-1)} \tag{8}$$

where λ_G is the number of triangles on vertex i. This definition suggests a quite efficient algorithm of computing clustering coefficient, since the number of triangles coming from a node (assuming there are no loops of length one at vertices) is simply the corresponding entry on the diagonal of A^3, where A is the adjacency matrix. We assume that $C_i = 0$ if $d_i = 0$ and $C_i = 1$ if $d_i = 1$. The average clustering coefficient of the whole graph is simply the average of coefficients computed for every node and hence it coincides with the normalized trace of A^3. Other interesting features of a graph include the information on how clustering coefficient is distributed over vertices of different degrees and, likewise, how the degree is distributed over nodes having clustering coefficient within a fixed range.

ACKNOWLEDGMENT

The authors would like to thank the Marshall of Kujawsko-Pomorskie county in Poland for financial support. Tomasz Schreiber gratefully acknowledges the support from the Polish Minister of Scientific Research and Information Technology grant 1 P03A 018 28 (2005-2007).

REFERENCES

[1] S. M. Bohte, "Spiking neural networks," Ph.D. dissertation, University of Leiden, Leiden, 2003.

[2] E. M. Izhikevich, *Dynamical systems in Neuroscience: The Geometry of Excitability and Bursting*. Boston: MIT Press., 2006.

[3] ——, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks*, no. 14, pp. 1569– 1572, 2003. [Online]. Available: http://www.nsi.edu/users/izhikevich/publications/spikes.pdf

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

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

[6] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, no. 74, pp. 47–97, January 2002.

[7] N.Davey, B. Christianson, and R. Adams, "High capacity associative memories and small world networks," in *Proc. IEEE International Joint Conference on Neural Networks (IJCNN 05)*, Budapest, Hungary, July 2004, pp. 159–165.

[8] H. F. Kwok, P. Jurica, A. Raffone, and C. van Leeuwen, "Robust emergence of small-world structure in networks of spiking neurons," *Cognitive Neurodynamics*, december 2006. [Online]. Available: http://www.springerlink.com/content/57241p34x7317641/fulltext.pdf

[9] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, p. 47, 2002. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0106096

[10] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the world-wide web," *Science*, vol. 401, pp. 130–131, Septmeber 1999.

[11] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, "Evolution of the social network of scientific collaborations," *Physica A*, vol. 311, no. 4, pp. 590–614, 2002.

[12] S. Redner, "How popular is your paper? an empirical study of the citation distribution," *European Physical Journal B*, vol. 4, no. 2, pp. 131–134, 1998.

[13] J. M. Montoya and R. V. S. V., "Small world patterns in food webs," *Journal of Theoretical Biology*, vol. 214, no. 3, pp. 405–412, February 7 2002.

[14] R. F. i Cancho and R. V. Solé, "The small-world of human language," *Proceedings of the Royal Society of London B*, vol. 268, no. 1482, pp. 2261–2265, November 7 2001.

[15] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, "The large-scale organization of metabolic networks," *Nature*, vol. 407, no. 6804, pp. 651–653, October 5 2000.

[16] E. Aarts and J. Korst, *Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing*. New York, NY, USA: John Wiley & Sons, Inc., 1989.

[17] H. Seyed-allaei, G. Bianconi, and M. Marsili, "Scale-free networks with an exponent less than two," *Physical Review E*, vol. 73, p. 5, 2006. [Online]. Available: doi:10.1103/PhysRevE.73.046113

[18] F. R. K. Chung, *Spectral graph theory (CBMS Regional Conference Series in Mathematics, No. 92)*, ser. Cbms Regional Conference Series in Mathematics. American Mathematical Society, 1997.

[19] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Annals of Mathematics*, vol. 62, no. 3, pp. 548–564, November 1955. [Online]. Available: http://www.jstor.org/view/0003486x/di961748/96p0020o/0

[20] ——, "Characteristic vectors of bordered matrices with infinite dimensions ii," *Annals of Mathematics*, vol. 65, no. 2, pp. 203–207, March 1957. [Online]. Available: http://www.jstor.org/view/0003486x/di961756/96p00014/0

[21] ——, "On the distribution of the roots of certain symmetric matrices," *Annals of Mathematics*, vol. 67, no. 2, pp. 325–327, March 1958. [Online]. Available: http://www.jstor.org/view/0003486x/di961762/96p00057/0

[22] I. J. Farkas, I. Derenyi, A.-L. Barabási, and T. Vicsek, "Spectra of "real-world" graphs: Beyond the semi-circle law," *Physical Review E*, vol. 64, p. 1, 2001. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0102335

[23] A.-L. Barabási and Z. N. Oltvai, "Network biology: Understanding the cell's functional organization," *Nat Rev Genet*, vol. 5, no. 2, pp. 101–113, February 2004. [Online]. Available: http://dx.doi.org/10.1038/nrg1272

[24] L. A. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of small-world networks." *Proc Natl Acad Sci U S A*, vol. 97, no. 21, pp. 11 149–11 152, October 2000. [Online]. Available: http://dx.doi.org/10.1073/pnas.200327197

[25] C. Koch and G. Laurent, "Complexity and the Nervous System," *Science*, vol. 284, no. 5411, pp. 96–98, 1999. [Online]. Available: http://www.sciencemag.org/cgi/content/abstract/284/5411/96