



ELSEVIER

Physica A 298 (2001) 553–561

PHYSICA A

www.elsevier.com/locate/physa

Geometrical structure of the neuronal network of *Caenorhabditis elegans*

Satoru Morita^{a,*}, Ken-ichi Oshio^b, Yuko Osana^c, Yasuhiro Funabashi^c,
Kotaro Oka^c, Kiyoshi Kawamura^c

^aDepartment of Systems Engineering, Shizuoka University, 3-5-1 Hamamatsu 432-8561, Japan

^bDepartment of Physiology, School of Medicine, Kinki University, Osaka 589-8511, Japan

^cFaculty of Science and Technology, Keio University, Yokohama 223-8522, Japan

Received 27 November 2000

Abstract

The neuronal network of the soil nematode *Caenorhabditis elegans* (*C. elegans*), which is a good prototype for biological studies, is investigated. Here, the neuronal network is simplified as a graph. We use three indicators to characterize the graph; vertex degree, generalized eccentricity (GE), and complete subgraphs. The graph has the central part and the strong clustering structure. We present a simple model, which shows that the neuronal network has a high-dimensional geometrical structure. © 2001 Elsevier Science B.V. All rights reserved.

PACS: 02.50.-r; 05.50.+q; 05.40.-1; 87.18.Sn

Keywords: Neuronal network; *C. elegans*; Graph theory; Geometrical structure; Small world

1. Introduction

Complexity of nervous systems is reflected in the complexity of their structural makeup [1]. The purpose of this paper is to characterize the structure of the neuronal networks. We deal with the soil nematode *Caenorhabditis elegans* (*C. elegans*), because all the connections of its neuronal network are well known (some data bases are available [2–4]). *C. elegans* is a small worm with a relatively simple nervous system. There are only 302 neurons in the adult hermaphrodites. The nervous system is separated into two units as follows. First, the pharyngeal nervous system is composed of 20 cells. The pharyngeal nervous system controls the rhythmical contraction of the

* Corresponding author. Tel.: +81-53-478-1226; fax: +81-53-478-1226.

E-mail address: morita@sys.eng.shizuoka.ac.jp (S. Morita).

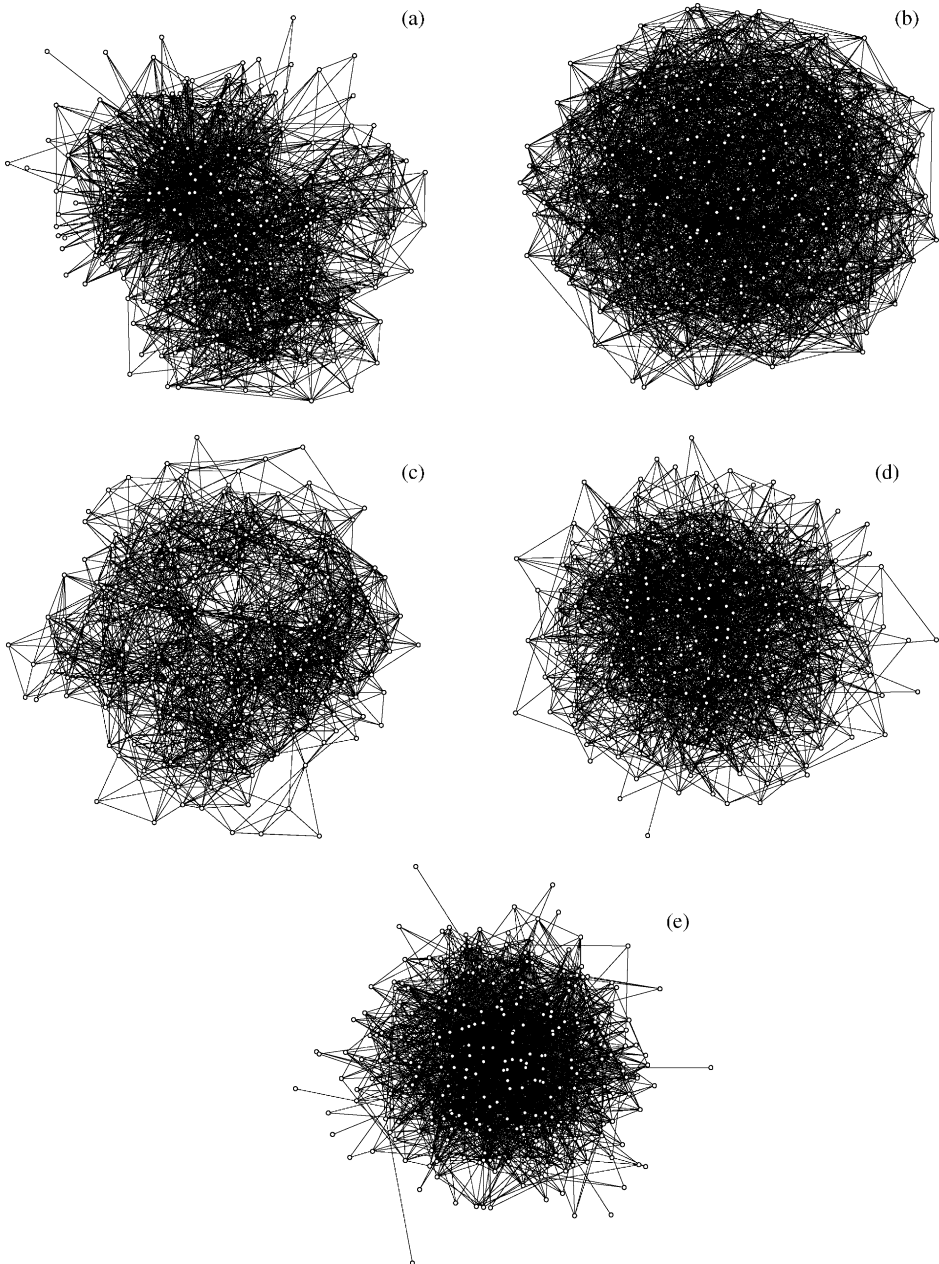


Fig. 1. Visual illustration of the neuronal network and the model networks: (a) G_{cc} , (b) random graph, (c) G_d for $d=4$, (d) $d=8$ and (e) $d=12$. Every network contains 269 vertices (open circles) and 2268 edges (solid lines). The network visualization is done using the pajek program [21].

pharynx to suck bacteria. The pharyngeal nervous system is nearly completely isolated from the rest. The synaptic connections among the pharyngeal neurons are well documented by Albertson and Thomson [5]. Second, the somatic nervous system consists of the rest neurons. The long processes from the somatic neurons construct bundles; the nerve ring, the ventral cord, etc. The synaptic connections among the 282 somatic neurons are studied by White et al. [6].

In this paper, the real neuronal network is simplified as a graph, which is a set of vertices connected to each other by edges. The internal structure of each cell is ignored. Thus, a vertex stands for a neuron. There are two types of connections; chemical synapse and gap junction. While the former is polarized, the latter is not. Furthermore, each pair of neurons often has more than one connection. For simplicity, type, direction and multiplicity of connection are not taken into account. This simplification is effectual for study of topological feature in diverse networks [7,8]. Thus, if there exists at least one synaptic connection between a pair of vertices, then they are linked by an edge. We mainly focus on the somatic nervous system. There are 282 vertices and 2268 edges. We use three indicators to characterize the graph; (1) vertex degree, (2) generalized eccentricity (GE), and (3) complete subgraphs [see Fig. 1].

There were some mathematical studies of the neuronal network. Some researchers have studied the neuronal network in terms of vertex degree [2,11,12]. However, only considering vertex degree is not sufficient to study the global structure, as we will show. Watts and Strogaz investigated three networks (*C. elegans*, collaboration of film actors and power grid) with path length and clustering coefficient [7]. The path length corresponds to the total average of GE, and the clustering coefficient relates to complete subgraphs of degree three. However, taking into account the number of complete subgraphs of high degree, we will show that their model is not appropriate for the neuronal network of *C. elegans*. In this paper, a simple model with geometrical structure is presented.

2. Characteristic of the neuronal network

First, we use the concept of vertex degree. The degree of a vertex v represents the number of edges meeting at v . There are 13 vertices whose degrees are 0 [17]. The connection of the 13 neurons was not reported in the experimental data by White et al. [6]. We deal with the graph G_{ce} which is obtained by the removal of these 13 isolated vertices. There are 269 vertices and 2268 edges. The solid line in Fig. 2(a) shows the degree sequence, which is the list of the degrees of the vertices which are ordered in a non-increasing way, for G_{ce} . To characterize G_{ce} , for a start, we consider the normal random (NR) graph with 269 vertices and 2268 edges [10]. Here, two ends of the edge are selected randomly with the equal probability. We generate 100 samples. The generated graphs are not always connected. We leave out disconnected ones. The dotted line in Fig. 2(a) shows the degree sequence for the NR graph, which is given by the Poisson distribution. The degrees of G_{ce} are more dispersed than that of the

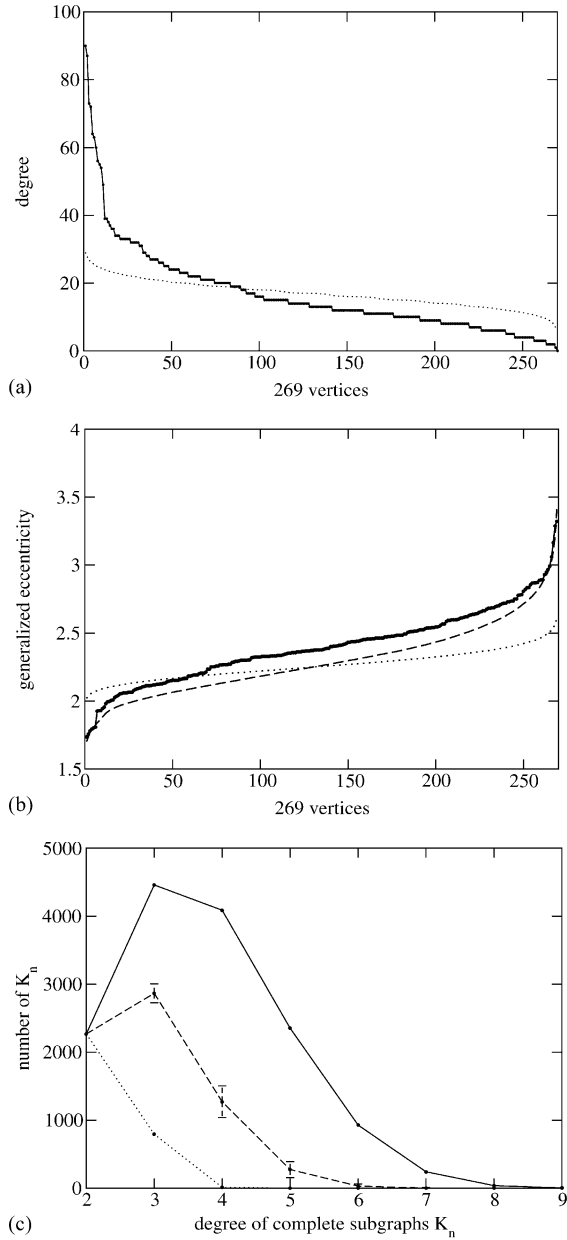


Fig. 2. Three indicators to characterize the neuronal network of *C. elegans*: (a) the degree sequence, (b) the GE sequence, and (c) the number of complete subgraphs. The solid, dotted and broken lines are for *C. elegans*, the normal random (NR) graphs and the biased random (BR) graphs, respectively. The lines for the two random graphs are obtained by the average over 100 samples. For the NR graph, the root-mean-squares (RMS) are nearly equal to the line width. For the BR graph, RMS of the GE are smaller than 0.05 except a few vertices which have large GE, and RMS of the number of complete subgraphs are shown by the error bars.

Table 1

(a) The vertices whose degrees are larger than 40, (b) 12 vertices whose generalized eccentricity (GE) is the smallest. The vertices are discriminated by names of the corresponding neurons. Almost *C. elegans* neurons are designated by a three-alphabet code with additional letters added to denote positional differences, e.g. L for left and R for right. The neurons in (a) and (b) almost coincide. They comprise inter-neurons which is biologically important. Chalfie et al. [19] studied the role of the inter-neurons by using laser ablation. The AVD* neurons are essential for backward movement to anterior touch stimuli; the PVC* neurons are essential for forward movement to posterior touch stimuli; the AVA* and AVB* neurons are necessary for normal coordinated movement. Moreover, AVE* neurons are like AVD*, but their outputs are restricted to anterior cord due to their shorter processes [3]

(a) “Hub” neurons						
Name	AVAL	AVAR	AVBR	AVBL	AVER	AVEL
Degree	90	87	73	72	64	63
Name	AVDR	PVCL	AVDL	PVCR	DVA	
Degree	60	56	55	54	49	
(b) “Center” neurons						
Name	AVAL	AVAR	AVBR	AVER	AVEL	AVBL
GE	1.74	1.76	1.78	1.79	1.80	1.81
Name	DVA	AVDR	AVDL	RIBL	PVCL	PVCR
Degree	1.92	1.93	1.93	1.94	1.96	1.98

NR graph. Especially, G_{ce} has some vertices with extremely large degrees. Table 1(a) shows the neuron names and degrees of the vertices whose degrees are larger than 40. More than 70% of the rest vertices in G_{ce} are linked directly to these vertices. Thus, this group is “hub”.

Achacoso and Yamamoto also indicated that the connections of the neuronal network of *C. elegans* is not uniformly weighted and some neurons seem to be more important than others [2]. Amaral et al. [11] reported that the tail of the distribution of the degrees is approximated by exponential decay, i.e., there is a single scale for connectivity. Amaral et al. [11] took into account only chemical synapse and separated incoming and outgoing degrees. In contrast, here, we use the total degree including three types of connections. In this case, the exponential decay is not seen. It is because there is a correlation among the three connections. Albert et al. [12–15] reported that the distribution of the degrees decays as a power-law in the network of world-wide web. The power-law decay is also seen in the network of citations of scientific papers [16]. However, for the degree of G_{ce} , a clear power-law is not seen.

Second, we consider distance on the graph. The distance refers to the minimum path length between a pair of vertices. Note that this distance is not the real distance between a pair of neurons in the body. Every pair of vertices in G_{ce} are joined by a path, i.e., G_{ce} is connected. The maximum distance is five and the average distance is 2.40. All vertices are connected with short length to one another, as was reported as small world in Refs. [7–9]. To study the feature of the distance in detail, we define the generalized eccentricity (GE) as follows. The GE of a vertex v is the average of the distances from v to other vertices. In graph theory, the eccentricity of a vertex v in a graph is defined as the distance from v to a vertex farthest from v [18]. While the

eccentricity is a natural number, the GE is a real number. The solid line in Fig. 2(b) shows the GE sequence, which is the list of the GE of the vertices in non-decreasing way, for G_{ce} . The dotted line in Fig. 2(b) is for the NR graph. The GE of G_{ce} is more dispersed than that of the NR graph. In G_{ce} , there are some vertices with relatively small GE. These vertices seem to be located at the center of the graph. Table 1(b) shows the neuron names and the GE of 12 vertices whose GE is the smallest. The “center” group almost coincides with the “hub” group. This result indicates that these neurons play an important role. Indeed, they are the inter-neurons which are essential for locomotion [19].

To study the cause of the result for GE, we introduce the biased random (BR) graph, which preserves the property of the degree sequence of G_{ce} . Here, two ends of the edge are chosen randomly with the probability which is given by the ratio of the empirical degrees. We generate 100 samples, where we leave out disconnected ones, again. For the BR graph (the broken line in Fig. 2(b)), the distribution of GE is similar to that for G_{ce} . Moreover, we study the relation between 10 vertices whose degrees are the largest and 10 vertices whose GE is smallest in the case of the BR graph. The number of the vertices which they share, averages 9.55. Therefore, the result for the GE of G_{ce} is mainly due to the large variety of the vertex degrees of it.

Third, we use the concept of complete subgraph. A complete graph is a graph in which every two vertices are connected directly. Although G_{ce} is not a complete graph, it includes a number of complete graphs as subgraphs. We count the number of complete subgraphs K_n with n vertices (shown by the solid line in Fig. 2(c)). For example, the numbers of K_2 and K_3 are the numbers of the edges and the triangles, respectively. The dotted line in Fig. 2(c) is for the NR graph. It is clear that G_{ce} has a lot of complete subgraphs K_n of high degrees n . This result indicates that G_{ce} has a clustering structure, i.e., there is a strong positive correlation among the connections. Some readers may think that this structure is also just a result of the large variety of the vertex degree as reported above. To test this, we use the BR graph again. The number of complete subgraphs of the BR graph is shown by the broken line in Fig. 2(c). The BR graph does not show so strong clustering as G_{ce} has. Therefore, this clustering structure is not due to the large variety of the vertex degree.

3. Simple model

So far, we have seen that the nervous system of *C. elegans* has the central group of the neurons and the strong clustering structure. Watts and Strogatz [7] proposed a small world model for biological and social networks with the clustering structure. The Watts and Strogatz (WS) model is made by random rewiring from a low-dimensional regular lattice. Thus, the WS model is the mixture of the low-dimensional lattice and the NR graph. Consequently, it is obvious that the WS model has smaller number of complete subgraphs of high degrees than the low-dimensional lattice. Thus, the WS small world model is not appropriate for the neuronal network of *C. elegans*.

We generate another model as follows: We consider the system with 269 vertices. Let us suppose that each vertex v has d characteristic indices, i.e., $a_i(v)$ for $i = 1, 2, \dots, d$. The values of the indices are given by the uniform distribution in $[0, 1]$. Then, we connect a pair of vertices (v, v') , if the difference of all pairs of the indices are smaller than a threshold value, i.e., $|a_i(v) - a_i(v')| < a_c$ for all i . The threshold value a_c is chosen so that the number of connections equals the empirical value 2268. This graph G_d has a geometrical structure, which changes with the index number d . Put simply, d is the dimension of the characteristic space. Note that the boundary condition is not periodic. Thus, the characteristic space has a solid boundary. As a consequence, the vertices near the center of the characteristic space form the central group. The numerical results are shown for three values of d in Fig. 3. For each d , we generate 100 different samples, where we leave out disconnected cases. Figs. 3(a) and (b) show that for larger value of d , the sequences of degrees and GE for G_d agree well with those for G_{ce} . On the other hand, Fig. 3(c) shows that too large value of d destroys the agreement for the clustering structure. It is because the limit $d \rightarrow \infty$ corresponds to the random connection. Thus, we obtain the best agreement for the middle value $d \sim 8$. This result suggests that G_{ce} has the geometrical structure, where the dimension of the characteristic space is approximately eight.

4. Discussions

We have investigated the somatic nervous system of *C. elegans*, which has the central part and the clustering structure. We proposed the model with the geometrical structure and anticipated its dimension. We can modify the definition of the distance or the distribution in the characteristic space. Such modifications yield quantitative changes, serious qualitative change is hardly seen. It is essential that the characteristic space has a solid boundary. Thus, we conclude that the somatic nervous system has the high-dimensional geometrical structure which is not seen in the NR or WS model. According to the study of cellular automata and coupled map lattices, high dimension is one of the crucial ingredients in the subtle build-up of correlations giving rise of the collective motion [20].

Whereas, we ignored the direction, multiplicity and type of connections, they are important for the neural information transmission. Future models will require to incorporate these attributions into the graph. The present approach is available for the neuronal network of the pharyngeal nervous system. However, the geometrical structure is not so clear, because there are too less number of vertices (only 20 neurons). It is an interesting question whether the neuronal networks of other animals have the similar topology. Since no data of the connections in the nervous systems of other animals is available now, the answer cannot be given here. A higher animal, which has more neurons than *C. elegans*, appears to have more complicated structure. Remembering the fact that even *C. elegans* has two separated subunits in the nervous system, the

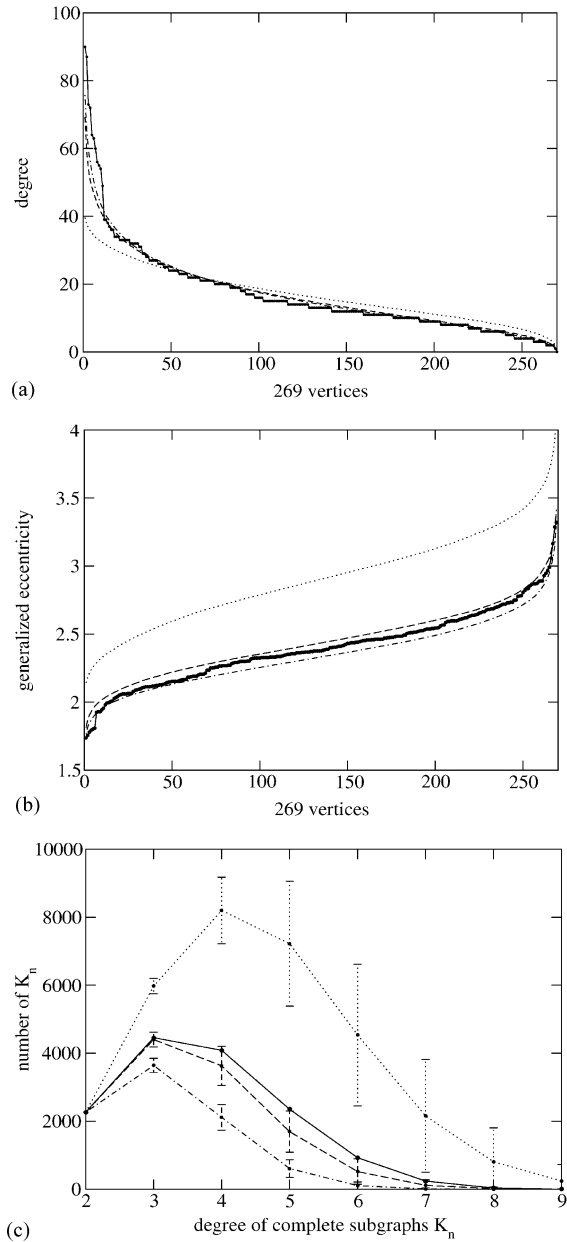


Fig. 3. The comparison between G_{cc} and G_d : (a) the degree sequence, (b) the GE sequence, and (c) the number of complete subgraphs. The solid lines are for G_{cc} . The dotted, broken and dotted-broken lines are for G_d of $d=4$, $d=8$ and $d=12$, respectively. The lines for G_d are obtained by the average over 100 samples. Whereas, RMS of the degrees are relatively large (the maximum is about 12 for $d=8, 12$) for vertices which have large degrees, they are smaller than 1 for the vertices whose degrees are small and medium values. RMS of the GE are smaller than 0.07 except a few vertices which have large GE. RMS of the number of complete subgraphs are shown by the error bars.

present approach may be expected to be valid for subunits rather than the whole of the nervous system.

Acknowledgements

We would like to thank H. Kagawa, R. Hosono and S. Mitani for their lectures, which motivated this research. We also thank H. Takano, S. Gomi, Y. Iwasaki, K. Omata and E. Akiyama for fruitful discussion and helpful comments. This research is supported by Japan Society for Promotion of Science under the contract number RFTF96I00102.

References

- [1] C. Koch, G. Laurent, *Science* 284 (1999) 96–98.
- [2] T.B. Acharso, W.S. Yamamoto, *AY's Neuroanatomy of C. elegans for Computation*, CRC Press, Boca Raton, FL, 1992.
- [3] R. Durbin, <http://elegans.swmed.edu/parts/>, 1995.
- [4] K. Oshio, S. Morita, Y. Osana, K. Oka, *C. elegans synaptic connectivity data*, Technical Report, CCEP, Keio Future No. 1, 1998.
- [5] D.G. Albertson, J.N. Thomson, *Philos. Trans. R. Soc. London Ser. B* 275 (1976) 299–325.
- [6] J.G. White, E. Southgate, J.N. Thomson, S. Brenner, *Philos. Trans. R. Soc. London Ser. B* 314 (1986) 1–340.
- [7] D.J. Watts, S.H. Strogatz, *Nature* 393 (1998) 440–442.
- [8] D.J. Watts, *Small Worlds*, Princeton University Press, Princeton, 1999.
- [9] M.E.J. Newman, *J. Statist. Phys.* bf 101 (2000) 819–841.
- [10] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.
- [11] L.A.N. Amaral, A. Scala, M. Barthélémy, H.E. Stanley, *Proc. Natl. Acad. Sci. USA* 97 (2000) 11,149–11,152.
- [12] A.-L. Barabási, R. Albert, *Science* 286 (1999) 509–512.
- [13] A.-L. Barabási, R. Albert, H. Jeong, *Physica A* 281 (2000) 69–77.
- [14] R. Albert, H. Jeong, A.-L. Barabási, *Nature* 301 (1999) 130–131.
- [15] R. Albert, H. Jeong, A.-L. Barabási, *Nature* 406 (2000) 378–382.
- [16] S. Redner, *Eur. Phys. J. B* 4 (1998) 131–134.
- [17] K. Kawamura, S. Morita, K. Oshio, Y. Funabashi, Y. Osana, K. Oka, preprint, 2000.
- [18] G. Chartrand, O.R. Oellermann, *Applied and Algorithmic Graph Theory*, McGraw Hill, New York, 1993.
- [19] M. Chalfie, J.E. Sulston, J.G. White, E. Southgate, J.N. Thomson, S. Brenner, *J. Neurosci.* 5 (1985) 956–964.
- [20] H. Chaté, P. Manneville, *Prog. Theor. Phys.* 87 (1992) 1–60.
- [21] <http://vlado.fmf.uni-li.si/pub/network/pajek/>.