From Sierpinski Carpets to Directed Graphs

Martin Hülse

Computer Science Department
Aberystwyth University
Aberystwyth, SY23 3DB, UK
msh@aber.ac.uk

Models of complex information processing based on artificial neural networks frequently apply fully connected or random graph structures. However, it is well known that biological neural systems operate on sparsely connected networks having properties quite distinct to random graphs. In this paper, a simple method is introduced for the deterministic generation of strongly connected digraphs. The method is inspired by Sierpinski carpets. Despite the large size of these digraphs, the distance between most of the nodes is short, that is, it scales logarithmically. It is further shown that important network properties, such as average degree and degree distribution, can directly be determined by the initial structure of this process. These findings lead to the formulation of general conditions providing a targeted generation of complex networks of arbitrary size. The circumstances under which these digraphs can show scale-free and small-world properties are discussed and finally possible applications of this method are outlined in the domain of artificial neural networks.

1. Introduction

Artificial neural networks (ANNs) represent a method of information processing that is inspired and motivated by the neural structures found in biological systems [1]. Therefore, it is not surprising that ANNs are frequently utilized as the basic building blocks for large-scale models in order to explore the nature of complex information processing exploited in animals and human beings.

The majority of such neural models are based on connectivity structures that match with the classical types of ANNs, such as multilayered perceptrons, Hopfield networks, or Elman networks [2, 3]. All of the classical network types establish only fully connected networks. The application of fully connected networks, however, might become crucial with respect to plausibility if they are intended to model biological systems. Fully connected ANNs can hardly represent very sparsely connected brain-like neural structures, if, as only one example [4], "in the mouse cortex only approximately 1 in 100 million of all possible connections are actually made."

An alternative, in particular for large-scale neural models, to overcome fully connected neural networks is the creation of random

https://doi.org/10.25088/ComplexSystems.19.1.45

graph structures [5]. Nevertheless, random graph models do not describe some essential properties of real-world networks [6].

Therefore, we argue, while modeling large-scale neural networks, alternatives for the projections between neural assemblies must be considered, including alternatives that go beyond random graphs and fully connected structures.

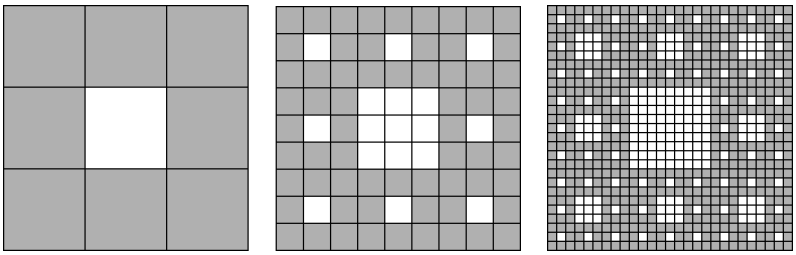
Furthermore, it becomes more and more common in the field of neural computation [7] that large-scale neural models are used for robot control [8, 9]. This leads very often to implementations of ANNs on special hardware devices, like massively parallel processor array VLSI circuits [10]. Such implementations on autonomous robots might be motivated as a proof of concept, as well as for targeting specific issues of embodiment [11]. However, autonomous robots usually have very limited computational resources, especially memory, compared with the performance provided by computer clusters or similar equipment. Hence, for performance reasons it is important to utilize highly connected and robust networks established by as few connections as possible.

The objective of this paper is to introduce a deterministic method for creating highly connected and structured neural systems formed by a number of connection magnitudes smaller than needed for fully connected networks. The generation process of such sparsely connected networks is inspired by fractal sets. This makes the resulting networks very distinct compared with random graphs and as we will see, they can cover, depending on the initialization, a wide range between fully connected graphs and connected graphs organized as rings. Due to the simplicity and the deterministic character of the generation process, this method seems to be a promising alternative for the generation of graphs and opens a wide field for applications in many areas of neural modeling.

Fractal sets, invented and promoted by Mandelbrot [12], are established tools for describing and modeling complex structures and processes, such as textures of surfaces or even the state space of chaotic attractors. Sierpinski carpets are well-known examples of mathematical shapes forming fractals (see Figure 1). Inherent properties of fractals, being self-similar and scale-free, can impressively be demonstrated with these sets. Therefore, we have asked what types of graphs or networks can be expected to emerge if the intermediate sets resulting from a generation process toward Sierpinski carpets are interpreted as adjacency matrices.

This paper presents an investigation of the properties of graphs constructed in the same fashion as Sierpinski carpets. As we will show, the result of this investigation is a simple method for a deterministic generation of strongly connected directed graphs. The manifold of possible graphs provided by this method is systematically analyzed for a low-dimensional case. But as we will see, this analysis already leads us to general conditions which guarantee robustness as well as specific degree distributions for arbitrary dimensions.

https://doi.org/10.25088/ComplexSystems.19.1.45



(a) (b) (c)

Figure 1. An example for a Sierpinski carpet. Its mask and initial form (a) as well as the first (b) and second (c) iteration of the process generating the Sierpinski carpet.

Based on these findings, we introduce two strategies that allow applying our graph generation method in the domain of ANNs. But before we discuss these aspects in detail, Section 2 introduces basic definitions and explains the process of graph generation. This includes the introduction of a sufficient condition maintaining the property of being strongly connected in general. After this, the next two sections describe essential network properties (e.g., shortest paths, degree distribution, clustering, and robustness) resulting from a systematic analysis of a representative subset. This is followed by a discussion summarizing our findings in a more general form leading to the outline of possible applications for generating ANNs, as aforementioned.

2. Generating Digraphs of Fractal Dimension

A directed graph/digraph G is a set of vertices V (sometimes also called *nodes*) and connections E (*edges*) between them. An edge e_{ji} only connects two nodes v_i and v_j, where e_{ji} is the incoming edge for v_j and the outgoing for v_i.

We call the total number of incoming and outgoing edges the degree k. In directed graphs for a single node, the number of incoming edges k_i can be different from the number of the outgoing edges k_o. Note that k_o and k_i are also referred to as out-degree and in-degree, respectively.

If there exists a path between each pair of nodes in the digraph, we call it *strongly connected*. In the case that node v_j can be reached from v_i, while there is no path back to v_i starting in v_j, the graph is only *connected*.

The structure of a digraph $G(V, E)$ can be represented by an adjacency matrix M. Each matrix element m_{ji} of an adjacency matrix can either be 0 or 1. The element m_{ji} is one, if and only if $e_{ji} \in E$.

https://doi.org/10.25088/ComplexSystems.19.1.45

As we show in the following, adjacency matrices build the bridge between Sierpinski carpets and directed graphs as well as giving us a process for the deterministic development of digraphs.

In Figure 1 an example for the construction of a Sierpinski carpet is illustrated. The process starts with the basic form of a square. Each side of this square is segmented into three equal sections defining an overall partition into nine identical squares. Some of these squares are labeled, as indicated by the gray coloring. We call this basic form a *mask*.

As can be seen in the figure, the Sierpinski carpet is easily generated by applying the same partitioning, defined by the mask, for all the labeled squares of the given form. In contrast to the original generation of Sierpinski carpets, now unlabeled squares are also subdivided in the same way but the resulting subsquares remain unmarked.

The partitioning of labeled and unlabeled squares, including the new labeling of the new set of squares, is what we call an *iteration*.

In this process a distinction has to be made between a mask and the form. The mask defines the partitioning while the form is the structure the mask is applied to. Indeed, a mask can be applied to any form. In the following, mask and initial form are always identical. All the resulting forms therefore are fully determined by the mask and the number of iterations. For an infinite number of iterations we get the Sierpinski carpet, that is, a set of fractal dimension.

We utilize this type of fractal generating process in order to generate directed graphs simply by interpreting the resulting set after n iterations as an adjacency matrix of a directed graph. Examples of 3-segmented forms transformed into digraphs are given in Figure 2. The labeled squares are interpreted as edges, that is, a gray color represents the value 1 in the corresponding adjacency matrix, while unlabeled squares indicate the 0 entries.

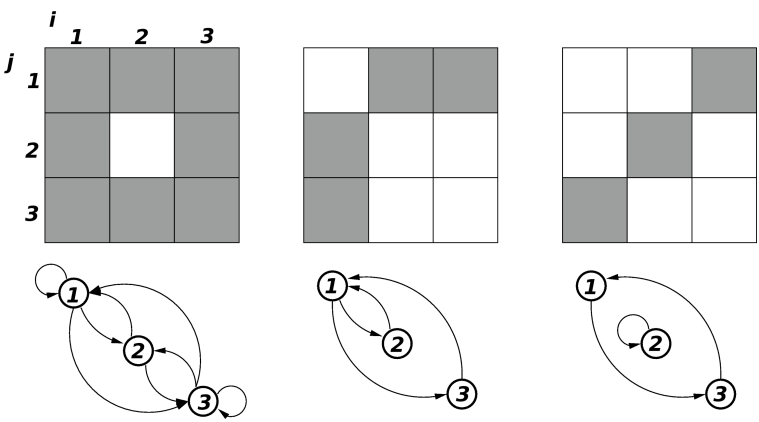


Figure 2. Examples of digraphs derived from 3-segmented forms.

https://doi.org/10.25088/ComplexSystems.19.1.45

Considering our transformation between patterns and adjacency matrices, we see that the Kronecker product [13] can directly be applied in order to define an algorithm that is isomorph to the process generating Sierpinski carpets:

$$\begin{aligned} D_0 &:= M_S \\ D_{i+1} &:= D_i \otimes M_S, \end{aligned}$$

where M_S is an adjacency matrix ($M(i, j) \in \{0; 1\}$) of dimension $S \times S$ representing the mask. In the following we refer to this algorithm as the *digraph generating process* (DGP).

Given the DGP, we want to know what kind of digraphs can be expected. An investigation of this question needs to focus on the masks M_S and the number of iterations i only, because these are the parameters that determine the resulting structures. For the 3-segmented case, we get $2^9 = 512$ different masks. Higher segmentations S generate $2^{(S \cdot S)}$, since an adjacency matrix has to have equal dimensions.

In order to cope with this exponentially increasing number of possible masks, we begin our analysis for $S = 3$. This analysis will provide insights about the interrelation between mask properties and the global structure of the resulting digraphs and therefore, will guide us into the huge space of digraphs spanned by the masks of higher segmentations.

The first constraint for our investigation is that we are interested in strongly connected digraphs only. This reduces the number of masks to be considered. We will show that this set can be further reduced by taking into account certain symmetries of the adjacency matrices, which maintain the property of being strongly connected. In that way we get a manageable number of remaining basic forms.

2.1 Labeling and Filtering of the Masks

We start by introducing a general numbering or labeling of the masks. As a unique numbering we have chosen the binary code derived directly from the structure of the mask/adjacency matrix (see Figure 3(a)). It can be seen that the labeled and unlabeled mask elements are interpreted as 1 and 0 of a binary number. However, this numbering is only unique if the mask segmentation is taken into account. Therefore, we use the symbol M_n^S, where n is the number that in binary representation corresponds to the mask of segmentation S. An example is given in Figure 3(b). It can be seen that the number M_{511}^3 refers to the 3-segmented mask in which all segments or entries are labeled. In the 4-segmented case, the number 511 represents a totally different mask structure.

https://doi.org/10.25088/ComplexSystems.19.1.45

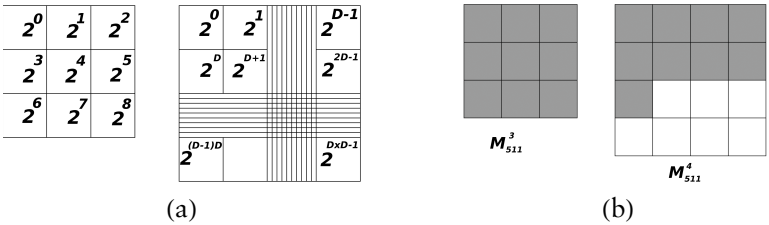


Figure 3. (a) Schema of the three-dimensional and D-dimensional mask indicating the value of each entry in order to derive the number of the binary code given by the mask. (b) Examples of the number 511, which in binary code represents different masks depending on the dimension.

2.2 Generating Strongly Connected Digraphs

As mentioned earlier, we only consider strongly connected digraphs. One necessary condition of being strongly connected is that each node has at least one outgoing and one incoming edge, $k_{o/i} > 0$. The value of k_i and k_o can be derived from the corresponding adjacency matrix M. Value k_o of node v_l is the sum over all entries in column l, while k_i is the sum over row l:

$$k_o(v_l) := \sum_j m_{j,l} \qquad k_i(v_l) := \sum_i m_{l,i}.$$

After applying the DGP it can be seen that each mask containing a node of in- or out-degree zero will inevitably result in graphs having at least one node with in- or out-degree zero. Thus, they are not strongly connected. Therefore, if a mask of segmentation S generates strongly connected digraphs, then this mask must have at least S labeled entries. In other words, a mask must contain at least as many labeled squares (edges) as columns/rows (nodes).

Apart from this condition, we have shown in [14] that each strongly connected digraph organized as a ring applied to the DGP generates digraphs that separate into disconnected subgraphs. In short, a mask representing a ring does not produce strongly connected or any connected digraphs at all.

On the other hand, we have proven that the following criterion guarantees strongly connected digraphs [14]: If a mask M_S represents a ring with at least one node having a self-connection, then the DGP generates a strongly connected digraph in each iteration. A self-connection is represented by a nonzero entry in the main diagonal of the adjacency matrix. According to this condition, the adjacency matrix of an S-segmented mask must contain at least $S+1$ edges: S edges to form a ring plus one self-connection.

Note that the introduced criterion is only a sufficient condition but applies to all segmentations ($S > 2$) and iterations. It is also worth not-

https://doi.org/10.25088/ComplexSystems.19.1.45

ing that this condition provides a huge set of masks generating strongly connected digraphs. For a given segmentation S there are $S!$ possible ring structures. Taking into account the self-connection needed, we get a set of $S \cdot S!$ masks generating strongly connected digraphs. Furthermore, each mask in this set can also be used as a substructure since additional connections do not destroy the property of being strongly connected.

2.3 Fractal Dimension of Strongly Connected Digraphs

It is easy to see that a mask that is fully labeled generates only fully connected digraphs. If we consider the given criterion, then the nontrivial cases of strongly connected digraphs are generated by masks with n labeled entries, where $S < n < S^2$. Interestingly enough, masks with this number of labeled segments generate Sierpinski carpets of fractal dimensions d_f between 1 and 2 [12], since

$$d_f = \frac{\log(n)}{\log(S)} \qquad S < n < S^2, \tag{1}$$

from which follows: $1 < d_f < 2$. Therefore, we define a graph G as a *strongly connected digraph of fractal dimension* if G is a strongly connected digraph and the result from the DGP applied to a mask of fractal dimension d_f with $1 < d_f < 2$.

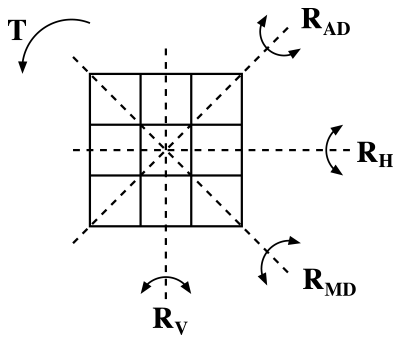


Figure 4. General rotation and reflection symmetries of Sierpinski carpets. Certain Sierpinski carpets remain the same applying a rotation (T) as well as under reflections along the middle horizontal or the vertical line (R_H and R_V) and along the two diagonals R_{MD} and R_{AD}.

2.4 Symmetrical Masks

Sierpinski carpets can be symmetric. Figure 4 shows the five possible transformations under which certain sets undergo no alterations. The

https://doi.org/10.25088/ComplexSystems.19.1.45

Sierpinski carpet shown in Figure 1 remains obviously the same for each of the five transformations.

If we apply these transformations to adjacency matrices, then only reflections along the two diagonals (R_{AD}, R_{MD}) as well as two successively applied rotations (T^2) keep a digraph strongly connected. Taking into account these three symmetries, we get 50 distinct 3-segmented masks of fractal dimension generating strongly connected digraphs. All other masks produce either no strongly connected digraphs at all or can be transformed into one of the 50 masks by a combination of the operations T^2, R_{MD}, and/or R_{AD}. Table 1 lists these masks indicating their fractal dimension. A 4-segmentation gives us 6692 unique masks out of $2^{16} = 65\,536$ possibilities.

$\lvert E_0 \rvert$	$d_f = \frac{\log(E_0)}{\log(S)}$	M_n^3	Σ
4	1.26	99, 102, 106, 114	4
5	1.46	79, 94, 103, 107, 110, 115, 118, 122, 171, 173, 174, 186, 229, 355	14
6	1.63	95, 111, 119, 123, 126, 175, 187, 189, 190, 231, 238, 245, 335, 359, 363, 371, 427	17
7	1.77	127, 191, 239, 247, 254, 351, 367, 375, 379, 431, 443	11
8	1.89	255, 383, 447, 495	4
			$\Sigma = 50$

Table 1. Numbers of unique 3-segmented masks generating strongly connected digraphs and their fractal dimension. Some of them do not match the introduced condition, but yet they create strongly connected digraphs.

3. Properties of Digraphs with Fractal Dimension

In this section we investigate some properties of the strongly connected digraphs constructed by the masks given in Table 1. As already mentioned, this list contains all unique 3-segmented masks that are strongly connected and create strongly connected digraphs via the DGP. This was explicitly tested because some of the masks do not match the sufficient criterion we introduced. For all masks, we applied five iterations and therefore the digraphs under investigation have $3^6 = 729$ nodes.

3.1 Shortest Paths

The most important feature of a strongly connected digraph is the average of the shortest paths and the evolution of this value over the iterations. Self-connections are not included in this calculation.

https://doi.org/10.25088/ComplexSystems.19.1.45

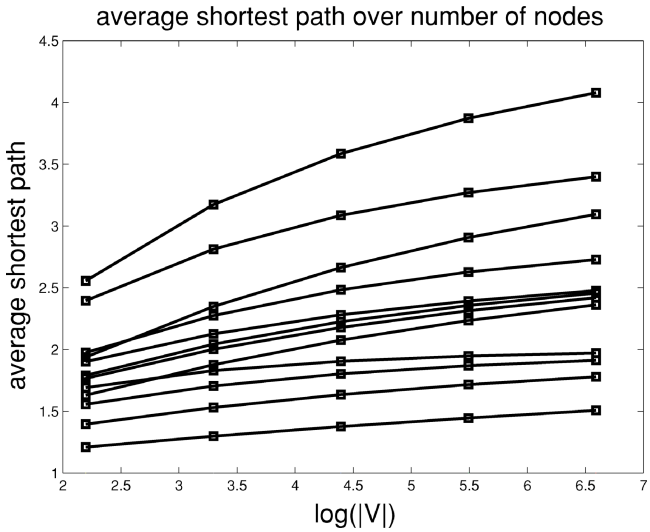


Figure 5. Diagram showing the evolution over five iterations of the average shortest path length with respect to the number of nodes (scaled logarithmically). All masks are plotted in this diagram and overlapping curves appear as one.

As indicated by the diagram in Figure 5, the average length of the shortest paths increases logarithmically with the number of nodes. The average length of the shortest paths correlates with the fractal dimension. The smaller the fractal dimension, the larger the shortest path between the nodes, on average.

3.2 Average Degree

A judgment about the shortest paths has to include the corresponding average degree. Figure 6 shows the evolution of the out-degrees over the iterations for all digraphs. It is obvious that the average degree correlates to the fractal dimension of the mask. Indeed, the average of the total degree $\overline{k}$ is determined by the ratio of the number of edges to the number of nodes:

$$\overline{k}_{\text{total}} = 2\,\frac{|E|}{|V|}.$$

However, for our fractal digraphs we know that

$$\left|E_i\right| = E_0^{i+1} \tag{2}$$

and

$$\left|V_i\right| = S^{i+1} \tag{3}$$

https://doi.org/10.25088/ComplexSystems.19.1.45

where $|E_i|$ and $|V_i|$ refer to the number of edges and nodes after $i > 0$ iterations while E_0 is the number of edges in the applied mask M_S. S is the segmentation of M_S, which gives the number of nodes. Hence, after i iterations we get the total degree:

$$\overline{k}_{\text{total}}(i) = 2 \cdot \left(\frac{E_0}{S}\right)^{i+1}, \qquad S < E_0 < S^2. \tag{4}$$

The average degree is obviously determined by E_0 and S. Both determine the fractal dimension of the mask. For $S = 3$ we have five values for E_0 ($\{4, 5, 6, 7, 8\}$). These five values result in the five distinct curves in the diagram of Figure 6, although all of the 50 unique 3-segmented masks are plotted.

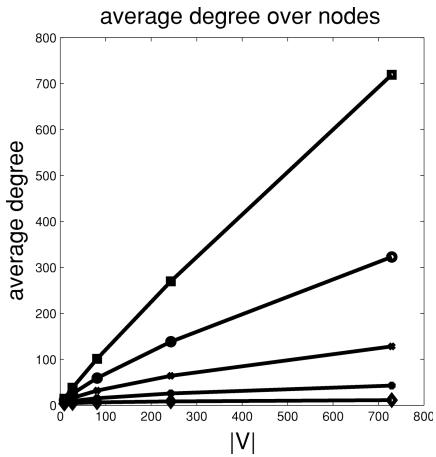


Figure 6. Diagram showing the evolution over five iterations of the mean of the in-degree k_i (equal to k_o) with respect to the number of nodes. As in Figure 5, all masks are plotted and overlapping curves appear as one.

Therefore, a general property of our fractal digraphs is that the average degree is exponentially increasing with respect to the number of nodes, since per definition $E_0 / S > 1$. Furthermore, the larger the fractal dimension, the larger the exponential increase. The same picture emerges if we consider either the in-degree or the out-degree only. In fact, we have

$$\begin{aligned} \overline{k}_i(i) &= \overline{k}_o(i) = \left(\frac{E_0}{S}\right)^{i+1} \\ \overline{k}_{\text{total}}(i) &= 2 \cdot \left(\frac{E_0}{S}\right)^{i+1} = 2 \cdot \overline{k}_i(i) = 2 \cdot \overline{k}_o(i) \end{aligned}$$

indicating that the difference between total degree and in- or out-degree is determined by a constant factor only.

https://doi.org/10.25088/ComplexSystems.19.1.45

3.3 Clustering

The third characteristic of graphs is the average clustering coefficient. We will apply the definition given in [15]. The resulting values over the iteration process for our digraphs is plotted in Figure 7. It can be seen that all values tend to decrease with respect to the size of the graph, that is, $|V|$.

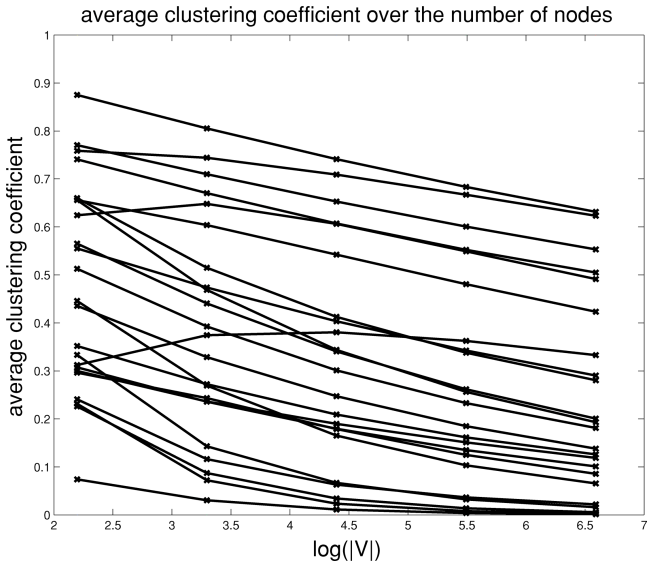


Figure 7. Clustering coefficient after the definition given in [15]. Once again, all 50 3-segmented masks are plotted.

The clustering coefficients of digraphs under investigation are often compared with the clustering values of random graphs $G_{n,p}$. A random graph belongs to the group $G_{n,p}$ if it is undirected and has n nodes, where each pair of nodes is connected with probability p [6]. Interesting for us is the fact that the average clustering coefficient of random graphs is equal to

$$c_r = p \tag{5}$$

with high probability [16].

In order to compare the clustering coefficient of random graphs with our digraphs of fractal dimension we define the ratio r of c_f to c_r:

$$r = \frac{c_f}{c_r} = \frac{c_f}{p}, \tag{6}$$

where c_r can be substituted by p.

https://doi.org/10.25088/ComplexSystems.19.1.45

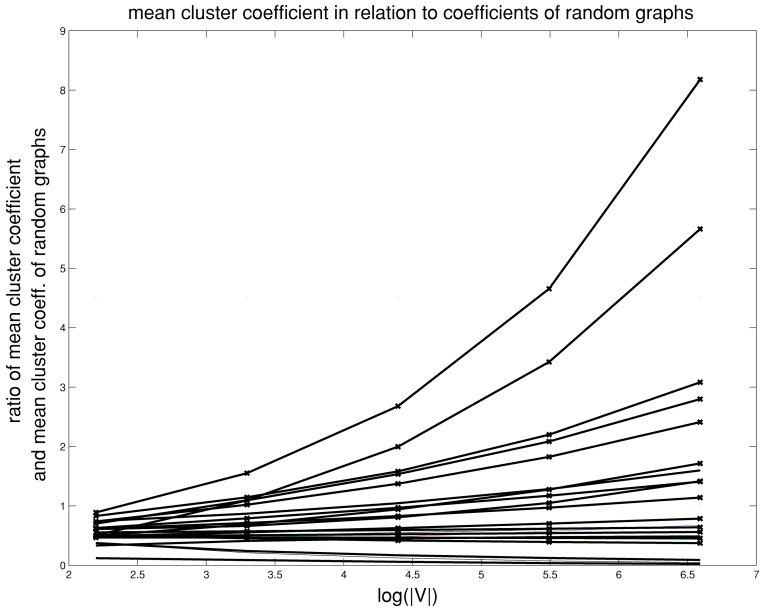


Figure 8. Evolution of $r = c_f / c_r$ over the number of iterations. In other words, the relation between the clustering coefficients of random graphs and digraphs of fractal dimension having the same size.

Assume a digraph of fractal dimension G_f with $|E|$ edges and $|V|$ nodes. Then we can estimate p of the corresponding random graph $G_{|V|,p}$ with the same number of nodes and edges in the following way:

$$p = \frac{|E|}{|V|^2} = c_r, \tag{7}$$

which is equal to the clustering coefficient c_r of $G_{|V|,p}$, with high probability. Hence, we have an estimation of the clustering coefficient c_r of a random graph having the same number of edges and nodes as G_f. Therefore, the ratio r of c_f to c_r can be written as:

$$r = \frac{c_f \cdot |V|^2}{|E|}, \tag{8}$$

where $|V|$ and $|E|$ are the number of nodes and edges in G_f and c_f is the clustering coefficient of G_f which was calculated empirically (results shown in Figure 7).

Figure 8 shows the evolution of r over the number of iterations. It can be seen that digraphs of small fractal dimension are more likely to

https://doi.org/10.25088/ComplexSystems.19.1.45

have a significantly larger clustering coefficient than a random graph of the same size. This relation is better illustrated later in Figure 9. On the other hand, one can see that the clustering coefficient of a fractal digraph can also be much smaller than in the corresponding random graph.

3.4 Degree Distribution

The last property under investigation in this paper is the degree distribution. Similar to the average degree, this distribution is fully determined by the mask. Let

$$d_i = k_o\left(v_i\right), \qquad 1 \le i \le S,$$

where S is the segmentation of mask M_S. Hence, d_i is the out-degree of node i. In the following we will only consider out-degrees. Nevertheless, the same argumentation also holds for the distribution of the in-degrees.

The out-degree of each node in a digraph after j iterations can be derived by solving

$$\left(d_1 + d_2 + \cdots + d_S\right)^{j+1}, \qquad j > 1,$$

in a symbolic manner to get a sum of S^{j+1} elements. Each summand is a product of $(j+1)$ factors. Each product represents the out-degree of a particular node in the resulting digraph. Whereas, each factor in the product represents an out-degree in the originally given mask M_S. Hence, the distribution of the out-degrees is given directly by the out-degree values in M_S and the iterations. As an example, assume we have

$$d_i < d_S, \qquad 1 \le i < S.$$

Therefore, after j iterations we have exactly one node, out of $n = S^{j+1}$ nodes, with the maximal out-degree ${d_S}^{j+1}$. Furthermore, let $d_{1,\ldots,S-1} = 1$ and $d_S > 1$. In such a case the number of nodes z_r with out-degree ${d_S}^{n-r}$ is given by

$$z_r = \binom{n}{n-r} \cdot (S-1)^r, \qquad 0 \le r < n, \tag{9}$$

where S is the segmentation of the mask and $n = S^{j+1}$ is the number of nodes after j iterations. The term $\binom{n}{n-r}$ can be substituted by

$$(n-r)^r \cdot \prod_{l=1}^{r} \left(\frac{1}{l} + 1\right) \tag{10}$$

https://doi.org/10.25088/ComplexSystems.19.1.45

and with n as the total number of nodes, we determine the probability $P(r)$ for selecting a node with degree $d_S^{(n-r)}$ in the following way:

$$P(r) = \frac{z_r}{n} = \frac{(S-1)^r \cdot (n-r)^r}{n} \cdot \prod_{l=1}^{r} \left(\frac{1}{l} + 1 \right). \tag{11}$$

According to the following estimation:

$$1 < \prod_{l=i}^{r} \left(\frac{1}{l} + 1 \right) \leq 2^r \tag{12}$$

we get:

$$\frac{((S-1)\,(n-r))^r}{n} < P(r) \leq 2^r \, \frac{((S-1)\,(n-r))^r}{n}, \qquad 0 \leq r < n. \tag{13}$$

It follows, that for this specific case, the tail of the corresponding out-degree distribution ($r > n/2$) is decreasing exponentially.

Obviously, the distribution of the out-degree has no characteristic scale and due to the construction process there will always be "a few nodes" (called hubs) with a degree that is magnitudes larger than the average [17]. As a consequence, we can say that there are digraphs of fractal dimension that can represent scale-free networks.

On the other hand, we can see that in-degree and out-degree are totally independent. They are determined by the corresponding degrees given in the mask. Hence, the distribution of the total degree is the sum of the in- and out-degree distributions.

3.5 Five Examples Representing the Five Possible Fractal Dimensions

The diagrams in Figure 9 summarize the main findings of this section by showing the qualities under investigation (average shortest path length, clustering coefficient, and the ratio of this coefficient in relation to random graphs) for five examples. Each mask has a different fractal dimension.

The mean value of the length of the shortest paths increases logarithmically. We can also see that the larger the fractal dimension, the shorter the paths connecting the nodes. Our empirical investigations of the clustering coefficients indicate that they tend to decrease over the iterations. Compared to random graphs having the same number of nodes and edges, however, one can see that masks with small fractal dimension (here M^3_{114}, M^3_{186}, and M^3_{335}) have significantly larger clustering coefficients. This difference even grows with the number of iterations.

https://doi.org/10.25088/ComplexSystems.19.1.45

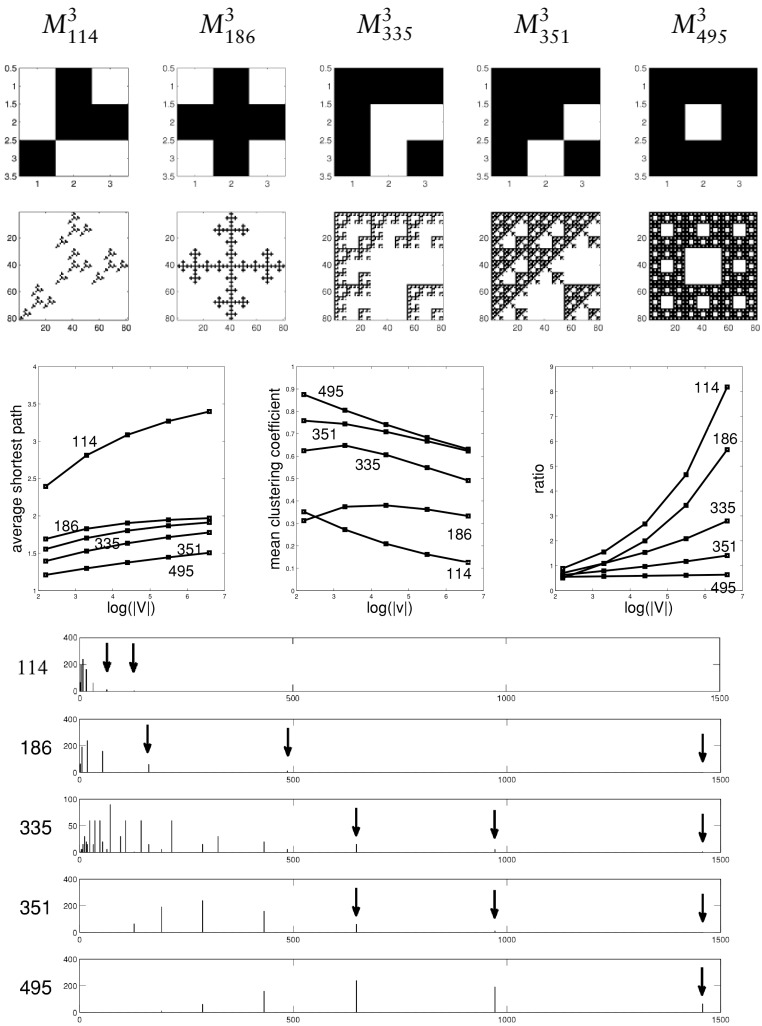


Figure 9. Characteristic qualities of five representative examples of digraphs of fractal dimension, including the distribution of the total degree. Each mask has a different fractal dimension. See text for explanations.

The bottom of Figure 9 shows the histograms of the total degree for each strongly connected digraph after five iterations. Hence, a total degree number of $\left(3^6 - 1\right) \cdot 2 = 1456$ can be expected. Some of the distributions have a shape very similar to the typical scale-free distribution. Namely, the digraphs resulting from M^3_{114}, M^3_{186}, or M^3_{335} have a "long tail" referring to a few nodes with significantly larger total degree, compared to the majority. The arrows indicate small but

https://doi.org/10.25088/ComplexSystems.19.1.45

nonzero entries in the histogram that cannot be seen in this illustration clearly.

4. Robustness

The robustness of digraphs can be investigated with respect to many specific properties. In the following, we only consider the property of being strongly connected and the average length of the shortest paths while nodes are removed. No matter which properties we are interested in, for any type of graph or network there are two types of robustness to consider. On the one hand, we want to estimate how essential properties of a network do change if a randomly chosen node is removed. On the other hand, it is important to know whether or not a graph contains single nodes that have a radical impact on a global scale. Nodes of the latter type make a network highly sensitive to targeted attacks because the failure of a single specific node causes a catastrophic failure of the whole network.

Due to the deterministic nature of our generation process, once more we can start to investigate these questions in a systematic manner.

4.1 Robustness under Attack

Whether or not a strongly connected digraph is not strongly connected after a particular node is removed is verified by removing it and testing explicitly if the digraph is still strongly connected. Nodes destroying the property of being strongly connected are called *crucial* nodes.

All digraphs during the first four iterations were analyzed in this way and Table 2 gives a summary of this experiment. One can see that apart from 16 digraphs, all others have at least one crucial node. Digraphs that have crucial nodes can be distinguished over the iterations with respect to the change in their number of crucial nodes as either exponentially increasing or remaining constant.

We cannot explain yet how this number of crucial nodes is related to the underlying structure of the mask and the DGP. However, it turns out that the number of crucial nodes is zero if the mask has no node with in- and out-degree less than two. This is formally written as

$$\forall\, i : k_o\left(v_i\right) > 1 \wedge k_i\left(v_i\right) > 1,$$

where v_i is a node in mask M_S.

https://doi.org/10.25088/ComplexSystems.19.1.45

E_0	M_n^3	1st	2nd	3rd	4th	E_0	M_n^3	1st	2nd	3rd	4th
4	99	7	15	31	63	6	231	1	1	1	1
	102	4	8	16	32		238	0	0	0	0
	106	4	8	16	32		245	0	0	0	0
	114	7	15	31	63		335	1	1	1	1
							359	2	2	2	2
5	79	1	1	1	1		363	1	1	1	1
	94	1	1	1	1		371	0	0	0	0
	103	4	8	16	32		427	0	0	0	0
	107	4	8	16	32						
	110	1	1	1	1	7	127	1	1	1	1
	115	2	2	2	2		191	1	1	1	1
	118	2	2	2	2		239	0	0	0	0
	122	4	8	16	32		247	0	0	0	0
	171	1	1	1	1		254	0	0	0	0
	173	2	2	2	2		351	0	0	0	0
	174	1	1	1	1		367	1	1	1	1
	186	1	1	1	1		375	0	0	0	0
	229	1	1	1	1		379	0	0	0	0
	355	2	2	2	2		431	0	0	0	0
							443	0	0	0	0
6	95	1	1	1	1						
	111	1	1	1	1	8	255	0	0	0	0
	119	1	1	1	1		383	0	0	0	0
	123	2	2	2	2		447	0	0	0	0
	126	1	1	1	1		495	0	0	0	0
	175	1	1	1	1						
	187	1	1	1	1						
	189	1	1	1	1						
	190	1	1	1	1						

Table 2. Numbers of crucial nodes. This value is given for the first four iterations. Hence, the largest digraphs have $3^5 = 243$ nodes.

The test of being strongly connected after removing two nodes gives the same result (see Table 3). We only remove two nodes if they are not crucial nodes. Hence, the nodes counted in Table 2 are not considered in Table 3. Again, the strong connectedness of a digraph cannot be destroyed by removing two nodes if the underlying mask contains only nodes with in- and out-degree larger than one.

https://doi.org/10.25088/ComplexSystems.19.1.45

E_0	M_n^3	1st	2nd	3rd	E_0	M_n^3	1st	2nd	3rd
4	99	0	0	0	6	231	3	3	4
	102	2	0	0		238	0	0	0
	106	2	0	0		245	0	0	0
	114	0	0	0		335	0	0	0
						359	0	0	0
5	79	0	0	0		363	2	3	4
	94	3	3	3		371	0	0	0
	103	0	0	0		427	0	0	0
	107	0	0	0					
	110	4	3	4	7	127	0	0	0
	115	4	6	8		191	0	0	0
	118	6	6	8		239	0	0	0
	122	0	0	0		247	0	0	0
	171	3	3	4		254	0	0	0
	173	6	6	8		351	0	0	0
	174	4	3	4		367	0	0	0
	186	0	0	0		375	0	0	0
	229	3	3	4		379	0	0	0
	355	4	6	8		431	0	0	0
						443	0	0	0
6	95	0	0	0					
	111	0	0	0	8	255	0	0	0
	119	2	3	4		383	0	0	0
	123	0	0	0		447	0	0	0
	126	3	3	4		495	0	0	0
	175	3	3	4					
	187	0	0	0					
	189	2	3	4					
	190	0	0	0					

Table 3. Numbers of pairs of nodes that cause the graph to become disconnected when deleted. A single node of these pairs does not destroy the connectedness.

This simple relation was also tested for the first two iterations of the 4-segmented masks. It holds for these cases as well. Therefore, although not formally proven, our experiments give evidence that if a given mask M_S represents a strongly connected digraph and each node in M_S has at least two outgoing and two incoming edges, then the property of being strongly connected of the resulting digraphs cannot be destroyed by removing one or two nodes.

4.2 Robustness under Failure

The experiments outlined in Section 4.1 indicate that the number of crucial nodes (if they are present at all) either remains constant or in-

https://doi.org/10.25088/ComplexSystems.19.1.45

creases exponentially with respect to the iterations. Nevertheless, the total number of nodes in the digraphs is exponentially increasing as well. In order to provide an estimation of the likelihood of destroying the property of being strongly connected if a randomly chosen node is removed, we have to consider the relative number of crucial nodes q in a given digraph. For the worst cases (i.e., M^3_{99} and M^3_{114}) we get the relation

$$q(i) = \frac{V_c(i)}{V(i)} = \frac{2^{(i+2)} - 1}{3^{i+1}}, \tag{14}$$

where $V_c(i)$ represents the number of crucial nodes and $V(i)$ is the total number of nodes in the digraph after i iterations. Thus, the term $q(i)$ represents the relative number of crucial nodes in the strongly connected digraphs resulting from M^3_{99} and M^3_{114} after i iterations. Hence, the probability of destroying the strong connectedness of these digraphs by removing a single node accidentally drops exponentially with respect to the number of iterations. Since this statement holds for the worst cases of 3-segmentation, we formulate a hypothesis for arbitrary segmentations in the following way: the probability of turning a strongly connected digraph of fractal dimension into a non-strongly connected digraph by randomly removing a single node is either zero or drops exponentially with respect to the number of iterations.

4.3 Changes of the Shortest Path Length

The question remains of how other network properties change if the digraph is still strongly connected after a random deletion of one or more nodes. Here, we have investigated this issue with respect to the average length of the shortest paths. In these tests the maximal value of the shortest paths in a given digraph is used for the normalization of all the shortest path lengths. Hence, the normalized values indicate their length relative to the longest shortest path in a digraph. Consequently, if one or more nodes are removed, the mean value of the relative length must be larger or equal to the mean value in the original digraph.

The three examples shown in Figure 10 indicate a change in the relative length of the shortest paths after removing up to 30 nodes. All removed nodes were selected in a way that the resulting digraph was still strongly connected. Two of these curves represent the very few examples where significant changes occur. Most of our 50 digraphs of fractal dimension show results very similar to the curve of M^3_{126}. The value does not change significantly for several reasons. The first, of course, is simply due to the fact that the larger the fractal dimension, the more connections our digraphs have and therefore, less impact on shortest path lengths if single nodes are removed. On the other hand,

https://doi.org/10.25088/ComplexSystems.19.1.45

nodes were randomly selected and only removed if the digraph remained strongly connected. Therefore, most of the digraphs were only cut on the periphery, where changes have minor impact on a global scale.

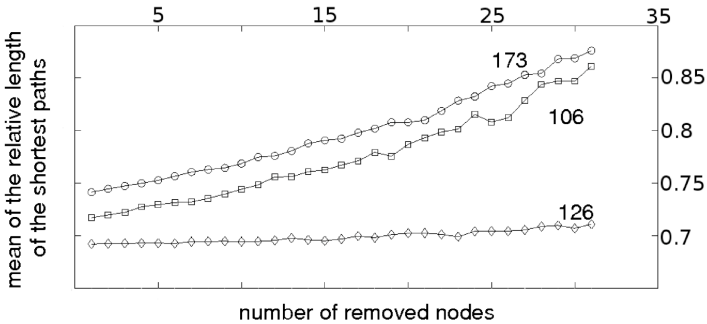


Figure 10. The plot shows the relative shortest path length after removing n nodes randomly and without destroying the strong connectedness. Results are shown for the digraphs generated by M^3_{106}, M^3_{173}, and M^3_{126}. Up to 30 nodes were deleted in each digraph. The relative length is related to the length of the longest shortest path in the original digraph, which has 81 nodes. Each deletion process was repeated 50 times.

In summary we can say, as long as the property of being strongly connected in our digraphs is not destroyed, the deletion of a few nodes has no significant impact on the average length of the shortest paths. According to our results it seems that the deletion of up to 30 nodes, out of 81, does not touch most of the shortest paths in the originally given digraphs at all.

5. Discussion

In the previous sections we analyzed digraphs generated with 3-segmented masks. Nevertheless, it was also shown that many properties of the resulting structures can directly be related to general mask properties, meaning that they apply for any segmentation S. Hence, we are able to determine and estimate graph properties resulting from masks of higher segmentations. This section briefly summarizes these findings in order to outline a more general picture of the directed graphs of fractal dimension. Finally, two methods for the generation of ANNs are discussed.

5.1 Being Strongly Connected

In general, we call a mask *complete* if it generates strongly connected digraphs in each iteration of the DGP. The sufficient condition we in-

https://doi.org/10.25088/ComplexSystems.19.1.45

troduced provides a huge set of complete masks for any segmentation. And, as we have argued, if a mask has a complete mask as a substructure then it also belongs to the set of complete masks.

However, it is still unknown which conditions will guarantee digraphs that are strongly connected for arbitrary adjacency matrices M and G combined via the Kronecker product $M \otimes G$. Nevertheless, as long as G and M are themselves complete masks or are the result of a complete mask, then the resulting digraph $(M \otimes G)$ is strongly connected. In consequence, one can think of applications where different complete masks are used during the generation process. This will provide a huge variety of strongly connected digraphs of arbitrary size. But this is an issue for future research or more concrete applications.

5.2 Parameterized Scale-Free Digraphs

Digraphs resulting from the DGP have illustrated that if they have scale-free properties then in- and out-degrees in the corresponding mask differ. (For example, consider masks where all nodes have the same in- and out-degree. In such a case, the degrees of the resulting digraphs are equal as well, which does not match with the characteristic degree distribution of scale-free networks.)

The resulting average degree and degree distribution can be derived easily from the values of in- and out-degrees in the mask and the number of iterations. Robustness issues can also directly be addressed under the consideration of minimal in- and out-degrees in the mask. As we saw, if each in- and out-degree in the mask is larger than one, then the generated networks are robust against targeted attacks. These findings hold for any segmentation without exception.

In summary, we are able to generate digraphs of characteristic average degree, degree distribution, and robustness just by "tuning" the in- and out-degrees in the mask irrespective of its segmentation.

5.3 Small-World Properties

Small-world properties of networks are not uniformly defined in the literature [6, 17]. Consequently, whether or not a graph establishes a small-world network somehow depends on the definition applied. Furthermore, small-world properties are very often defined in relation to random graphs. This class, however, does not well match the deterministic nature of our fractal digraphs. It is therefore not surprising that digraphs of fractal dimension are small-world networks as well as they are not. It depends on the given definition or, in other words, on the point of view.

One definition for small-world networks, given in [6], relates the average degree to the shortest paths. According to the definition provided, a graph is a small-world network if the average length of the shortest paths scales logarithmically with respect to the number of nodes, while the average degree does not change.

https://doi.org/10.25088/ComplexSystems.19.1.45

As mentioned earlier, the average of the total degree $\overline{k}$ of our digraphs increases exponentially with respect to the iterations:

$$\overline{k} = \left(\frac{S+m}{S}\right)^{i+1} \tag{15}$$

and so this criterion is not fulfilled. Again, S is the segmentation and i represents the iterations, while $S < (S+m) < S^2$. Nevertheless, one can also see in equation (15) that the average degree does not significantly change if i and m are kept in certain ranges. For instance, set $S = 10$ and consider complete masks with less than 15 edges ($m \in \{1, 2, 3, 4\}$). The average degree before the first iteration is less than 1.5. After $i = 6$ iterations we have strongly connected digraphs established by $n = 10^6$ nodes, while the average degree is still less than 5 and the average of the shortest paths scales with $\log(n)$. Thus, this specific example shows that $\overline{k}$ does not change significantly over the first six iterations while the number of nodes n rises immensely ($n \in \{10, 100, 1000, 10^5, 10^6\}$) and yet the shortest paths scale with $\log(n)$. Therefore, according to the given definition, it can be argued that these digraphs have small-world properties.

Another characterization of small-world networks is based on the clustering coefficient [6]. In random graphs $G_{p,n}$ the clustering coefficient tends to be $O(n^{-1})$ for large n (number of nodes) and small p (probability that two nodes are connected), while small-world graphs are characterized by a $O(1)$-relation [6]. In other words, in small-world networks the clustering coefficients remain constant, which is obviously not the case for our digraphs (see Figure 7). Again, the decrease of the clustering coefficient scales logarithmically. For a specific application this could mean that the change is not significant. Therefore, under certain circumstances our digraphs might be considered as small-world networks.

The last characterization of small-world networks we discuss relates the clustering coefficient to random graphs, as done in equation (6). The clustering coefficients of small-world networks tend to be considerably higher than for random graphs of the same size (i.e., equal number of nodes and edges) [15]. We already discussed this relation and the results are plotted in Figure 8 for $S = 3$. It can be seen in our diagram that there are digraphs of fractal dimensions having clustering coefficients that are magnitudes larger than those of the corresponding random graphs. We also argued previously that this relation correlates to the number of connections, meaning that the lower the fractal dimension, the larger the coefficient in relation to random graphs. Furthermore, the diagrams in Figures 8 and 9 indicate that for some digraphs of fractal dimension the clustering coefficient increases much faster with respect to the number of nodes than it does for random graphs. Therefore, it can be argued that strongly connected

https://doi.org/10.25088/ComplexSystems.19.1.45

digraphs of small fractal dimension have small-world properties, because their clustering coefficient is much larger than for random graphs.

Unfortunately, the lower the fractal dimension, the less the absolute value of the clustering coefficient (see Figure 9). The classical Watts-Strogatz networks for generating small-world networks start with a clustering coefficient of 0.5 [6, 15, 17]. Whether or not this value has to be considered as a threshold for networks with small-world properties is out of the scope of this paper.

In summary, our fractal digraphs do not belong to the class of small-world networks in general. However, for specific parameter settings such network properties can be expected to emerge.

5.4 From Carpets to Neural Networks

In general, there are two strategies for the implementation of ANNs based on the introduced method (see Figure 11). First, the adjacency matrix/the graph can directly be interpreted as a neural network containing recurrences of any kind. Second, the adjacency matrix can be seen as a feed-forward projection between two neuron layers of the same size. Here, the term "size" refers to the number of neurons in a layer. Obviously, a chain of different feed-forward connections can be constructed in that way.

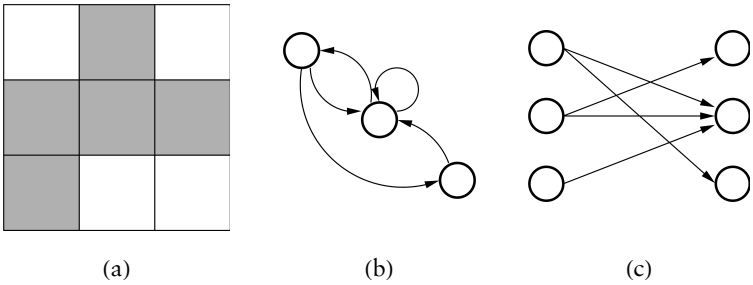


Figure 11. Two ways of transforming a given adjacency matrix (left) into ANNs. First, the graph is directly interpreted with recurrent neural connections (middle). Second, the adjacency matrix as a description of a feed-forward network (right).

The recurrent case might be an interesting method for the generation of reservoirs of nonlinear dynamics. Based on random graphs, this has been done in the echo-state [18] or liquid-state-machine [19] approach.

The latter case might become an object of investigation within the neural Darwinism approach to the function of the brain [20]. According to this approach, an essential element for the brain function is to match specific signal configurations to neural groups that respond in a specific manner. Obviously, this matching must be sufficiently spe-

https://doi.org/10.25088/ComplexSystems.19.1.45

cific in order to allow distinction among different signals, called recognition. However, more important within the neural Darwinism approach is the argumentation that such a matching must be degenerated. The assumption is that there is more than one way to recognize a signal, that is, one signal configuration activates different neural groups as well as one neural group can be activated by different signal configurations. Two extremes of degeneration can be distinguished: a nondegenerated (unique) matching on one side and the completely degenerated matching on the other side. The neural Darwinism approach claims that the variability of brain functions occurs within neural organization somehow located between these two extremes of non- and complete degeneration.

It is interesting to see that the introduced digraphs of fractal dimension create networks between these two extremes. The examples shown in Figure 12 represent only a simple schema. But it is not hard to imagine that the fractal dimension and degree distribution of a graph determine the level of degeneration. Therefore, we argue that within the neural Darwinism approach the introduced digraphs of fractal dimension might be a promising substrate for future research in order to model brain-like mechanisms of adaptation.

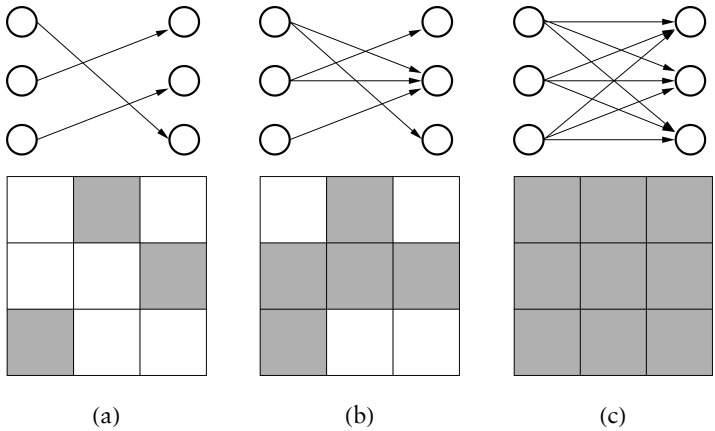


Figure 12. Three examples of feed-forward connections between two neural layers of an ANN. (a) Shows a nondegenerated matching between input and output signal. Each neuron in the left layer only activates one neuron in the right layer. Such a projection is established by a ring. (b) Shows a degenerated matching formed by a digraph of fractal dimension. (c) Shows a completely degenerated matching where each neuron on the left activates each neuron on the right. This projection is based on a fully connected digraph.

https://doi.org/10.25088/ComplexSystems.19.1.45

6. Conclusion

In this paper we have introduced a method, called the digraph generating process (DGP), for the deterministic generation of directed and strongly connected graphs. The DGP is inspired by fractal sets called Sierpinski carpets. Strongly connected digraphs result from structures that correspond to Sierpinski carpets of fractal dimensions between 1 and 2. Therefore, we refer to them as strongly connected digraphs with fractal dimension.

Exclusively aiming for strongly connected digraphs, we introduced a general criterion allowing the design of masks that guarantee the generation of strongly connected digraphs. For better or for worse, this criterion does not cover all possible masks generating strongly connected digraphs. At this point, we must be content to leave this problem to future investigations. However, we have outlined that the given criterion is valid for all mask segmentations and therefore, it provides a huge number of masks and digraphs.

We have shown that the DGP creates complex network structures very distinct to random graphs and that the average of the shortest paths scale logarithmically with respect to the number of nodes. Additional essential properties of complex networks (e.g., average degree, degree distribution, clustering coefficient, and robustness) can be derived directly from the structure of the mask, the initial structures of the DGP. Furthermore, conditions were discussed under which scale-free or small-world networks can be expected to emerge.

The simplicity and the deterministic character of the DGP support an application of this method in many areas related to complex network research. Here, our emphasis was on the domain of artificial neural networks (ANNs). Indeed, more and more models of complex information processing based on ANNs take into account complex network structures [4]. We think the introduced digraphs of fractal dimension provide a promising framework for this context, not only because they match real-world network properties better than random graph structures, but mainly because the DGP provides a variety of complex network structures generated in a systematic and reproducible way. And therefore, the global performance of these complex networks can directly be compared according to their global structural properties, such as fractal dimension, average degree, degree distribution, clustering coefficient, robustness, and so forth.

Acknowledgment

The author would like to thank Dr. Vassilis C. Mavron and Dr. Frank Pasemann for fruitful comments on the topic of this paper and gratefully acknowledges the support of EPSRC through grant EP/C516303/1.

https://doi.org/10.25088/ComplexSystems.19.1.45

References

[1] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing*, Cambridge, MA: MIT Press, 1986.

[2] J. L. Elman, "Finding Structure in Time," *Cognitive Science*, **14**, 1990 pp. 179-211.

[3] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, Redwood City, CA: Addison-Wesley Pub. Co., 1991.

[4] W. Chen, R. Adams, L. Calcraft, V. Steuber, and N. Davey, "Connectivity Graphs and the Performance of Sparse Associative Memory Models," in *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN 2008)*, Hong Kong, IEEE, 2008 pp. 2742-2749.

[5] B. Bollobás, *Random Graphs*, London: Academic Press Inc., 1985.

[6] M. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, **45**(2), 2003 pp. 167-256.

[7] M. A. Arbib, *The Metaphorical Brain 2*, New York: John Wiley and Sons, Inc., 1989.

[8] M. Ito, K. Noda, Y. Hoshino, and J. Tani, "Dynamic and Interactive Generation of Object Handling Behaviors by a Small Humanoid Robot using a Dynamic Neural Network Model," *Neural Networks*, **19**(3), 2006 pp. 323-337.

[9] S. Wermter, G. Palm, and M. Elshaw (eds.), *Biomimetic Neural Learning for Intelligent Robots*, Heidelberg, Germany: Springer, 2005.

[10] D. R. W. Barr, P. Dudek, J. M. Chambers, and K. Gurney, "Implementation of Multi-Layer Leaky Integrator Networks on a Cellular Processor Array," in *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN 2007)*, Orlando, FL, IEEE, 2007 pp. 1560-1565.

[11] R. Pfeifer and C. Scheier, *Understanding Intelligence*, Cambridge, MA: MIT Press, 1999.

[12] B. B. Mandelbrot, *The Fractal Geometry of Nature*, San Francisco: W. H. Freeman, 1983.

[13] H. Eves, *Elementary Matrix Theory*, Dover Publications, 1980.

[14] M. Hülse, "Generating Complex Connectivity Structures for Large-Scale Neural Models," in *Artificial Neural Networks Proceedings, Part II (ICANN 2008)*, Prague, Czech Republic (G. Goos, J. Hartmanis, and J. van Leeuwen, eds.), Berlin: Springer, 2008 pp. 849-858.

[15] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Chichester, UK: Princeton University Press, 1999.

[16] K. A. Lehmann and M. Kaufmann, "Random Graphs, Small-Worlds and Scale-Free Networks" in *Peer-to-Peer Systems and Applications* (R. Steinmetz and K. Wehrle, eds.), Berlin: Springer, 2005 pp. 57-76.

[17] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford, UK: Oxford University Press, 2003.

https://doi.org/10.25088/ComplexSystems.19.1.45

[18] H. Jaeger and H. Haas, "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science*, **304**(5567), 2004 pp. 78-80.

[19] W. Maass, T. Natschläger, and H. Markram, "Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations," *Neural Computation*, **14**(11), 2002 pp. 2531-2560.

[20] G. M. Edelman, *Neural Darwinism*, Oxford, UK: Oxford University Press, 1989.

https://doi.org/10.25088/ComplexSystems.19.1.45