We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Controllable Subspace as a New Characterization of Influence of Nodes in Complex Networks

Jiuhua Zhao

Additional information is available at the end of the chapter

Abstract

In this chapter, we investigate the influence of a node on a network. By virtue of the classical control theory, the influence of a node is represented by the controllable subspace, which is further transformed into a specified graph named general-cacti. An algorithm is developed to calculate the influence, which contains the searching of different kinds of circles and the longest path in the directed graph. Moreover, eight real networks are studied and simulations show that (1) a node in dense and homogeneous networks could have more influence comparing with nodes in sparse and heterogeneous networks, (2) any single studied classical centrality measure including output degree, betweenness and PageRank could not rank the influence and (3) a node with high rank in all these three measures could have more influence.

1. Introduction

With the development of society and technology, nowadays lots of industry and manmade systems have become larger and larger. Different from traditional systems, these systems have a large amount of state variables and complex connection between states. Usually the variables, which capture the connection between states, are unable to measure and at the same time these variables can be easily changed for reasons of temperature, humidity and so on. Hence, these complex systems are difficult to analyze by using classical control theory. In the 1970s, a new control conception, structural controllability, was first introduced by Ching-Tai Lin in 1974 [1]. In Ref. [1] the complete structural controllability of systems with single input was studied. Shields and Pearson [2] extended the result of Lin to multiple input structural systems. Glover and Silverman [3] simplified the work of Shields and Pearson. The conception of structural controllability allows part of systems of the same structure, though with measure

of zero, be uncontrollable. Mayeda and Yamada [4] introduced the conception of strong structural controllability to eliminate this kind of uncontrollability. Reinschke et al. [5] studied the conception in multiinput and multioutput systems. Jarczyk et al. [6] revisited the previous results and gave out a new graph-theoretic characterization of strong structural controllability. The solid foundation of structural controllability had been laid down from then on. For those structural uncontrollable systems, the structural controllable subspace is studied by Hosoe [7] and Poliak [8]. All these studies shed light on the study on the systems where the variables, which capture the connections between states, are unmeasured or mutable. When the system becomes larger, the methods and algorithms mentioned by the above literature studies become inapplicable for the reason of complexity of computation. Recently, Liu et al. [9] combined the structural controllability theory with the complex network, by using the algorithm of maximum match to calculate the minimum input nodes for complete control and methods of complex network science to analyze the properties of the input nodes. Since then more and more studies have been dedicated to controllability of complex network. Nepusz and Vicsek [10] used edge dynamics to describe the controllability of complex networks. Banerjee and Roy [11] exploited the property of driver nodes and made some supplement for the result of Liu et al. All the existing literature studies about the controllability of complex networks are focused on complete controllability. How about those complex networks, who are not controllable with the given diver nodes? Among hundreds even thousands or billions nodes, what role does a single node actually play? How many nodes will it affect, if it is chosen as a driver node? What properties do those nodes with powerful control ability have? All these questions motivated us to study the problem when networks are not fully controlled.

The chapter is organized as follows. In Section 2, some useful preliminaries and the model descriptions are introduced. In Section 3, an algorithm to search for the influence of a single node is proposed and we apply it to several real complex networks in Section 4. Finally, Section 5 contains concluding remarks.

2. Preliminaries and model description

Consider a linear time-invariant control system

$$\dot{x} = Ax(t) + Bu(t) \tag{1}$$

where $x(t) \in R^n$, $u(t) \in R^m$ and we assume that A and B are structural matrices, which means all elements in the matrix are either fixed (zero) or unfixed (indeterminate) and the indeterminate parameters are unrelated. Hereafter Eq. (1) is called a structural system and denoted by (A, B). Let $[AB]$ denote the matrix of the system and $G(V, E)$ denote the digraph of the system. We generate a digraph $G(V, E)$ from (A, B), with $n + m$ nodes, $V = v_1, ..., v_n, v_{n+1}, ..., v_{n+m}$, where $v_1, ..., v_n$ represent the state nodes of (A, B) and $v_{n+1}, ..., v_{n+m}$ represent the control nodes. And all those edges are generated as follows: for every fixed entry m_{ij} of $[AB]$, the graph contains an oriented edge (v_i, v_j). For a structural system (A, B), his corresponding fixed system is denoted by $(\hat{A}, \hat{B})$, where unfixed parameters in (A, B) are filled by one class of unrelated real numbers.

Definition 1: A system (A, B) is structural controllable if all its possible fixed systems $(\hat{A}, \hat{B})$ are controllable.

The above definition was first introduced by C-T Lin. He reduced the property of structural property to the property of the graph of the system.

Definition 2 (stem) [1]: A stem is an elementary path originating from an input vertex (see **Figure 1**).

Definition 3 (bud) [1]: A bud is an elementary cycle with an additional edge that ends, but not begins, in a vertex of the cycle. Vertex $v1$ is called the origin of the bud (see **Figure 2**).

Definition 4 (cactus) [1]: A cactus is defined recursively as follows. A stem is cacti. Given a stem S_0 and buds $B_1, B_2, \cdots, B_p$, then $S_0 \cup B_1 \cup \cdots \cup B_p$ is a cacti if for every $i(1 \leq i \leq p)$ the origin vertex of B_i is also the origin of an edge in $S_0 \cup B_1 \cup \cdots \cup B_i$ and is the only vertex belonging the same time to B_i and $S_0 \cup B_1 \cup \cdots \cup B_i$. A set of disjoint cacti is called cactus (see **Figure 3**).

We call the red node in **Figure 3**, i.e., the input vertex, driver node and the nodes connected to driver node are defined as driven nodes(e.g., node $d1$ and $d2$ in **Figure 3**.
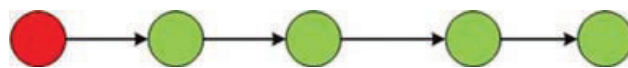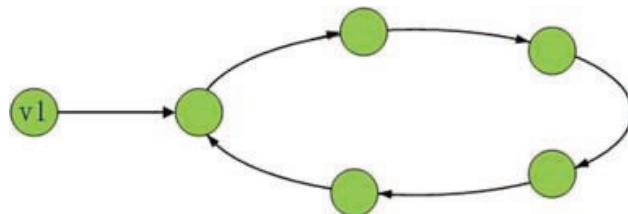
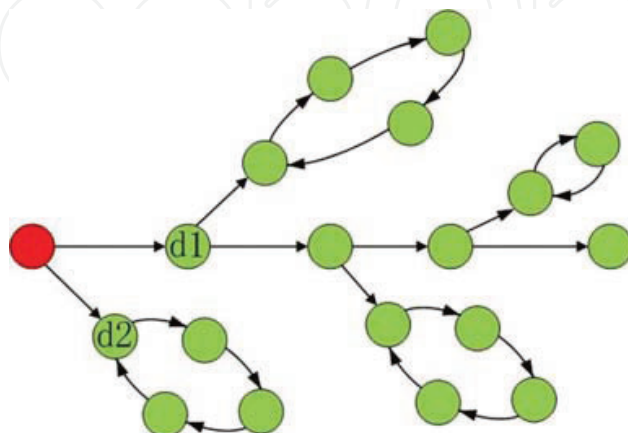

Figure 1. Stem.

Figure 2. Bud.

Figure 3. Cacti.

Theorem 1 (Lin) [1]: The single input system pair (A, b) is structurally controllable iff the graph of (A, b) is spanned by cacti.

Many works have been done to study the complete controllability of a system. Hosoe considered the problem for the first time, where the system is not fully controllable. He defined the maximal dimension of controllability matrix of the structural system as the generic dimension of controllable subspace, denoted by $d_c = \text{gennericrank}[A \quad AB \quad \cdots \quad A^{n-1}B]$. And give out two methods for determining the generic dimension of controllable subspace of a system. We show the graph-theoretic method here (see Theorem 1).

Theorem 2 [2]: Let d_c be the generic dimension of the controllable subspace of the structural system (A, B) and $G(A, B)$ represents the graph of (A, B). Assume that $[A\ B]$ is irreducible. Then $d_c = \max_{\tilde{G} \in \mathcal{G}} \{|E(\tilde{G})|\}$ where $\mathcal{G}$ denotes the subgraphs of $G(A, B)$ which is defined by and $|E(\tilde{G})|$ represents the number of edges in graph $\tilde{G}$. Comparing with the graphic properties of complete controllable system in Lin [1], the graph of controllable subspace contains more circles than cactus. Those circles are not directly connected to the cactus in the form of buds. They are connected to the cactus through at least one node. We define the circle in $\tilde{G}$ as a general-bud. Then replacing all the buds in cacti by general-buds, we can have the definition of general-cacti (see **Figure 4**). In **Figure 4**, those circles with dash lines are the defined general-bud.
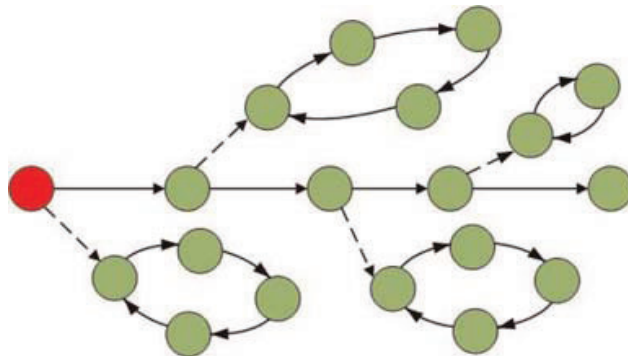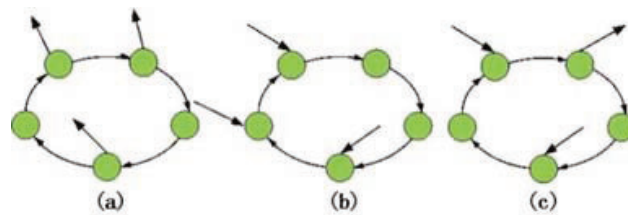

Figure 4. General-cacti.

In our research, we consider the influence of a single node in the complex network. By the influence of node i, we mean the controllable subspace of the system when we only control node i. Then we denote the structural input vector as b_i, where all elements in vector b_i are fixed zero except for the ith element. Therefore, we can just study the controllable subspace of structural system (A, b_i), whose corresponding structural matrix and digraph are denoted as $[A\ b_i]$ and G_i. Let us define the controllability of a node i in a complex network as P_i. Then we can easily derive P_i from Theorem [1], i.e.,

$$P_i = \max_{\tilde{G}_i \in \mathcal{G}_i} \{|V(\tilde{G}_i)|\} \tag{2}$$

where $\mathcal{G}_i$ denotes the set of general-cacti originated from node i in digraph G_i. To ignore the influence of the network size, the influence of node i is normalized as $p_i = P_i/N$.

3. Algorithm

To find the largest general-cacti, we classify the circle in the network into three categories. The first one is circles with no inbound (see **Figure 5a**). This kind of circles cannot be arrived by the input, so they are not in our general-cacti. The second one is circles with no outbound edges (see **Figure 5b**). This kind of circles can only act as general-buds in the general-cacti without affecting the competition of two different general-cactus. The third one is circles with both inbound and outbound edges (see **Figure 5c**). These circles can also be part of a stem, which means these nodes affect the dimension of general-cacti.

Figure 5. Circles of three categories.

Our instinct is exploring the longest path to act as a stem and then add all the circles as the general-buds. Nevertheless, the third kind of circles makes it difficult to search for the largest general-cacti. As shown in **Figure 6**, the blue stem in (b) is longer than the red one in (a), but the circle can also act as a general-bud to increase the dimension of controllable subspace. Hence, the dimension of general-cacti in (a) is larger than the one in (b).
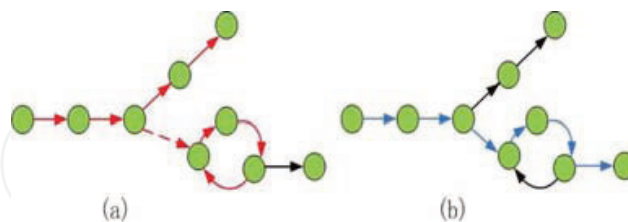
Figure 6. (a, b) Two different cacti in the same network.

However, if the circle of the third kind is not large enough, the choice of the longest path will not affect the influence of the origin node significantly. When the circle is large enough, we set it as a general-bud to prevent it from participating in the stem forming. The definition of "large" affects the accuracy of our method. However, from the real database we find that the circles are either very small or very large. Our method is efficient in real networks. Thus far we can give an approach to discover the influence of a single node in the complex network.

3.1. Procedure to discover the influence of a single node

- Search the reachable nodes from node i in the original graph and delete all those unreachable nodes.

- Search for the second category circles in the current graph, count the number (N_1) of their nodes, then delete all those nodes and the edges connected to them.

- Search for large circles in the current graph, count the number (N_2) of their nodes, then delete all those nodes and the edges connected to them.

- Search for the longest path originate from the input in the current graph, count the number (N_3) of the nodes, then delete all those nodes and the edges connected to them.

- Search for all the circles left in the current graph, count the number (N_4) of the nodes, then delete all those nodes and the edges connected to them.

- $P_i = N_1 + N_2 + N_3 + N_4$

The first, second and fourth steps of this procedure are all about search for the directed circle in the digraph. We recursively delete the nodes without any input or without any output. Then we can get all the nodes in circle. Using the depth-first search method, we can separate these nodes into different circles and count for the number of nodes of each circle. For the circle, which has small circles inside, we just count the largest circle.

The third step of this procedure is the most difficult one. To the best of our knowledge, there exist no optimal algorithms to search for the longest path in the digraph. The existing algorithms of searching for the shortest path cannot be modified to search for the longest path, for the existing of directed circles. We modify the classical depth-first search method to search for the longest path in a digraph with directed circles (see Algorithm 1).

Algorithm 1 Algorithm to search for the longest path in a digraph

---

Require: $v \neq NULL$

Ensure: $Long(v)$

   $deep \Leftarrow deep + 1$, $visited(v) \Leftarrow 1$,

   $vexdeep(v) \Leftarrow deep$,

   $vex \Leftarrow$ the vertex v directs to

   while $vex \neq NULL$ **do**

   if $visited(v) = 0$ and ($circle(v) \neq 0$ or

   $deep \geq vexdeep(v)$) **then**

     $Long(vex)$

     $visited(v) \Leftarrow 0$

$vex \Leftarrow$ the vertex v directs to

$deep \Leftarrow deep - 1$

else if $visited(v) = 1$ **then**

    for all vertex i in the current circle **do**

        $circle(i) \Leftarrow 1$

    end for

    $vex \Leftarrow$ the vertex which has the same origin as
        vex

else

    $vex \Leftarrow$ the vertex which has the same origin as
        vex

end if

end while

In the algorithm, we use vector *visited* and *circle* to mark the node. Node i is visited (in a circle) if $visited(i) = 1$ ($circle(i) = 1$), otherwise unvisited (not in a circle). We use *deep* to record the current depth and vector *vexdeep* to record the depth of every vertex. The most significant difference between our method and the depth-first search method is that we allow one vertex to be detected more than one time and at the same time we record the circles in the network for searching in other directions. The time complexity of our algorithm is $O(NM)$, where N is the number of nodes in the network and M denotes the number of edges.

4. Simulation results

We now apply the proposed algorithm to several kinds of real networks. It is infeasible to search all the nodes to find out the most influential node, for the unbearable complexity. Hence, we rank the node by their ODCR (outcome-degree centrality rank), BCR (betweenness centrality rank) and PR (page rank), then analysis the first three nodes of every categories. Form **Table 1**, we can find that nodes in dense and homogeneous networks, like network Caenorhabditis elegans, network Seagrass, network s838, network s420 and network s208, can control more percentage of nodes than nodes in sparse and heterogeneous networks. This phenomenon implies that nodes are more influential in dense and homogeneous networks. We think it is reason why dense and homogeneous networks need less driver nodes for full structural control which is observed in [9].

| Type	Regulatory			Neuronal	Food web	Electronic circuits		
Name	TRN-Yeast-1 [12]	TRN-Yeast-2 [12]	TRN-EC-2 [12]	Caenorhabditis elegans [13]	Seagrass [14]	S838 [12]	S420 [12]	S208 [12]
N	4441	688	418	297	49	512	252	122
L	12873	1079	519	2345	226	819	399	189
n_D	96.5%	82.1%	75.1%	16.5%	26.5%	23.2%	23.4%	23.8%
p_{ODR1}	0.97%	0.29%	0.96%	58.59%	24.49%	31.64%	32.54%	34.43%
p_{ODR2}	1.10%	0.44%	0.48%	58.59%	22.45%	2.73%	31.35%	24.60%
p_{ODR3}	1.15%	0.29%	0.72%	56.9%	16.33%	26.95%	3.17%	31.97%
p_{BCR1}	1.15%	0.73%	0.96%	0.34%	24.49%	31.64%	3.17%	5.74%
p_{BCR2}	1.10%	0.44%	1.20%	58.59%	14.29%	7.42%	3.17%	34.43%
p_{BCR3}	1.04%	0.29%	0.48%	58.59%	10.20%	1.76%	3.57%	18.03%
p_{PR1}	1.15%	0.29%	0.96%	0.34%	10.20%	31.64%	32.54%	34.43%
p_{PR2}	1.10%	0.44%	0.48%	58.59%	14.29%	2.34%	31.35%	5.74%
p_{PR3}	0.97%	0.44%	0.72%	58.59%	12.24%	2.73%	5.95%	12.30%

Table 1. The influence of the studied nodes in eight real networks.

From **Figure 7**, we find influence of node does not rank according to any single studied centrality measure. In fact, in network Yeast-1, nodes ODR3, BCR1 and PR1 are the same node and it has largest influence 1.15% among the considered nodes. In network Yeast-2, nodes ODR2, BCR2 and PR2 are the same node and it has the largest influence 0.44%. In network EC-2, nodes ODR1, BCR1 and PR1 are the same node and it has the largest influence 0.96%. In network elegans, nodes ODR1, BCR2 and PR2 are the same node and it has the largest influence 58.59%. In network Seagrass, nodes ODR1 and BCR1 are the same node and it has the largest influence 24.49%. In network s838, nodes ODR1, BCR1 and PR1 are the same node and it has the largest influence 31.64%. In network s240, nodes ODR1 and PR1 are the same node and it has the largest influence 32.54%. In network s208, nodes ODR1, BCR2 and PR1 are the same node and it has the largest influence 34.43%. Hence, we can conclude that if a node has a high ODCR, BCR and PR at the same time, this node can have lager influence.

5. Conclusions

In this chapter, we have introduced the notion of the influence of a node in complex networks and gave out a useful algorithm to count for that property. Meanwhile we applied this method to exploit the influence of nodes in some real networks. Results have shown that nodes in dense and homogeneous networks were more influential than nodes in sparse and heterogeneous networks. And in the same network node

which has high rank in ODR, BCR and PR simultaneously could be more influential. Next we will try to study the characteristic of the most influential node in a complex network and the influence of a set of nodes.
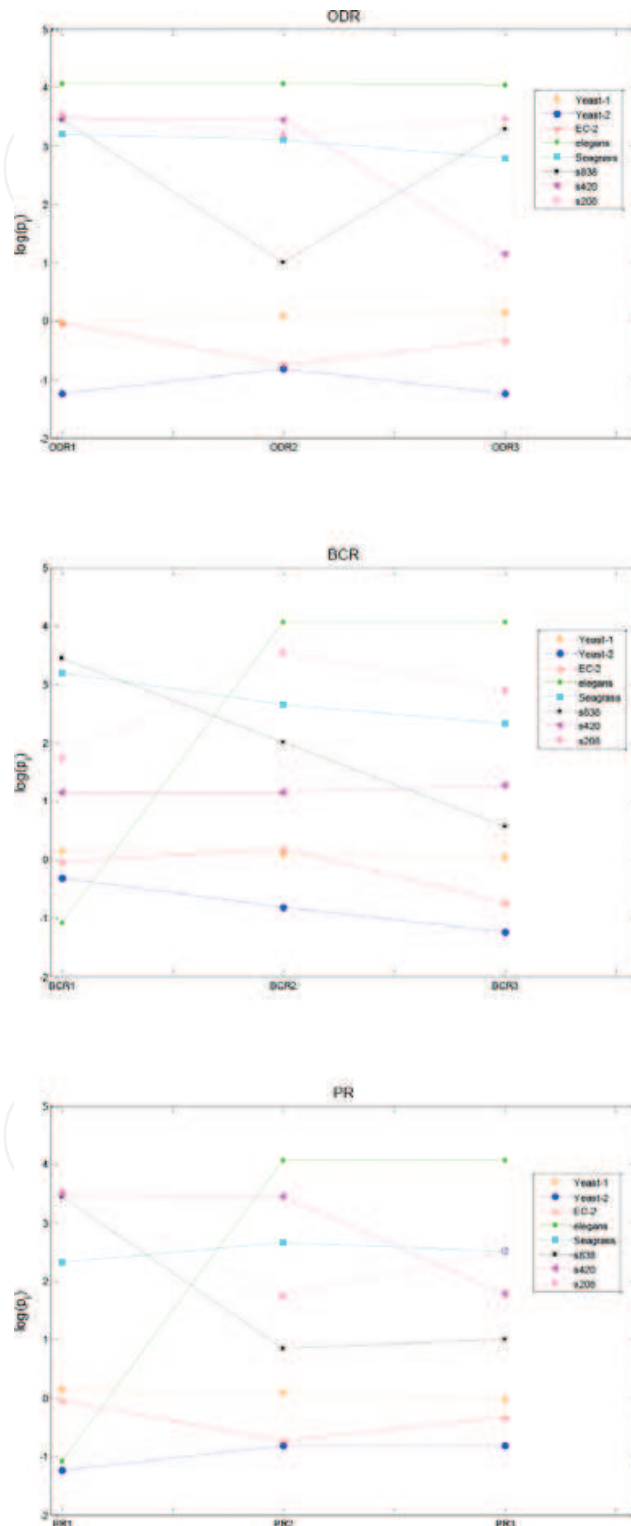
Figure 7. The influence of top three nodes according to each centrality measure in eight real networks.

Author details

Jiuhua Zhao

Address all correspondence to: jiuhuadandan@sjtu.edu.cn

Department of Automation, Shanghai Jiao Tong University and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China

References

[1] C.-T. Lin, "Structural controllability," IEEE Transactions on Automatic Control, vol. 19, no. 3, pp. 201–208, 1974.

[2] R. W. Shields and J. B. Pearson, "Structural controllability of multiinput linear systems," IEEE Transactions on Automatic Control, vol. 21, no. 2, pp. 203–212, 1976.

[3] K. Glover and L. M. Silverman, "Characterization of structural controllability," IEEE Transactions on Automatic Control, vol. 21, no. 4, pp. 534–537, 1976.

[4] H. Mayeda and T. Yamada, "Strong structural controllability," SIAM Journal on Control and Optimization, vol. 17, no. 1, pp. 123–138, 1979.

[5] K.J. Reinschke, F. Svaricek and H.D. Wend, "On strong structural controllability of linear systems," in Proceedings of the 31st Conference on Decision and Control, pp. 203–208, 1992.

[6] J. C. Jarczyk, F. Svaricek and B. Alt, "Strong structural controllability of linear systems revisited," in Proceedings of the 50th Conference on Decision and Control and European Control Conference, pp.1213–1218, 2011.

[7] S. Hosoe, "Determination of generic dimensions of controllable subspace and its application," IEEE Transactions on Automatic Control, vol. 25, no. 6, pp. 1192–1196, 1980.

[8] S. Poliak, "On the generic dimension of controllable subspace," IEEE Transactions on Automatic Control, vol. 35, no. 3, pp. 367–369, 1990.

[9] Y.-Y. Liu, J.-J. Slotine and A.-L. Barabasi, "Controllability of complex networks," Nature, vol. 473, no. 7346, pp. 167–173, 2011.

[10] T. Nepusz and T. Vicsek, "Controlling edge dynamics in complex networks," Nature Physics, vol. 8, pp. 568–573, 2012.

[11] S. J. Banerjee and S. Roy, "Key to network controllability," vol. http://arxiv.org/abs/1209.3737, 2012.

[12] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." Science, vol. 298, no. 5594, pp. 824-827, 2002.

[13] Modha, Dharmendra S. and Raghavendra Singh. "Network architecture of the long-distance pathways in the macaque brain." Proceedings of the National Academy of Sciences, vol. 107, no. 30, pp. 13485–13490, 2010.

[14] Christian, Robert R. and Joseph J. Luczkovich. "Organizing and understanding a winter's seagrass foodweb network through effective trophic levels." Ecological Modelling, vol. 117, no. 1, pp. 99–124, 1999.