Modeling complex biological systems using Scale Free Networks

RADU DOBRESCU
POLITEHNICA University of Bucharest, ROMANIA
Splaiul Independentei 313, Faculty of Control & Computers

Abstract: - The paper proposes a model that relieves the characteristics of several complex systems having a similar scale free network architecture. The properties of this kind of networks are compared with those of other methods which are specific for studying complex systems: nonlinear dynamics and statistical methods. We place particular emphasis on scale free network theory and its importance in augmenting the framework for the quantitative study of complex biological systems. The advantages and limits of this model in understanding the structure of cellular signaling networks involved in Sepsis phenomena are finally discussed.

Key-Words: - complexity, scale free networks, modeling; self similarity; cellular signaling, sepsis

1 Introduction

The definition of complexity, or more exactly of a complex system, should be richer than that of algorithmic complexity, and should express the level of interconnectedness and interdependencies of a system, not just the instruction set for creating the system - the amount of effort it takes to use those instructions must addressed as well. In a complex system it is often the case that the utility of a structure or process is expressed at the next higher level of organization relative to the process itself. Unlike entropy and the related concept of information, complexity is not extensive, nor is it entirely intensive. What is clear though is that complexity concerns a specific description, which is of course dependent on the technology and subjective capabilities of the observer. Anyway, we can consider that a complex system is a system with a large number of elements, building blocks or agents, capable of interacting with each other and with their environment. The interaction between elements may occur only with immediate neighbors or with distant ones; the agents can be all identical or different; they may move in space or occupy fixed positions, and can be in one of two states or of multiple states. The common characteristic of all complex systems is that they display organization without any *external* organizing principle being applied.

Another problem is how to measure the complexity. The first and still classic measure of complexity is that introduced by Kolmogorov, which is (roughly) the shortest computer program capable of generating a given string. This quantity is in general uncomptable, in the sense that there is simply no algorithm which will compute it. Moreover, the Kolmogorov complexity is maximized by random strings, so it's really telling us what's random, not what's complicated, and it's gradually come to be called the "algorithmic information." Let also remind that Bennett proposed a measure for the computing complexity, which he called the logical depth of a system [1]. The basic idea is that a system should be called complex, or *logically deep*, if that system can be generated by a few simple rules, but those rules require a long time to run. So, for example, a human body is complex in that it is specified by a relatively small amount of information encoded in DNA, but it takes a great deal of processing to get from that DNA to the human body.

So, if we can not define exactly the complexity and also we can not measure its dimensions, let try to model at least the complex systems. Because we met more and more example of complex systems.The stock market, cities, metabolic pathways, ecosystems, the Internet or the human brain, are all complex. We can ask what do have in common all these systems. In the last few years the answer that has emerged is that they all share similar network architectures. Network theory has become one of the most visible pieces of the body of knowledge that can be applied to the description, analysis, and understanding of complex systems and is now an essential ingredient their study. But it is not the only solution. Actually, as will be shown in the next section, network theory competes with other two complex systems modeling methods..

2 Tools for complex systems models

In a rough sense, the current toolbox used in tackling complex systems involves three main areas: (i) nonlinear dynamics and chaos, (ii) statistical physics, including discrete models, and (iii) network theory.

2.1 Nonlinear dynamics and chaos

Nonlinear dynamics and chaos in deterministic systems are now an integral part of science and engineering. The theoretical foundations are well agreed upon mathematical definitions of chaos, many of them formally equivalent. However, because of its relative novelty and, in much case, counterintuitive nature, there are still many misconceptions about chaos and its implications. Extreme sensitivity to initial conditions does not mean that prediction is impossible. Memory of initial conditions is lost within attractors but the attractor itself may be extremely robust. In particular chaotic does not mean unstable. Chaos means that simple systems are capable of producing complex outputs. Many techniques have been developed to analyze signals and to determine if fluctuations stem from deterministic components. There are numerous applications in geophysics, physiology and neurophysiology [2].

2.2 Statistical physics

Statistical physics brought three very important conceptual and technical advances: 1. It lead to a new conception of prediction; 2. It circumvented classical mechanics and it casted solutions in terms of ensembles; 3. It introduced the concept of discrete models, ranging from the cellular automata to agent-based models. In the 1970s, fundamental advances occurred in our understanding of phase transitions and critical phenomena leading to the development of two important new concepts: *universality* and *scaling* [3]. The finding, in physical systems, of universal properties that are independent of the specific form of the interactions gives rise to the intriguing hypothesis that universal laws or results may also be present in complex social, economic and biological systems. The *scaling hypothesis* which arose in the context of the study of critical phenomena led to two categories of predictions, both of which have been remarkably well verified by a wealth of experimental data on diverse systems. The first category is a set of relations, called *scaling laws* that serve to relate the various critical-point exponents characterizing the singular behavior of the order parameter and of response functions. The second category is *data collapsing*. Another fundamental concept arising from the study of critical phenomena is *universality*. For systems in the same universality class, exponents and scaling functions are the same in the vicinity of the critical point. This fact suggests than when studying a given problem, one may pick the most tractable system to study and the results one obtains will hold for all other systems in the same universality class. Fractal analysis seems to be one of the most promising tools.

In what concern *discrete models* the main assumption in that some phenomena can and should be modeled directly in terms of computer programs (algorithms) rather than in terms of equations. Cellular automata are the simplest example of discrete time and space models that were developed with the computer in mind. Examples of the application of cellular automata exist in physical, chemical, biological and social sciences; they can be as simple as elementary predator-prey models between a handful of species and as complex as the evolution of artificial societies. Discrete, or agent-based, modeling has been extremely successful because of the intuition-building capabilities it provides and the speed with which it permits the investigation of multiple scenarios. For this reason discrete modeling has led in some cases to a replacement of equation based approaches in disciplines such as ecology, traffic optimization, supply networks, and behavior-based economics.

2.3 Networks

The third element in the toolbox is networks. A network is a system of nodes with connecting links. Once one adopts this viewpoint, networks appear everywhere. Consider some examples from two main fields: a) biological networks: autonomous nervous systems of complex organisms, a network of neurons connected by synapses, gene regulation networks, a network of genes connected by cross-regulation interactions or metabolic networks, a network of metabolites connected by chemical reactions, b) social networks, like e-mails services, Internet and the World Wide Web. The structure of such social networks was formalized exactly by using *random graphs*, in which the existence of a link between any pair of nodes has probability *p*. Erdos, in collaboration with Renyi, pursued the theoretical analysis of the properties of random graphs obtaining a number of important results, including the identification of the percolation threshold, that is, the average number of links per node necessary in order for a random graph to be fully connected, or the typical number of intermediate links in the shortest path between any two nodes in the graph. Another important class of network is represented by the *Small-world networks,* that have as main

characteristic the so-called small-world phenomenon, which is defined by the co-existence of two apparently incompatible conditions, (i) the number of intermediaries between any pair of nodes in the network is quite small - typically referred to as the six-degrees of separation phenomenon and (ii) the large local redundancy of the network, i.e., the large overlap of the circles of neighbors of two network neighbors. The latter property is typical of ordered lattices, while the former is typical of random graphs.

Models of biological systems, especially of biochemical networks have been developed to make statistical predictions of the types of dynamics based on network topology and interaction bias [4]. For values of mean connectivity chosen to correspond to real biological networks, these models predict disordered dynamics and offer a direct connection to the chaos approach. However, chaotic dynamics seems to be absent from the functioning of a normal cell. While these models use a fixed number of inputs for each element in the network, recent experimental evidence suggests that several biological networks have distributions in connectivity. We therefore study networks with distributions in the number of inputs, K, to each element according to a power law (scale free). The topology of scale-free biochemical networks, characterized by a wide distribution in the number of inputs per element, may provide a source of order in living cells. Recently, Watts and Strogatz [5] proposed a minimal model for the emergence of the small-world phenomenon in simple networks. In their model, small-world networks emerge as the result of randomly rewiring a fraction p of the links in a d-dimensional lattice (Fig. 1). The parameter p enables one to continuously interpolate between the two limiting cases of a regular lattice (p = 0) and a random graph (p = 1).

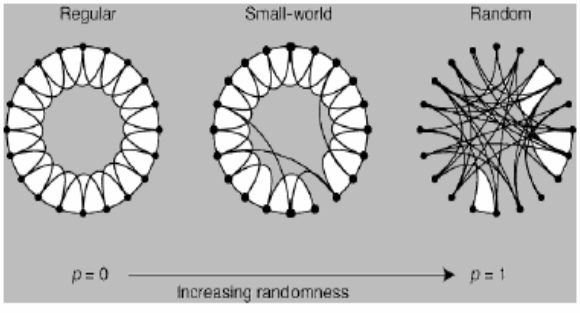

Fig.1. Small-world networks model generation

Watts and Strogatz probed the structure of their small-world network model via two quantities: (i) the mean shortest distance L between all pairs of nodes in the network, and (ii) the mean clustering coefficient C of the nodes in the network. For a d-dimensional lattice one has $L \sim N^{1/d}$ and $C = O(1)$, where N is the number of nodes in the network; for a random graph one has $L \sim \ln N$ and $C \sim 1/N$.

3 Scale-free networks

An important characteristic of a graph that is not taken into consideration in the small-world model of Watts and Strogatz is the degree distribution, i.e., the distribution of number of connections of the nodes in the network. The Erdos-Renyi class of random graphs has a Poisson degree distribution, while lattice-like networks have even more strongly peaked distributions - a perfectly ordered lattice has a delta-Dirac degree distribution. Similarly, the small-world networks generated by the Watts and Strogatz model also have peaked, single-scale, degree distributions, i.e., one can clearly identify a typical degree of the nodes comprising the network. Against this theoretical background, Barabasi and coworkers found that a number of real-world networks have a scale-free degree distribution with tails that decay as a power law [6]. These networks were called Scale Free Networks (SFN).

3.1. General SFN properties.

This Barabasi and Albert suggested that scale-free networks emerge in the context of growing network in which new nodes connect preferentially to the most connected nodes already in the network. Specifically,

$$p_i(n+1) = \frac{k_i(n)}{\sum_{-n_0+1}^{n} k_i(n)}$$

where n is the time and number of nodes added to the network, n_0 is the number of initial nodes in the network at time zero, k_i is the degree of node i and $p_i(n + 1)$ is the probability of a new node, added at time n+1 linking to node i.

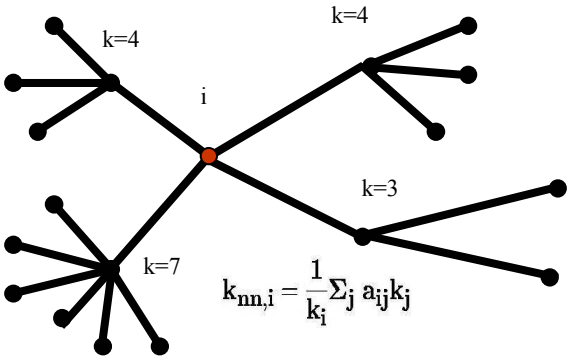

Fig.2. A scale free network graph

As illustrated in Figure 2, as time ticks by the degree distribution of the nodes becomes more and more heterogeneous since the nodes with higher degree are the most likely to be the ones new nodes link to. Significantly, scale-free networks provide extremely efficient communication and navigability as one can easily reach any other node in the network by sending information through the "hubs", the highly-connected nodes. The efficiency of the scale-free topology and the existence of a simple mechanism leading to the emergence of this topology led many researchers to believe in the *complete* ubiquity of scale-free network. Note that scale-free networks are a subset of all small-world networks because (i) the mean distance between the nodes in the network increases extremely slowly with the size of the network and (ii) the clustering coefficient is larger than for random networks.

3.2. Diameter of scale-free networks

It was shown that scale-free networks with degree exponent $2<\lambda<3$ possess a diameter $D \sim \ln \ln N$, smaller even than that of random and small world networks [7]. If the network is fragmented, we will only be interested in the diameter of the largest cluster (assuming there is one). In this study we consider the diameter of a Molloy-Reed scale-free network definite as the *average* distance between any two sites on the graph. Actually, it easier still to focus on the radius of a graph, $L \equiv \langle l \rangle$ as the average distance of all sites from the site of highest degree in the network. The diameter of the graph D is restricted to $L \leq D \leq 2L$ and thus scales like L .

3.3. Minimal graphs and lower bound

Cohen, *et al*., show that the radius of any scale-free graph with $\lambda>2$ has a rigorous lower bound that scales as ln ln N. It is easy to convince oneself that the smallest diameter of a graph, of a given degree distribution, is achieved by the following construction: Start with the highest degree site, then connect to each successive layer the extant sites of highest degree, until the layer is full. By construction loops will occur only in the last layer [8]. To bound the radius L of the graph, we will assume that the low degree sites are connected randomly to the giant cluster. On the other hand, if we start uncovering the graph from any site - provided it belongs to the giant component – then within a distance l_2 from this site there are at least l_2 bonds. Since $l= l_1+l_2$, all sites are at a distance of order ln ln N from the highest degree site, and $L=$ ln ln N is a rigorous lower bound for the diameter of scale-free networks with $\lambda>2$. In a similar way one can demonstrate that the scaling of $D \sim \ln \ln N$ is actually realized in the general case of *random* scale-free graphs with $2<\lambda<3$. For $\lambda>3$ and $N>>1$, k is independent of N , and the radius of the network is $L \sim \ln N$.

4 Using SFN in complex biological systems modeling

In order to exemplify the properties of a SFN based complex biological system model we have chosen to model the behavior of the immune system under the action of a Multiple Organ Dysfunction Syndrome (MODS). Because in MODS host-related factors are associated with patient outcome, the focus shifted to the study of host response in trauma, shock and sepsis. The immune system is a highly complex and integrated system which has evolved to provide the organism with substantial defences against pathogenic organisms. In order to perform this function, the immune system has evolved strategies that allow successful elimination of a wide variety of pathogens, including viruses, bacteria and parasites. At its turn, the immune response to pathogens is a complex, highly regulated system involving numerous interactions between different cell types.

4.1. Interactions in the immune systems

There are numerous neural-immune, endocrine-immune and behavioral-immune interactions. For example, proinflammatory cytokines cause activation of hypothalamic-pituitary-adrenal axis. Sympathetic enervation of lymphoid tissue has variable effects, including potentiating B Cell IgM antibody response. Lymphocites bear receptors for substance P, somatostatin, vasoactive intestinal peptide, corticotropine-releasing factor and adrenocorticotropic hormone and all mediate significant immune alteration. Hence the host response to sepsis must be considered as a systemic response, involving multiple interactions between metabolic, neural, endocrine, behavioral and immune processes. In addition to these interactions, the immune system can be considered as a highly interconnected self-reactive nonlinear network. There are two repertoires of antibody variable regions (V-regions and site of antigen binding); one reactive to self, and the other nonreactive to self. There is a high degree of autoreactivity or connectivity between elements in the immune system, specifically between variable regions (V-regions) on antigen-binding sites of immunoglobulin.

The connectivity of the elements of the immune system is related to disease states, which we hypothesize may be secondary to emergent

properties of the immune system. In summary, normal immune function includes a complex non-linear network of self-reactive immunoglobulin; the study of the connectivity of this network may provide valuable information regarding altered emergent properties of the immune system, which lead to autoimmune disease. It has been demonstrated that cytokines have multiple effects and exhibit marked interdependence. and redundancy (multiple cytokines with the same effect). The effects of cytokines may dampen, or amplify the effects of other cytokines. For example, TNF-α and IL-1 act synergistically, whereas IL-10 decreases serum TNF-α and 1L-1 [9]. The effect of a cytokine may vary with its concentration in a nonlinear fashion. For example, TNF-α will cause neutrophilia at low doses, and neutropenia at high doses in healthy humans [10]. Feedback loops, either positive or negative, are everywhere within the cytokine-cellular network. The number of mediators are extremely large, and continue to be identified. The inflammatory cellular and cytokine network truly consists of a complex nonlinear system, which will be named in the following as a protein interaction network (PIN) which can be mathematically represented as graphs whose nodes symbolize proteins and edges connect pairs of interacting proteins. All interaction data sets exhibit a non-trivial topological structure of the networks, showing a broad connectivity distribution P(k); i.e., the probability that any given protein interacts with k other proteins. This feature implies the statistical abundance of ‘‘hubs’’, that is nodes with a large connectivity, and prompt to a complex architecture that has found further support in the non-trivial correlation and hierarchical features observed in the networks topology. At the same time, topological information is being exploited in predictive methods for protein functional assignment and theoretical models are being developed for the formation of PINs.

In order to find some correspondence between the so different approaches for sepsis modeling, we choose the same quality indicator for the immune response, the endotoxin tolerance defined as a reduced capacity of the host (in vivo) or of the cultured macrophage/monocyte (in vitro) to respond to LPS (lipopolysaccharide) activation following a first exposure to this stimulus [11]. At its turn, the best indicator of the endotoxin tolerance was considered the proinflammatory cytokine TNF-α (actually, the functional state of the monocyte). The difference between the TNF-α concentrations at an initial state (steady state) and after a measured interval was defined as Δ cytokine (the term was borrowed from the paper of Flondor and Vasilescu [12]). One can take as a typical antiinflammatory cytokine IL-10 (interleukin-10). For the construction of a network model, we have considered the monocyte, the main actor of the innate immunity as a central node of a network, which can be disconnect from this network when the phenomenon of endotoxin tolerance (marked by a significant value of Δ cytokine) occurs.

4.2. Construction of the model graph

In order to obtain significant results, we start by considering a set of standard metrics of the model graph (the size and the number of interactions of the network, together with the size of the largest component). An important feature of the graph is the average degree, where the degree of a given node is defined as the number of its connections. The average degree is done by $\langle k \rangle = 2l/n$ with l being the total number of links in the graph and n the number of nodes. Another characteristic of the graph local cohesiveness is provided by the clustering coefficient, defined for any vertex (node corresponding to the protein i) as the fraction of connected neighbors of i : $C_i = \frac{2l_i}{k_i(k_i-1)}$ where l_i is the number of links connecting neighbors of i and $k_i(k_i-1)/2$ the number of possible connections among neighbors. A more global quantity for characterizing the graph is the mean clustering coefficient $\langle C \rangle = (1/n)\sum_{1}^{n} C_i$ where the average is over all the *n* proteins in the network. Direct measurements of the degree distribution for metabolic networks show that these nets exhibit a scale-free degree distribution [13]: $P(k)=ck^{-\lambda}$, $k=m,..,K$, where $c \approx (\lambda-1)m^{(\lambda-1)}$ *is* a normalization factor, and *m* and *K* are the lower and upper cutoffs for the connectivity of a node, respectively.

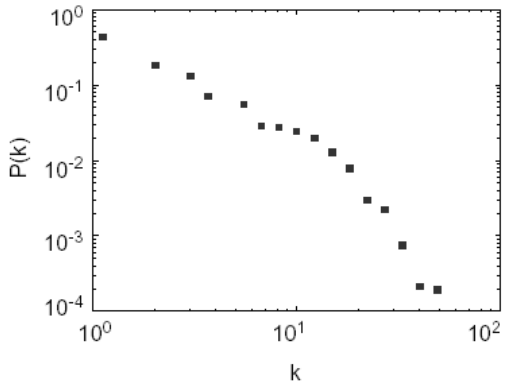

Fig. 3 Degree distribution P(k) for the network model

The degree distributions of our graph model presented in fig.3 correspond to a power-law exponent $\lambda \approx 2,1$ and cut-off $K \approx 30$ and represent the best fit to the data. One can observe that the degree fluctuations are much larger than the average value, that render the system an example of scale-free behavior

The problem to solve is that limited information-processing capabilities have a significant and quantifiable effect on the large-scale structure of growing networks. We find that information filtering leads to an exponential truncation of the in-degree distribution for networks growing under conditions of preferential attachment. We can conclude that our scale-free graph model can express the dynamics of the immune system, especially because the nodes with the largest degree have important roles

5 Conclusions

The need for enhanced computational ability is most evident when one attempts to couple large numbers of individual units into highly interactive and largely parallel networks. The proliferation of information transferred in such networks introduces the need for these systems that provide a framework for classifying information, spatial statistics for analyzing patterns, and dynamic simulation models that allow the integration of information across multiple spatial, temporal, and organizational scales. It is impossible to ignore the apparent universality of pair interactions among the various elements of a complex system. Instead of chance and randomness, one must consider a high degree of internal order that governs the system organization. Each node selected in order to be discussed as an element in a network of interacting constituents, ensures to spot and quantify the interplay between behavior, structure and function. It can be approached from the bottom up, moving from cells to modules, or from the top to the bottom, starting from the network's scale-free and hierarchical nature and moving to the specific modules. In either case, it must be acknowledged that structure, topology, network usage, robustness and function are deeply interlinked. The edge complexity could be reduced in different respects.. For instance a clustering of a network derived from the connectivity distribution of the nodes might show sub-networks of intense communication or the impact of distinct nodes for the whole system. Network modeling, quantitative analysis and laboratory experiments have to be combined in various ways to gain new insights.

References:

[1] C. H. Bennett, How to define complexity in physics, and why, in W. Zurek, ed., *Complexity, entropy and the physics of information*, Addison-Wesley, 1990, pp.137-148

[2] L. Glass, M.C. MacKey, *From Clocks to Chaos: The Rhythms of Life,* Princeton University Press, Princeton, NJ, 1988

[3] H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena,* Oxford University Press, Oxford, 1971

[4] J. Fox, C. Hill - From topology to dynamics in biochemical networks, *Chaos: An Interdisciplinary Journal of Nonlinear Science,* Vol'11, Issue 4, December 200, pp. 809-850

[5] D. J.Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature,* 393, 1998, pp. 440-442.

[6] A-L. Barabasi, R. Albert, Emergence of scaling in random networks, *Science*, 286, 1999, pp. 509-512

[7] S. N. Dorogovtsev, J. F. F. Mendes, Natural Scale of Scale-free Networks. *Phys. Rev. E* **63**, 2001, p. 62101

[8] Cohen, R., K. Erez, D. ben-Avraham and S. Havlin (2000). Resilience of the Internet to Random Breakdown. *Phys. Rev. Lett.* **85**, pp. 4626–4628

[9] C.A. Dinarello, Proinflammatory and antinflammatory cytokines as mediators in the pathogenesis of the septic shock, *Chest*, 1997, 112: 321S-329S

[10] A. Seely and N. Christou, "CCM Multiple organ dysfunction syndrome: Exploring the paradigm of complex nonlinear systems", *Critical Care Medicine*, 28(7):2193-2200, July 2000.

[11] H. Fan and J. Cook J, "Molecular mechanisms of endotoxin tolerance", Journal of Endotoxin Research (2004), vol. 10, No.2, pp. 71-84

[12] P. Flondor and C. Vasilescu, "The dynamic of inflammatory processes", in R. Dobrescu (editor) – *Proceedings of the second Int. Symp. on Interdisciplinary Applications of Fractal Analysis*, POLITEHNICA Press, Bucharest, pp. 33-38

[13] R. Dobrescu, M. Dobrescu, D. Popescu, St. Mocanu - Differential Equations Systems versus Scale Free Networks in Sepsis Modeling, WSEAS Transactions on Information Science and Applications, Issue 12, Vol. 2, dec.2005, pp. 1746-1751