From Parts-to-Whole: A Holistic Theory of Optimal Design and Adaptation in Complex Networks

Venkat Venkatasubramanian
Laboratory for Intelligent Process Systems
School of Chemical Engineering
Purdue University
Tel:(001) 765 494 0734
venkat@ecn.purdue.edu

Leaelaf Hailemariam
Laboratory for Intelligent Process Systems
School of Chemical Engineering
Purdue University
Tel:(001) 765 496 7362
lhailema@ecn.purdue.edu

Arun Giridhar
Laboratory for Intelligent Process Systems
School of Chemical Engineering
Purdue University
Tel:(001) 765 496 7362
agiridha@ecn.purdue.edu

ABSTRACT

Recent research in network science has shown that complex networks in diverse application areas seem to exhibit several universal features. In this paper, we describe a holistic theory of optimal design and adaptation of complex networks based on the universal organizing principle of network structure optimization to meet overall survival or performance objectives. We demonstrate how a network self-organizes by evolutionary adaptation to maximize its survival or performance fitness, which in turns depends, under given environmental survival pressure, on the network's efficiency, robustness and cost. The theory provides a general conceptual framework for integrating network topology with survival objectives, selection pressures, evolutionary adaptation and performance optimization in a single coherent formalism. Further, we investigate another related and important question, namely, how "parts-level" network properties such as the vertex degree determine "system-level" properties such as survival or performance. We also examine the implications of uncertainty in the environment on optimal network design. Using the concepts of interaction 'values' and network 'satisfaction' as generic performance measures, we show that the underlying organizing principle is to meet an overall performance target with reliable performance under maximum environmental uncertainty. This leads to the emergence of the power law degree distribution, often displayed in complex networks, as a consequence of the Maximum Entropy Principle. This notion also leads to other distributions under various limiting conditions.

Keywords

Complex networks, power laws, maximum entropy, optimization, adaptation, evolution, systems biology, networks science

1. Introduction

Complex networks are found in a wide variety of domains such as biology, ecology, sociology, engineering and so on. One of the outstanding problems in network science is explaining and predicting the emergence of self-organized network structures with very interesting properties [1,2]. Such networks, found in diverse applications such as supply chains, computer and communication networks, metabolic networks, food webs etc., often exhibit similar topological features. Recently, there have been attempts to propose mechanisms, such as, preferential attachment [3], or different optimization formulations [4-6] for the emergence of the scale-free topologies for such networks. These results provide valuable insights into the structure of scale-free networks; however, the questions of why and how these different network configurations emerge, their significance, and common underlying design principles, if any, remain to be clarified further. Another central question is how "local" network properties such as individual vertex degrees, which are node-level or part-level properties determine the "global" or system-level properties such as network survival or performance. In other words, *how do we go from parts to whole* in a complex system? We summarize the progress we have made in addressing these questions in this paper.

We describe how a complex network optimizes its structure via self organization by evolutionary adaptation to maximize its overall survival or performance fitness [7]. The theory provides a general conceptual framework for integrating network topology with survival objectives, selection pressures, evolutionary adaptation and performance optimization in a single coherent formalism. Further, we also investigate the implications of environmental uncertainty on optimal network organization. The design requirement of reliable performance under maximum uncertainty, as a consequence of the maximum entropy principle, leads to the emergence of power law degree distributions, often displayed in complex networks, as well exponential or Poisson distributions, thus explaining all three regimes observed in reality as different manifestations of the same underlying phenomenon within a novel unified theoretical design framework [8].

2. Self-Organization by Evolutionary Adaptation

Venkatasubramanian *et al.* [7] proposed a framework for self-organization of a network by evolutionary adaptation, modelled after Darwin, in which the network's objective is to maximize its chances of overall survival by adapting its configuration according to the environmental pressure. The basic premise is that existing biological networks exhibit certain characteristic configurations and properties because the same helped them survive the test of time and natural selection. A network typically serves to transport material, energy, and/or information; thus the idea of survival is included in the more general term network 'performance' in terms of meeting its design objectives. Therefore, the novel hypothesis is that although human-engineered networks have not necessarily 'emerged' by evolutionary adaptation, the underlying design principles that led to their creation could be very similar to those that caused biological networks to evolve to their present forms.

The overall performance is modelled as a function of a *short-term* and a *long-term* component, subject to *cost* or *resource* constraints in a given environment. These two objectives, short and long-term performance, are often conflicting, requiring a trade-off in the design. In general, they depend on two critical measures of the system: *efficiency* (η_E) and *robustness* (η_R) respectively. Efficiency is a measure of the effectiveness of the network configuration to accomplish its functions. Robustness accounts for the extent to which the network is able to carry out its functions despite suffering some damage, such as the removal of some nodes and/or edges.

Depending on the functional goal and environment, the network structure is *optimized* for efficiency and "worst case" robustness subject to cost constraints. For a given network acting under a specified environment, the optimization problem is to maximize the network's overall performance fitness, G:

$$\max_{configuration} \quad G = \alpha\, \eta_E + (1-\alpha)\, \eta_R \tag{1}$$

The parameter α models the environmental or selection pressure on the network. When α=1, network performance or survival depends entirely on its efficiency (i.e. short-term) with no regard for robustness. At the other extreme, when α=0, only worst case robustness i.e. long-term performance is paramount with no regard for efficiency. For intermediate values of α, the environment demands both efficiency and robustness to varying degrees. Thus, by varying α from 0 to 1, different selection pressures can be imposed on network performance.

Network Design under Environmental Uncertainty

The optimization framework discussed above considers a given, imposed environment (specified by α) for network organization. However, a network's design or organization should ideally enable it to perform or survive under a wide variety of operating conditions or environments. The implications of such environmental uncertainty on network organization have been addressed by Venkatasubramanian *et al.* [8]. The local or node-level properties of a network, such as individual vertex degrees, are connected to its global or overall properties via "values" of interactions defined between pairs of nodes. The notion of value here is generic, which in practice, may be a design variable like productivity, money, effort, efficiency, time, CPU, bandwidth, emotional support etc., or some combination of such variables. In addition, there may be various domain-specific constraints on these variables. In general, each such relationship may '*cost*' the member a certain amount, similar to the expenses associated with creating and maintaining computer lines, roadways, transaction lines, friendships, etc. in real networks. From a holistic perspective, the design principle for macroscopic organization would be to meet a certain level of overall value or *global satisfaction* for the network as a whole. However, a reliable design would not presume that only certain conditions will prevail in the future and hence limit the design to perform well only under those conditions. That is, one would not want to *bias* the design for a specific operating environment, particularly if the nature of future environments is unknown, uncertain or unpredictable. Hence, the network design should reflect this *inherent uncertainty* about future operating environments and *minimize the bias* or any unwarranted assumptions about them. In other words, a reliable design should accommodate as much uncertainty about the future operating environments as allowed by the constraints. Then, this requirement of meeting a desired level of global satisfaction with reliable performance under maximum uncertainty leads, via the Maximum Entropy Principle [9-11], to a power law as the optimal degree distribution. When the number of edges is very small, the Maximum Entropy Principle leads to an exponential distribution. The same formalism also produces a Poisson distribution, when the number of edges is very large, due to the high level of redundancy in the network. The theory predicts all three regimes as different manifestations of the same underlying phenomenon within a unified theoretical design framework. The above mechanism may help explain emergence of power law behavior in metabolic networks in their degree distributions (Jeong et al., 2000; Wagner and Fell, 2001; Ma and Zeng, 2003) and metabolic flux distributions (Almaas et al., 2004). Our framework lays the ground work for a novel approach to model, analyze, and design complex biological networks and human-engineered networks.

REFERENCES

[1] Albert, R., Barabási, A.-L. (2002). *Rev. Mod. Phys.*, 74, 47.

[2] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A.-L. (2000). *Nature* 407, 651.

[3] Barabási, A.-L., Albert, R. (1999). *Science* 286, 509.

[4] Carlson, J. M., Doyle, J. (2000). *Phys. Rev. Lett.,* 84, 2529.

[5] Valverde, S., Cancho, R. F., Solé, R. V. (2002). *Europhys. Lett.*, 60, 512.

[6] Newman, M. E. J., Girvan, M., Farmer, J. D. (2002). *Phys. Rev. Lett.,* 89, 028301:1-4.

[7] Venkatasubramanian, V., Katare, S. R., Patkar, P. R., Mu, F-P. (2004). *Comput. Chem. Eng.*, 28, 1789.

[8] Venkatasubramanian, V.; Politis, D. N.; Patkar, P. R. (2006), AIChE Journal, 52(3).

[9] Jaynes, E. T. (1957a). *Phys. Rev.*, 106, 620.

[10] Jaynes, E. T. (1957b). *Phys. Rev.*, 108, 171.

[11] Cover, T. M., Thomas, J. A. (1991). Elements of Information Theory. *John Wiley & Sons*, New York.

[12] Wagner, A., Fell, D.A., Proc. R. Soc. Lond. B. (2001), 268, 1803.

[13] Ma, H-W, Zeng, A-P.(2003). *Bioinformatics.,* 19(11), 1423.

[14] Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N. , Barabási, A.-L. (2004). *Nature,* 427, 839.