Emergent of networks based on relationship among elements

Yasuhiro Suzuki and Hiroshi Tanaka
Bio-Informatics,
Medical Research Institute,
Tokyo Medical and Dental University
Yushima 1-5-45, Bunkyo, Tokyo 113 JAPAN

Abstract

It is known today there are also many networks existing in living systems like signal transduction, transcriptional regulation, immergence response and like so on. These networks, like all the other networks, are based on relationships between elements. In this contribution, we propose a network model of living systems. It possesses information of connections among nodes, which stand for critical element in the network. We formulate it as a simple model of probability theory and find interesting phenomena in it.

1   Introduction

With the development in biotechnology and bioscience these years, we are now able to investigate living things in a molecular level. The basic genetic information existing in living systems can be roughly depicted by central dogma theory and many phenotypes can be explained by genes regulation. Genes regulation is so complicated that search the "regulation map" bit by bit is the only and efficient way at the moment. Our knowledge of regulation networks today is an accumulation of all the experimental results over the past and is fulfilled with a increasing speed. Today, the advents like microarray, large-scale sequencing machines enable us to acquire quantitive data in a very short period of time, which make it a critical point for us to analyse the data. As traditional biological analysis is only powerful in analyzing a single gene most of the case, the system biological study turns out to be more and more important in life science.In many fields, we can find that theory establishment and development in advance is usually instructive and beneficial.Therefore to challenge the understanding of network, especially regulatory networks of genes shall be promising. To clarify gene network,it is usually necessary to model a system by synthesizing the empirical data at first, in which System biology approach is necessary. There is few models existing, for example, Kauffuman's boolean network model,has been a instructive model in analyzing gene network though it is still insufficient.

Small worlds is a new system used to deplore the property of Complex systems. It originally comes from social experiments, for example, two people in the world chosen randomly were found to be connected with each another by a short cut chain of intermediate acquaintances, and this phenomenon is called "small world effect". These days, there are many modelings and theoretical researches of this phenomena in investigating all kinds of networks.

The structure of Small worlds was founded in social research at first. Stanley Milgram preformed a simple experiment in 1960s. He took a number of letters addressed to a stockbroker acquaintance of his in Boston, Massachusetts, and distributed them to a random selection of people in Nebraska. His instruction were that the letters were to be sent to their addressee by passing them from person to person, and that, in addition, they could be passed only to someone whom the passer new on a first-name basis. Since it was not likely that the initial recipients of the letters were on a first-name basis with a Boston stockbroker, their best strategy was to pass their letter to someone whom they felt was nearer to the stockbroker in some social sense: prehaps someone they knew in the financial industry, or a friend in Massachusetts. A reasonable number of Milgram's letters did eventually reach their destination, and Milgram found that it had only taken an average of six steps for a litter to get from Nebraska to Boston. He concluded, with a somewhat cavalier disregard for experimental niceties, that six is therefore the average number of intermediate acquaintances separating two randomly chosen humans. He labeled third situation "six degrees of separation", a phrase which has since passed into popular folklore (Guare 1990[5]) [7].

This experiment is not precisely controlled,

however, we can find these kind of structures in many systems such as WWW documents,movie star's friends network, or mathematicians of Erdos's co-workers and likewise. In a network, there are hubs and each of them holds many neighbors, which make all the nodes be linked to each other in the end. This is called a "**small world**".

2 Model

A simplest model of network is random graphs. It can explain small-world effect. Suppose there is some number N nodes and on average each node has z out degree. This means there are $\frac{1}{2}Nz$ connections between edges. This kind of graph is called Random graph. Random graphs have been studied in mathematics community [3] however Kauffuman use it to model gene networks. In general, a random graph dose not show clustering. In a graph the probability that two of nodes will be neighborhoods is no grather than the probability that two randomly chosen nodes will be.

We are interested in the process of a network's growth. However, this is seldom addressed in models presented by other groups of small worlds and random graphs. Real networks in living systems are so complex that it is difficult to find the underlying principle from numerous phenomena, although in living systems there are many networks and elements which are connected based on some rules instead of random connection. Thus we model a network based on the relationship among elements.

Technically the model is composed of directed graph. Each node of the graph remembers the history of visitors (the name of nodes that came to the node). The history gives the probability of transition from the node, and the transition probability from node i to j is defined as

$$\frac{\Sigma n_m}{\Sigma n_i}.$$

where Σn_m denotes the number of in degree from node i to m and Σn_i denotes the total number of in degree of node i. And when the edge from node i came to node m then the probability is changed as

$$\frac{\Sigma n_m + 1}{\Sigma n_i + 1},$$

thus transition probability is changed dynamically. This model is closely related to "Polya's urn[4]" which is a famous model of probability theory. Polya's urn is defined as follows;

Polya's urn

Supposing there is a urn with b black balls and r red balls inside. When we select a ball randomly and put it back to the urn, s and c balls that are painted by the same color of the selected ball and d balls that are painted by the the other color. The procedure is repeated.

Our model allows the condition when there are more than two kind of balls as c = 1 and d = 0. When there are two kinds of colors, it is well investigated [4]. Thus we defined the above model as the network of generalized Polya's urn as follows;

Network of generalized Polya's urns

Each node has a "urn" in which balls labeled node number except for own number. First a ball is randomly selected from the urn, and then two balls are returned to the urn, the added ball is labeled the same node number with the selected ball.

3 Experimental Results

Since it is not easy to analyze the problem theoretically, we performed simulations in order to investigate the process of emerging network in this model.

First, we assume 100 nodes and preformed a simulation until 150,000 edges were linked. The figure 4 displays the network emerged. In this figure 100 nodes on the edge of a cycle and each edge denotes there are more than five times transitions between the nodes.

Shift lock effect in a network   However, the number of nodes increases up to about 130 the topology of network changes and it is not a connected graph under the same conditions in the previous case. The figure ?? illustrates it. In this figure 200 nodes on a cycle and each edge illustrates more than five times transitions between the nodes.

In this case, once some nodes connected strongly then the shift of connection is locked and by positive feedback among these connections, they are strongly connected other than them. Thus, the connections can not spread over to another nodes.

By introducing randomness in this network shift lock effect is disappeared. In the above model in every time a ball is selected randomly and put two balls back to the urn but if only one ball is backed then the connection dose not

strengthen in such a case. Hence, by using this method, we introduce randomness in the model. And the ratio of inputting two balls to the urn is 1/3. In this case, the network becomes a connected graph.

3.1 Hierarchical structure of edges

Although it is difficult to see the hierarchical structure of edges in the figures 1, 2 there emerges the structure in fact. Hierarchical structure of edges means that at the top of the structure there are highly used edges and based on it there are secondary used edges and the structure is formed like cascade.

4 Discussion

Biotechnologies like DNA chip is widely used today to provides genome wide information fast and informative to discover gene networks. Combination with genomic database and use of System biology, now we are closer to find the correlation and regulation between genes and genes. It even can be imagined that in the near future we can obtain various and lots of data from living systems much easier and quicker. Thus the observation and analysis of those data will be much critical as ever. As traditional researches accumulate our knowledge of living systems from a bottom-to-top way, we are easier to pinpoint some specific gene or pathway, while on the hand, to understand the whole living system from a top-to-down way still remains complicated as before. Our method is not sufficient, but we are keep refining it with the belief that basic regulation rules exist in living systems. Method proposed here shall be helpful to answer the questions like what genes are critical in the process of senescence or cancer; What genes are important in controlling a human being's weight referring to the balance between synthesis and catabolism, or what is the regulation rules common in species. The theory can even give hint to solve the problem in finding the simplest and basic element for a living thing, which is a fantastic goal for biologists and computer scientists for a long time.

Acknowledgment

References

[1] Bollobas, B., Random Graphs, Academic Press, 1985

[2] Epstain, C. and R.Butow, Microarray technology -enhanced versatillity, persistent challaenge, Curr. Opin. Biotech, 11, 36-41, 2000

[3] Erdos, P., and A. Reny, On random graphs Publicationes Mathematicae, 6, 290-297, 1959.

[4] Feller W., An Introduction to Probability Theory and Its Applications, I, 1957.

[5] Guare, J., Six Degrees of Separation, A Play, Vintage, 1990.

[6] Kauffman, S.A., The Orgin of Order, Oxford Univ. press, 1993.

[7] Newman, M.E.J., Small Worlds, Santa Fe Institute Working Paper, 1999.

[8] Smogyi, R., and C. A. Sniegoski, Modeling the Complexity of Genetic Networks: Understanding Multigenetic and Pliiotropic Regulation, Complexity 1(6), 45-63, 1996.

[9] Smogyl, R, et.al., Cluster Analysis and data visualization of largescale gene expression data, Pasific Symposium on Biocomputing,3,42-53,1999.

[10] http://www.cs.virginia.edu/oracle/

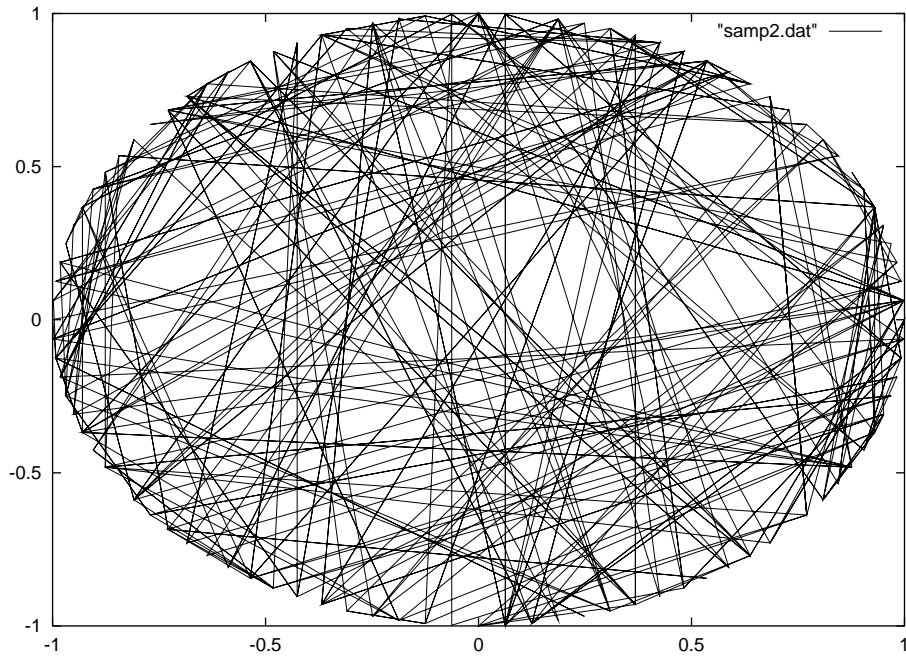[11] Watts,D.J, Small Worlds, Princeton Univ. Press, 1999.
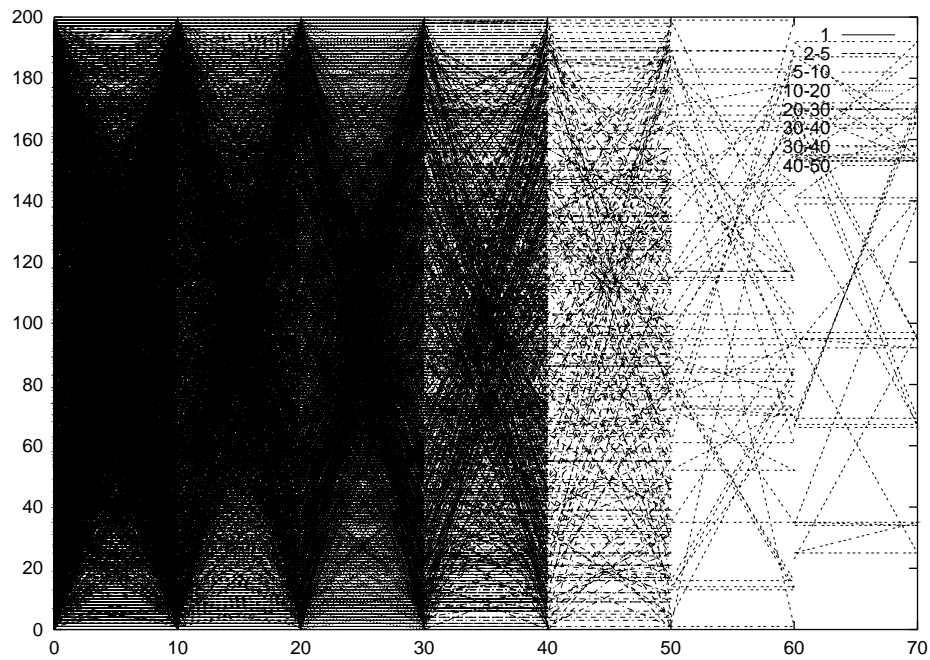
Figure 1: Polya's urns Network

Figure 2: Clustering of nodes of Polya's urn network