Generalized Maximum Entropy Principle, Superstatistics and Problem of Networks Classification

Bahruz Gadjiev

International University for Nature, Society and Man, 19 Universitetskaya street, Dubna, 141980, Russia

Abstract. In the presented paper the results of superstatistics methods are given to study growing networks with exponential topology. Topologies of strongly inhomogeneous growing networks are determined. In the framework of maximum entropy method the probability distribution of real networks is derived and their classification is discussed.

Keywords: Generalized maximum entropy principle, Superstatistics, Networks classification
PACS: 05.20.-y, 05.70.Ln, 89.70.Cf, 05.90.+m, 89.75.-k

INTRODUCTION

In the past few years it was established that in the basis of the complex architecture of real-world networks there are universal principles. Analysis of real world complex networks shows that many of them have a scale-invariant topology. The network approach has proved itself a powerful tool for analyzing the structural complexity of systems [1,2].

Classification of real-world networks shows that they can be divided into random network Erdos-Renyi and small-world network of Watts-Strogatz and growing scale-invariant networks (uncorrelated and correlated) [3]. The data analysis shows that for every complex system is a network with nontrivial topology. From the data analysis we determine the order for the set of millions nodes and links, forming a complex network. At present simple algorithms are developed to generate random, correlated and uncorrelated scale-invariant networks. It is possible to study the stability of these networks and random failures and targeted attacks, as well as the spreading processes on networks. Despite the fact that in the real world growing networks combined mechanisms of growth act, the basic growth principles of networks are well known.

This paper is devoted to the study of nonequilibrium networks with spatio-temporal fluctuations of intensive parameters on a large scale.

SUPERSTATISTICAL APPROACH TO STUDYING COMPLEX SYSTEMS

Complex systems in nature exhibit a rich structure of dynamics, described by a mixture of different stochastic processes on various time scales. Such dynamical processes with time scale separation are described in the framework of the superstatistical approach [4, 5, 6, 7]. To explain the idea of the superstatistical approach we start with a simple example.

Consider processes subject to both additive and multiplicative noises described by the dimensionless stochastic differential equation of the form

$$\dot{u} = f(u) + g(u)\xi(t) + \eta(t) \tag{1}$$

where $u(t)$ is a stochastic variable, f and g are arbitrary functions and $\xi(t)$ and $\eta(t)$ are uncorrelated and Gaussian-distributed zero-mean white noises, hence satisfying

$$\langle \xi(t)\xi(t') \rangle = 2M\delta(t - t') \text{ and } \langle \eta(t)\eta(t') \rangle = 2A\delta(t - t') \tag{2}$$

where $M > 0$ and $A > 0$ are the noise amplitude and stand for multiplicative and additive, respectively [8].

The Fokker-Plank equation for probability density $P(u,t)$, associated to equation (1) , can be obtained from the Kramers – Moyal expansion

$$\frac{\partial P(u,t)}{\partial t} = \sum_{n \ge 1} \left(-\frac{\partial}{\partial u} \right)^n \left[D^{(n)} P(u,t) \right] \tag{3}$$

where the coefficients are given by

$$D^{(n)}(x,t) = \frac{1}{n!} \lim_{\tau \to 0} \frac{[u(t+\tau) - x]^n}{\tau} \bigg|_{u(t)=x} \tag{4}$$

Using the Stratonovich definition of stochastic integral, one gets

$$D^{(1)}(u,t) = f(u) + Mg(u)g'(u) \tag{5}$$

$$D^{(2)}(u,t) = A + M[g(u)]^2 \tag{6}$$

while $D^{(n)}(u,t) = 0$ for $n \ge 3$.

Then, the Fokker-Planks equation has the form

$$\frac{\partial P(u,t)}{\partial t} = -\frac{\partial j(u,t)}{\partial u} \tag{7}$$

where

$$j(u,t) = D^{(1)}(u,t)P(u,t) - \frac{\partial}{\partial u}\left[D^{(2)}(u,t)P(u,t)\right] \tag{8}$$

is the current.

Now suppose that $g(u) = \sigma$, where σ is constant and $f(u) = -\gamma u$. For further study we shall assume that $A = 0$. In this case $D^{(1)}(u,t) = -\gamma u$ and $D^{(2)}(u,t) = M\sigma^2$. Then, using the $\frac{\partial P(u,t)}{\partial t} = 0$ the stationary Fokker-Plank equation can be written as

$$-\frac{\partial j(u,t)}{\partial u} = 0 \tag{9}$$

We will restrict to the stationary solutions for no flux boundary conditions (i.e., such as $j_{st}(-\infty) = j_{st}(\infty) = 0$). Then using the $P(u,\pm\infty) = 0$, we get $j_{st}(u) = 0$. In this case, one obtains

$$\frac{\partial}{\partial u}\left[M\sigma^2 P_{st}(u)\right] = -\gamma u P_{st}(u) \tag{10}$$

The solution of this equation can be written as follows

$$P_{st}(u) = P_0 e^{-\frac{1}{2}\beta u^2} \tag{11}$$

where $\beta = -\frac{\gamma}{2M\sigma^2}$.

In superstatistical approach “local equilibrium” is meant in a generalized sense for suitable observables of the system dynamics under consideration. In the long term, the stationary distribution of a superstatistical inhomogeneous system arises as superposition of a local factor $e^{-\frac{1}{2}\beta u^2}$ with various values of β weighted with a global probability density $f(\beta)$ to observe some value β in a randomly chosen cell.

While on the time scale T the local stationary distribution in each cell is Gaussian $p(u|\beta) = \left(\frac{\beta}{2\pi}\right)^{\frac{1}{2}} e^{-\beta u^2}$ the distribution describing the long-time behavior of the entire systems for $t >> T$

$$p(u) = \int_0^\infty \left(\frac{\beta}{2\pi}\right)^{\frac{1}{2}} e^{-\beta u^2} f(\beta) d\beta, \tag{12}$$

exhibits non trivial behavior.

For example, if $f(\beta)$ is χ^2-distribution of degree n equation (12) generates Tsallis statistics, with entropic index q given by $q = 1 + \frac{2}{n+1}$:

$$p(u) = \frac{1}{\left(1 + (q-1)\beta_0 u^2\right)^{1/(q-1)}}. \tag{13}$$

Discussing physically relevant superstatistical universality classes it should be noted that there are three physically relevant universality classes: (1) χ^2-statistics (=Tsallis statistics), (2) inverse χ^2- superstatistics and (3) lognormal superstatistics [6]. These arise as universal limit statistics for many different complex systems. In principle of course, other classes of universality are possible.

SUPERSTATISTICAL APPROACH TO STUDYING COMPLEX NETWORKS

If a given set of N nodes is connected by a fixed number of links in a completely random manner, the result is a random network Erdos – Renyi, whose degree distribution is Poissonian i.e., the probability that a randomly chosen node has degree k is given by $p(k) = \frac{\lambda^k e^{-\lambda}}{k!}$, where $\lambda = \bar{k}$ is the average degree of all nodes in the network. In [9] it was assumed that λ in random networks is fluctuating according to distribution $\Pi(\lambda)$. In this sense a network with any degree distribution can be presented as a 'superposition' if random networks with the given degree distribution are $p(k) = \int_0^\infty d\lambda \Pi(\lambda) \frac{\lambda^k e^{-\lambda}}{k!}$. In [9] it was shown that a power-law functional form of $\Pi(\lambda)$ leads to degree distribution of Zipf-Mandelbrot form, $p(k) \sim (k_0 + k)^{-\gamma}$, which is equivalent to a q-exponential with an argument of k/κ, and given the substitutions, $\kappa = (1-q)k_0$ and $q = 1 + 1/(1 + (1/\gamma))$.

Consider the growing networks with exponential degree distribution. Initially, the system has m_0 unconnected nodes. Every time the system added a node with m ($m \le m_0$) edges, in accordance with a uniform distribution the process of linking with existing network nodes takes place. In this case, the degree distribution of the network represented by the expression

$$P(k) = \frac{1}{m} e^{-\frac{k}{m}} \tag{14}$$

where $m = \int_0^\infty kP(k)dk$ — average degree of node.

The essential difference between the networks of Erdos-Renyi and exponential networks is as follows. The principle of connection edges with nodes in the networks of Erdos-Renyi and exponential networks is the same. However, in networks of Erdos-Renyi all nodes have the same age, then, as in exponential networks, age sites are different. This, in particular, leads to the fact that the exponential networks are assortative, while the network Erdos - Renyi are uncorrelated. It arises because of the principles distinction of networks formation.

Let us discuss the applicability of the principles of superstatistics to analyze networks with exponential topology.

Let the system have local worlds where there is growth of networks with an exponential distribution $p(k|m)$, but to the average degree m of local worlds are distributed according to some distribution $f(m)$. Then, using Bayes' theorem the distribution of the entire system can be obtained in the form [10]

$$P(k) = \int_0^\infty p(k|m) f(m) dm \tag{15}$$

In fact, we are dealing with the integral equation, namely, the known $P(k)$ and $f(m)$ and need to find a solution $p(k|m)$. In the case of networks with exponential degree distribution we are dealing with the integral equation in the form

$$P(k) = \int_0^\infty p(km) f(m) dm \tag{16}$$

In our case, we multiply both sides of this equation by k^s and integrate from zero to infinity. Using the definition of the Mellin transformation $\tilde{g}(s) = \int_0^\infty g(k) k^{s-1} dk$, we get

$$\int_0^\infty f(m) dm \int_0^\infty p(km) k^{s-2} dk = \int_0^\infty P(k) k^{s-1} dk \tag{17}$$

In the inner integral we introduce variable $\varsigma = km$ and after integration we lead to

$$\int_0^\infty f(m) m^{1-s} dm \int_0^\infty p(\varsigma) \varsigma^{s-2} d\varsigma = \int_0^\infty P(k) k^{s-1} dk \tag{18}$$

Using the definition of the Mellin transformation we obtain

$$\tilde{f}(-s)\tilde{p}(s-1) = \tilde{P}(s) \tag{19}$$

By the Mellin inversion formula

$$f(k) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{f}(s) k^{-s} ds \tag{20}$$

we determine the solution of the integral equation.

For a network with the exponential topology, we have the degree distribution $P(km) = \frac{1}{k} kme^{-km}$. Assume that the degree distribution of the entire system is a Tsallis distribution

$$P(k) = \frac{1}{\left[1 + (q-1)k_0 k\right]^{\frac{1}{q-1}}}, \tag{21}$$

where $\Gamma(x)$ is gamma function, we find that the probability distribution of average degrees in local worlds has the form of gamma distribution

$$f(\langle k\rangle)=\frac{1}{\Gamma\left(\frac{1}{q-1}\right)}\left\{\frac{1}{(q-1)k_0}\right\}^{\frac{1}{q-1}}\langle k\rangle^{\frac{1}{q-1}-1}e^{-\frac{\langle k\rangle}{(q-1)k_0}} \qquad (22)$$

Let a degree distribution of entire system be defined by function of the parabolic cylinder $D_{-2}(z)=\frac{e^{-\frac{z^2}{4}}}{\Gamma(2)}\int_0^\infty e^{-zx-\frac{x^2}{2}}xdx$, which asymptotically behaves as $\sim z^{-2}e^{-\frac{z^2}{4}}$. As we consider a network with an exponential degree distribution $P(km)=\frac{1}{k}kme^{-km}$, we find that

$$f(m)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}m^2} \qquad (23)$$

Let's note that normal distribution can be received as a limiting case binomial distribution $B_k(n,p)=\binom{n}{k}p^k(1-p)^{n-k}$ at $n\to\infty$ and $p=\frac{1}{2}$. Assuming $\varphi_n(u)=\sigma B_k(n,p)$, where σ is mean square deviation and taking a limit $n\to\infty$ we get normal distribution $\lim_{n\to\infty}\varphi_n(u)=\varphi(u)=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$. Here $\varphi(u)$ is a probability density.

Thus, in case of an exponential network when fluctuations of average degrees in the local worlds is the normal distribution then degree distribution of entire system described by distribution $\sim k^{-2}\exp(-(k^2/4))$.

Let temporary evolution of a local world be carried out in correspondence to the following algorithm: at the initial moment in a network there are m_0 isolated nodes and in each discrete time in a network the new vertex with m edges are added. Every edge of new vertex is connected with node presented at a network randomly.

Let at each moment of time from each local world a new local world is born where the growth of a network with exponential topology takes place at subsequent moments of time regardless of other local worlds.

As the mean value of the vertex degree is determined in each local world as $\langle k\rangle = m$, the number of edges that arrive with the given vertex in the local worlds should be taken from the distribution $\rho(\langle k\rangle = m)$. Therefore, the vertices that come into various local worlds have different number of edges; the distribution of these edges' number is described by the distribution $\rho(m)$. The cohesiveness of such a system may be achieved with homogeneous connection of local worlds at the expense of edges that arrive into the network. Then the Bayes theorem leads us to the distribution for the entire system in the following form:

$$\int_0^\infty d\beta\rho(\beta)\beta e^{-\beta k} = P(k) \tag{24}$$

It is obvious that different distributions $\Pi(x)$ lead to different probability distributions $P(k)$.

However it is possible to show also, that exponential network topology with scale invariant $\Pi(\beta)$ fluctuations of β can not exit.

PRINCIPLE MAXIMUM ENTROPY TO STUDYING COMPLEX NETWORKS

The study of results of data on networks topology shows that basically topology of the studied networks are described by distribution

$$p(u) = \frac{1}{\left(1 + (q-1)\beta_0 (u-a)^\nu\right)^{1/(q-1)}} \tag{25}$$

where $\nu \in [1,2]$. Let's discuss the derivation of such type of distributions in the framework of a maximum entropy principle. Note that Shannon entropy has been extended in several ways. One particular generalization is Havrda-Chavrat α -entropy

$$h_\alpha(x) = \frac{1 - \int_{-\infty}^{\infty} f^\alpha(x)dx}{\alpha - 1} \tag{26}$$

Shannon entropy is the special case of (26) as $\alpha \to 1$.

Using the maximum entropy principle we shall show that an entropy (26) with restrictions

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{27}$$

and

$$\int_{-\infty}^{\infty} |x-a|^b f(x)dx = g_0 \tag{28}$$

leads to the distribution function

$$f(x) = \frac{\left[1 - (1-\alpha)\frac{\beta}{\alpha\chi} g\left(|x-a|^b\right)\right]^{\frac{1}{\alpha-1}}}{Z} \tag{29}$$

Really, according to a method of Lagrange we form expressions

$$\Lambda = h_\alpha(x) - \alpha \int_{-\infty}^{\infty} f(x)dx - \beta \int_{-\infty}^{\infty} |x-a|^b f(x)dx \tag{30}$$

and then from condition $\frac{\partial \Lambda}{\partial f(x)} = 0$ we find

$$\frac{\alpha}{1-\alpha} f^{\alpha-1}(x) - \alpha_0 - \beta |x-\beta|^b = 0 \quad (31)$$

Here α_0 and β are Lagrange factors. Equation (31) may be rewritten as

$$\frac{\alpha}{1-\alpha} \chi - \alpha_0 - \beta g_0 = 0 \quad (32)$$

from which we define α_0. Substitution of expression α_0 into equation (31) allows us to define

$$f(x) = \chi^{\frac{1}{\alpha-1}} \left[1 - (1-\alpha) \frac{\beta}{\alpha\chi} g\left(|x-a|^b\right) \right]^{\frac{1}{\alpha-1}} \quad (33)$$

Here $g\left(|x-a|^\beta\right) = g_0 + |x-a|^\beta$.

After integration of equation (33) we find $\chi^{\frac{1}{1-\alpha}} = \int_{-\infty}^{\infty} \left[1 - (1-\alpha) \frac{\beta}{\alpha\chi} g\left(|x-a|^b\right) \right]^{\frac{1}{\alpha-1}} dx$

and after introducing notation $Z = \chi^{\frac{1}{1-\alpha}}$, we get

$$f(x) = Z^{-1} \left[1 - (1-\alpha) \frac{\beta}{\alpha\chi} g\left(|x-a|^b\right) \right]^{\frac{1}{\alpha-1}}. \quad (34)$$

where $Z = \int_{-\infty}^{\infty} \left[1 - (1-\alpha) \frac{\beta}{\alpha\chi} g\left(|x-a|^b\right) \right]^{\frac{1}{\alpha-1}} dx$.

From equation (34) it follows that if $b=1$ at $\alpha \to 1$ we get an exponential distribution whereas $b=2$ we get a normal distribution.

Thus, variable α allows us to make classification of systems on the level of a degree of complexity of systems whereas b allows us to define a distribution function of the initial simple systems.

REFERENCES

1. G. Bianconi and A.-L. Barabasi, *Eurohysics letters* **54** (4), 436 – 442 (2001).
2. S. N. Dorogovtsev and J. F. F. Mendes, *Phys. Rev. E* **63**, 056125 1-19 (2001).
3. R. Albert and A.-L. Barabasi, *Reviews of Modern Physics* **74**, 47-97 (2002).
4. Superstatistics: Recent developments and applications,, in 'Complexity, Metastability and Nonextensivity', eds. C. Beck, G. Benedek, A. Rapisarda, and C. Tsallis, World Scientific (2005)
5. C. Beck, *Contemp. Phys.* **50**, 495-510 (2009).
6. C. Beck, *Braz. J. Phys.* **39**, 357 (2009).
7. C. Beck and E. Van der Straeten, *Phys. Rev. E* **78**, 051101 (2008).
8. C. Antenedo and C. Tsallis, *Journal of mathematical Physics* **44**, N 11, p. 5149 (2003).
9. S. Thurner, *Europhysics News* **36,** No 6, p. 218 (2005)
10. C. P. Robert, N. Chopin, J. R. Harold, *Journal: Statistical Science* **24**, No. 2, pp. 191-194 (2009).