American Journal of Mathematical and Computational Sciences
2021; 6(1): 1-5
http://www.aascit.org/journal/ajmcs

American Association for Science and Technology

The Comparison of Entropy and Parameter Estimation in Complex Networks

Zhao Chunxue

School of Mathematics and Statistics, Anyang Normal University, Anyang, China

Email address
zhaochunxue66@163.com

Citation
Zhao Chunxue. The Comparison of Entropy and Parameter Estimation in Complex Networks. *American Journal of Mathematical and Computational Sciences*. Vol. 6, No. 1, 2020, pp. 1-5.

Received: November 25, 2020; **Accepted**: January 8, 2021; **Published**: January 22, 2021

Abstract: Entropy is used to describe the uniform distribution of any kind of energy in space. The more uniform the energy distribution, the higher the entropy. When the energy of a system is uniformly distributed, the entropy of the system reaches its maximum. Most complex systems can be described to network model, which includes a large number of nodes and complex connection relationships, Large number of networks show seemingly unrelated, but exist many striking similarities. According to different degree distribution, the network is divided into four kinds: regular network, random network, small-world network, scale-free network. Entropy is also a very important indicator which describes the heterogeneity of the networks. The scale-free network shows a non-homogeneous nature and a kind of sequence the complex network emerges. There have been some researches on using entropy to study complex networks. In the paper, we quantify the scale-free properties of the complex network by using the entropy theory and maximum likelihood estimation (MLE). We first review two kinds of entropy and prove their consistency; then we investigate the relationship of the parameter estimation among MLE, the moment estimator and the entropy of fitness scale-free network; finally, we gain the entropy of random network, which provides theoretical support for the practical application of entropy.

Keywords: Complex Network, Entropy, Maximum Likelihood Estimation, Moment Estimator

1. Introduction

Almost all complex systems can be described to network model, which includes a large number of nodes and complex connection relationships [1-3], such as various networks in the life sciences(cellular networks, protein networks, neural networks, ecological networks), Internet network, social networks, scientific collaboration networks. Large number of networks show seemingly unrelated, but exist many striking similarities. According to different degree distribution, the network is divided into four kinds: regular network, random network, small-world network, scale-free network. In the work, we briefly review random network and scale-free network. For the network of random equal probability, the node degree distribution follows Poisson distribution [4], namely, $P(k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$. Afterwards, A. L. Barabasistudied the topology structure of WWW network [5] and find its properties is more complex than the random network, it turned out to be a continuously decreasing curve following the power law. It shows there are many web sites which have a small links, small number of web sites which have moderate number of links and several sites(hubs) which have a large number of links. Namely, the structure of WWW network is dominated by several sites which have many links [6]. A. L. Barabasi called the network to scalefree networks, which degree distribution follows $p(k_i)=\dfrac{k_i^{-\gamma}}{\xi_1(\gamma)}$,where k_i is the degree of the node i in complex network, γ is the parameter measuring network connection. Subsequently, the investigation showed the property is not unique to WWW network and there is universal significance, such as internet router connections [2], journal paper reference networks [7], and even sexual contact networks [8]. More recently, people generalized the scale-free network and get the fitness scale-free network, which degree distribution follows $p(k_i)=\dfrac{\eta_i k_i^{-\gamma}}{\xi_2(\gamma)}$, where, η_i is the fitness degree of node i, $\xi_2(\gamma)$ is Riemann zata, and defined as $\xi_2(\gamma)=\sum_{j=1}^{N}\eta_j k_j^{-\gamma}$,

k_i, γ are similar to those of the scale-free network.

The scale-free network shows a non-homogeneous nature and a kind of sequence the complex network emerges. It is not uniform, which degree distribution curve is continuously decreasing and the probability connected to other k nodes is proportional to $k^{-\gamma}$. γ characterizes non-homogeneity to some extent, γ increases, the decreasing speed of the degree distribution curve increases, then the non-homogeneity of networks is clearer.

In addition, the entropy is also used to measure the non-homogeneous nature of the network. Shannon [9] first introduced the thermodynamic entropy into the information theory and see the entropy as the uncertainty of a random event and a measure of the amount of information. Specifically, if the value of the random variable X is x_i, $i = 1,2,...,n$,and $x = \{x_i\}$ are pairwise incompatible, the probability of x_i is p_i , $i = 1,2,...,n$, $\sum_{i=1}^{n} p_i = 1$, Shannon proved $H(X) = -c\sum_{i=1}^{n} p_i \log(p_i) \quad (c > 0)$ is the only function satisfied the following conditions:

(i) H is the continuous function of $p_1, p_2,..., p_n$,

(ii) H get the maximum if and only if $p_1 = p_2 = ... = p_n$;

(iii) $H(X) = H(Y) + H(X/Y)$, where, $Y = f(X)$, $H(X/Y)$ is conditional entropy of X under the condition we know Y.

At the point, $H(X)$ is called the entropy of X. Let $c = 1$, people called $H(X) = -\sum_{i=1}^{n} p_i \log(p_i)$ to the traditional entropy.

If the distribution of random variable X is continuous, which distribution density function is $f(x)$, the entropy of X is defined as follows: $H(X) = -\int_R f(x)\log(f(x))$, where, R is the definition domain of $f(x)$.

Since the entropy is introduced, as a measure of the system stability [10, 11], it has become an important tool for studying the complex system and been extensively studied. For example, Cai [12] introduced a Caveman network and its evolution rules, the theoretical analysis and simulation experiments indicated that the Caveman network can effectively evaluated the sensitivity of different structure entropies on evolution process of the network and reflected the difference of ability to identify the properties of network complex of entropy indices. Shen [13] first took the micro-blog transmission network evolution as an example and applied the macro indicators in the study of the network's micro evolution based on the standard network structure entropy as a measurement index of ordering. Li [14] introduced four hybrid ratios into the unified hybrid network model, which were more in line with the randomness, the uncertainty and the variable growing in the real world network. Yan [15] proposed a new method for identifying key nodes in a complex network by means of combining the idea of the entropy weight method into the AHP algorithm. Cai [16] considered the difference between "node" and "edge" to define a new network structure entropy, and made theoretical analyses and the simulation experiments on regular network, random network and scale-free network. Zhao [17] proposed criteria stability entropy index based on the number of nonoverlapping paths for describing the invulnerability variation with nonoverlapping paths' number between nodes. Xu [18] proposed a virtual network mapping algorithm based on the entropy weight method. Long [19] explored a kind of network interaction mechanics process by taking advantage of entropy, built the network diffusion's system complexity of finite volume model and discussed the influence of local topology and routing capacity on the complexity. Liu [20] aligned clauses for Shi Ji ancient and modern parallel corpora using maximum entropy model and Back Propagation neural network model. He [21] proposed the uncertainty and complexity calculation method for network organization structure using entropy theory. Li [22] presented a evaluation method that can evaluate the switch performance in smart grid based on AHP-Entropy method and fuzzy-comprehensive evaluation theory. Zhu [23] proposed a computation model for network evolution based on entropy theory through summarizing the recent study of cooperation network evolution and provided more research perspectives for further analysis. Cheng [24] studied a novel network attack strategy evaluation method based on the conditional Shannon entropy and variable precision rough set. According to the diversity of micro grid's topology, through analyzing the theories of wavelet transform, singular value decomposition and extended shannon-entropy, the wavelet singular entropy could measure the fault signal, a fault diagnosis methods for the micro grid system was proposed by integrating the wavelet singular entropy with the self organizing feature map neural network [25]. Because the real-time modeling is difficult on thermal system and the model precision is not high and the convergence rate of neural network decreases dramatically when there are too many inputs, the BP NN modeling method based on information entropy was proposed in which the attribute reduction based on the model of approximation decision entropy was used [26]. The precision of user identification is low since the subjective weighting algorithms ignore the special meanings and effects of attributes in applications, to solve the problem, an information entropy based multiple social networks user identification algorithm was proposed [27]. Pan [28] developed an adaptive traffic classification using entropy-based detection and incremental ensemble learning, assisted with embedded feature selection; in order to update the classifier timely and effectively, the entropy-based detection utilizes sliding window technique to measure the statistical difference between the previous and current traffic samples by counting and comparing all instances with respect to their feature stream membership. Wang [29] studied the entropy theory of distributed energy for internet of things.

In the paper, we study the properties of the complex networks by way of entropy. First, we prove the consistency between the traditional entropy and the entropy with history information in section 2. Then we study the parameter estimation of the fitness scale-free network by methods of MLE, the moment estimation, the entropy and find the relationships among them in section 3. At last, we get the traditional entropy of the random network.

2. Comparison of Entropy

In order to seek the relationship between the traditional entropy and the entropy with history information, we first prove the equivalence of solutions of linear transformation before and after.

Theorem 2.1 The solutions of the objective functions which satisfies equality constrained linear transformation, namely,

$$\begin{aligned} &\max \mathrm{aF(x)+b} \quad \mathrm{st: ag(x)=0} \\ &\Leftrightarrow \max \mathrm{F(x)} \quad \mathrm{st: g(x)=0(i=1,2,\cdots,m)} \end{aligned} \tag{1}$$

proof

Construct Lagrange function for the objective functions $aF(x)+b$ is as follows,

$$Z(x,\lambda)=aF(x)+b+\sum_{i=1}^{m}\lambda_i a g_i(x),$$

then we get the partial derivative:

$$\frac{\partial Z}{\partial x_k}=a\frac{\partial F}{\partial x_k}+\sum_{i=1}^{m}\lambda_i a\frac{\partial g_i}{\partial x_k}=0,(k=1,2,\cdots,m)$$

eliminate a of on both sides and get

$$\frac{\partial F}{\partial x_k}+\sum_{i=1}^{m}\lambda_i\frac{\partial g_i}{\partial x_k}=0,(k=1,2,\cdots,m)$$

This is the same as Lagrange partial derivative equation of the primary target function. So they have the same solution, the equivalence is proved.

Subsequently, we prove the consistency between the traditional entropy and the entropy with history information of the complex networks using theorem 2.1. We suppose that the historical degree distribution $q(k_i)$ is uniform distribution, namely, $q(k_i)=\frac{1}{N}$, where, N is the number of node in the complex network.

theorem2.2

$$\min\sum_{i=1}^{N}p(k_i)\log\frac{p(k_i)}{q(k_i)}\Leftrightarrow\max-\sum_{i=1}^{N}p(k_i)\log p(k_i)$$

proof

$$\begin{aligned}&\sum_{i=1}^{N}p(k_i)\log\frac{p(k_i)}{q(k_i)}\\&=\sum_{i=1}^{N}p(k_i)\log(Np(k_i))\\&=\sum_{i=1}^{N}p(k_i)\log(N+\log p(k_i))\\&=\log N+\sum_{i=1}^{N}p(k_i)\log(p(k_i))\end{aligned}$$

Using the result of theorem 2.1, the theorem is proved.

From theorem 2.2, we find when the connection probability of all nodes in the history information is equal, the entropy model with historical information and the traditional entropy model are equivalent. Since each network varies, the entropy model with historical information shows its superiority. We see the historical information $q(k_i)$ as a weight vector, the weight vector reflects the connection of different nodes, then different explanations for the weight vector produce different degree distribution model. The choice of the weight value is worth further study.

3. The Parameter Estimation of Fitness Model

The scale-free network and its generalized model with a power-law distribution are important models of studying complex systems. The determination of the power-law exponent also becomes a focus point both in theory and application studies. The linear fitting of empirical data is often used for solving the exponent of scale-free distribution, but it is not accurate. The following work shows the parameter estimation of the scale-free fitness model based on MLE, the moment estimation and the entropy.

TMLE is based on the assumption that what has happened is the event originally to occur with the maximum probability. MLE of the exponent is the estimation of $\log k$. The log-likelihood function of the fitness scale-free model is given by

$$l(\gamma\,|\,k)=\prod_{i=1}^{N}\frac{\eta_i k_i^{-\gamma}}{\xi_2(\gamma)}$$

$$\begin{aligned}&L(\gamma\,|\,k)\\&=\log l(\gamma\,|\,k)\\&=\sum_{i=1}^{N}(\log\eta_i-\gamma\log k_i-\log\xi_2(\gamma))\\&=\sum_{i=1}^{N}(\log\eta_i-\gamma\log k_i)-N\log\xi_2(\gamma)\end{aligned}$$

Where, $l(\gamma \mid k)$ is the likelihood function of γ giving the observed data $k = k_i, 1 \le i \le N$, and $l(\gamma \mid k)$ is a log-likelihood function. The maximum can be obtained by finding the zero of the derivative of the log-likelihood function. Let

$$\frac{dL(\gamma \mid k)}{d\gamma} = -\sum_{i=1}^{N} \log k_i - N \frac{1}{\xi_2(\gamma)} \frac{d\xi_2(\gamma)}{d\gamma} = 0$$

we can get

$$\frac{\xi_2^{'}(\gamma)}{\xi_2(\gamma)} = -\frac{1}{N} \sum_{i=1}^{N} \log k_i \quad (2)$$

$$\frac{d^2 L(\gamma \mid k)}{d\gamma^2} = -\frac{1}{N} \frac{\xi_2^{''}(\gamma)\xi_2(\gamma) - (\xi_2^{'}(\gamma))^2}{\xi_2^2(\gamma)} < 0 \quad (3)$$

Then $(\gamma, l(\gamma \mid k))$ satisfied "Eq. (2)"and "Eq. (3)"is the maxima.

To get the relation between the moment estimator and MLE, we let

$$Y = \log(k), EY = \overline{Y}$$

And $k_1, k_2, ..., k_N$ be the observed values of k in the moment estimator, then we have

$$E \log(k) = \overline{\log(k)},$$

i.e. $\sum_{i=1}^{N} \log k_i \frac{\eta_i k_i^{-\gamma}}{\xi_2(\gamma)} = \frac{1}{N} \sum_{i=1}^{N} \log(k_i)$,

namely $(\log \xi_2(\gamma))' = -\sum_{i=1}^{N} \log(k_i)$,

which is equivalent to the result of MLE.

The above parameter estimate has very limited meaning without the analysis and assessment of its goodness of fit. Since the entropy obtained by Shannon is finite, we give the entropy of fitness model as follows:

$$H(k) = -\sum_{i=1}^{N} p(k_i) \log p(k_i)$$

$$= -\sum_{i=1}^{N} p(k_i) \log \frac{\eta_i k_i^{-\gamma}}{\xi_2(\gamma)}$$

$$= -\sum_{i=1}^{N} p(k_i) \log \eta_i + \gamma \sum_{i=1}^{N} p(k_i) \log k_i + \sum_{i=1}^{N} p(k_i) \log \xi_2(\gamma)$$

$$= -\sum_{i=1}^{N} p(k_i) \log \eta_i + \gamma E(\log k_i) + \log \xi_2(\gamma)$$

$$= -\frac{L(\gamma)}{N} - \sum_{i=1}^{N} p(k_i) \log \eta_i$$

From the entropy and MLE of the fitness scale-free network, we notice that likelihood function L get the maxima, entropy H gain the minimal value. On the contrary, when likelihood function L get the minimal value, entropy H gain the maxima.

It is important to notice the relation between the value of random variable and the entropy of the distribution. The former is a measure of the variation, while the entropy is the measure of uncertainty of the probability distribution. The entropy is only dependent on the probability distribution and has no relation with the random variable value. The two concepts are not equivalent except one of them taking the value of zero.

At last, we give the entropy of the random network. During the process of calculation, we use the stirling factorial formula $\log x! = x \log x - x$. In the random network, the degree distribution obeys the following probability distribution:

$$p(k_i) = \frac{\lambda^{k_i} e^{-\gamma}}{k_i!}$$

Then

$$H(k) = -\sum_{i=1}^{N} p(k_i) \log p(k_i)$$

$$= -\sum_{i=1}^{N} p(k_i) \log \frac{\lambda^{k_i} e^{-\gamma}}{k_i!}$$

$$= -\sum_{i=1}^{N} k_i p(k_i) \log \lambda + \gamma \sum_{i=1}^{N} p(k_i) + \sum_{i=1}^{N} p(k_i) \log(k_i!)$$

$$= -\log \lambda \sum_{i=1}^{N} k_i p(k_i) + \gamma + \sum_{i=1}^{N} p(k_i)(k_i \log k_i - k_i)$$

$$= -\log \lambda Ek + \lambda + \sum_{i=1}^{N} k_i p(k_i) \log k_i - Ek$$

$$= -(1 + \log \lambda) Ek + E(k \log k) + \gamma$$

We can quantitatively analyze the parameter of random network in order to understand the random network better, which needs be further studied by methods of the entropy.

4. Conclusions

In the paper, we study the traditional entropy and the entropy with history information of the complex network and get their consistence. Then we compare different ways of parameter estimation and find MLE and the moment estimator have the same result, but, MLE and the entropy is converse, which provides theoretical support for our future study of complex systems. We'll continue to study entropy and the relationship between different kinds of entropy in the future.

Acknowledgements

The research was supported by National Natural Science Foundation of China (11801012) and Anyang Normal University

Students innovation Fund project (X2020104790142).

References

[1] S. Wasserman and K. Faust, Social network analysis: methods and applications. Cambridge, England: Canbridge University Press, 1994.

[2] M. Faloutsos, P. Faloutsos and C. Faloutsos, “On power-law relationships of the internet topology,” Comput Commun Rev., 1999, 29, 251-262.

[3] S. Lawrence and C. L. Giles, “Searching the world wide web,” Science, 1998, 280, 98-100.

[4] P. Erdos and A. Renyi, “On random graphs,” Publicationes Mathematicae, 1959, 6, 290-297.

[5] A. L. Barabasi, R. Albert and H. Jeong, “Mean-field theory for scale-free random networks,” Physica A, 1999, 272, 173-187.

[6] R. Albert, H. Jeong and A. L. Barabasi, “Diameter of the world-wide web,” Nature, vol. 401, pp. 130-131, 1999.

[7] S. Redner, “How popular is your paper? An empirical study of the citation distribution,” Eur. Phys. J. B., 1998, 4, 131-134.

[8] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley and Y. Aberg, “The web of human sexual contacts,” Nature, 2001, 411, 907-908.

[9] C. E. Shannon, “A mathematical theory of communication,” Bell Syst, Tech. J., 1948, 27, 379-423.

[10] A. G. Wilson, “A statistical theory of spatial distribution models,” Transportation Research, 1967, 1, 253-269.

[11] A. G. Wilson, Entropy in urban and Regional modelling. London, England: PionPress, 1970.

[12] M. Cai, H. F. Du and F. W. Marcus, “Caveman network and its application in analysis of complex network entropy,” Systems Engineering-Theory and Practice, 2017, 37, 2403-2412.

[13] Q. Shen, Y. Huang, N. Ma and Y. J. Liu, “Entropy reduction point in complex networks: Taking the evolution of micro-blog transmission network as an example,” Mathematics in Practice and Theory, 2015, 45, 282-290.

[14] Y. LI, J. Q. Fang, Q. Liu, “An entropy approach to complexity of networks generated with the unified hybrid network model: complexity of complex systems,” Science and technology review, vol. 35, pp. 56-62, 2017.

[15] D. Yan, S. B. Zhang, K. Zong and Z. H. Hu, “Identification of key nodes in a complex network based on AHP-entropy method,” Journal of Guangxi University (Nat. Sci. Ed.), vol. 41, pp. 1933-1939, 2016.

[16] M. Cai, H. F. Du, Y. K. Ren and F. W. Marcus, “A new network structure entropy based node difference and edge difference,” Acta. Phys. Sin., vol. 60, pp. 1105131-1105139, 2011.

[17] J. X. Zhao, “Evaluation method of network invulnerability based on nonoverlapping paths entropy,” Application Research of Computers, vol. 32, pp. 825-826, 851, 2015.

[18] Q. Xu, H. Y. Yi, J. Zhu, et al, “Virtual network mapping algorithm based on entropy weight method,” Computer Engineering and Applications, vol. 51, pp. 94-99, 2015.

[19] Y. H. Long, “Analysis of entropy into finite volume models scale-free network diffusion,” Journal of Hubei University of Arts and Science, vol. 36, pp. 9-12, 2015.

[20] Y. Liu, N. Wang, “Comparison of clause alignment based on maximum entropy model and Back Propagation neural network model,” Computer Engineering and Application, vol. 51, pp. 112-117, 2015.

[21] X. J. He, Y. J. Tan, Y. Y. Wu, “Entropy-based on measurement of the complexity of network organization and its empirical analysis,” Systems Engineering, vol. 34, pp. 154-158, 2016.

[22] Y. Z. Li, Q. Liu, B. Gao, et al, “Probe into evaluation method of switch network performance in smart grid based on AHP-entropy method,”ShanXi Electric Power, vol. 44, pp. 29-33, 2016.

[23] Y. X. Zhu, “The measure research of cooperation network evolution based on entropy theory,” Journal of Intelligence, vol. 36, pp. 183-188, 2017.

[24] R. Cheng, J. Lei, J. Cheng “Evaluation method of network attack strategy based on conditional Shannon entropy and VPRS,” Journal of WuT (Information and management engineering), vol. 39, pp. 162-167, 2017.

[25] L. Qiu, Y. Z. Ye, C. D. Jiang “Fault diagnostic method for microgrid based on wavelet singularity entropy and SOM neural network,” Journal of Shandong University (Engineering Science), vol. 47, pp. 118-122, 2017.

[26] H. R. Sun, R. Wang, J. Y. Geng “Thermal system modeling based on entropy and BP Neural network” Journal of System Simulation, vol. 29, pp. 226-233, 2017.

[27] Z. Wu, H. T. Yu, S. X. Liu, et al, “User identification across multiple social networks based on information entropy,” Journal of Computer Applications, vol. 37, pp. 2374-2380, 2017.

[28] W. B. Pan, G. Cheng, X. J. Guo, et al, “An adaptive classification approach based on information entropy for network traffic in presence of concept drift,” Chinese Journal of Computers, vol. 40, pp. 1556-1571, 2017.

[29] Z. L. Wang, “Entropy theory of distributed energy for internet of things,” Nano Energy, 58, pp. 669-672, 2019.