Entropy description of a cooperation–competition system*

Xu Xiu-Lian(徐秀莲)[a], Fu Chun-Hua(傅春花)[a],
Liu Chun-Ping(刘春平)[b], and He Da-Ren(何大韧)[a]†

[a] College of Physics Science and Technology, Yangzhou University, Yangzhou 225009, China
[b] College of Mathematics Science, Yangzhou University, Yangzhou 225009, China

(Received 21 February 2010; revised manuscript received 12 March 2010)

Understanding the cooperation–competition dynamics is a long-standing challenge in studying complex systems. Inspired by the idea of Shannon entropy, we define competition information entropy and propose an entropy evolution model. The analytic results of the model of the relation between competition gain distribution parameters and entropy, as well as the relation between entropy and time are compared with empirical results obtained in 14 real world systems. They are found to be in good agreement with each other.

Keywords: cooperation–competition, entropy, competition gain, evolution model

PACC: 0565, 0590, 0175, 0250

1. Introduction

Simultaneous cooperation–competition of complex system elements has been a fascinating research topic in commercial,[1,2] ecological,[3−5] social,[6−8] biochemical,[8] economical,[8−10] and some technological systems.[11−19] Different models have been proposed for the description of the cooperation–competition dynamics.[4−10] Among the systems, the cooperation–competition networks described by bipartite graphs are of interest,[11−19] which contain two types of nodes.[20] One is called the "acts", which can be organizations, events, or activities, and the other is named the "actors", who are participants. The actors may be human beings, such as movie actors who take part in screening a movie (an act), or athletes who join a sport event (an act) in the Olympic Games.[18] They may also be something abiotic, but influenced by human beings. The actors, for example, can be some languages, which coexist in a geographical or political region (an act). After long consideration, the language can be regarded as being a competition to be used by more people. Cocktail ingredients may be mentioned as another example for actors, which compete in a cocktail to occupy more content.[18] Often in each act some actors make concerted efforts to accomplish a task and often create a type of production. The production often induces several kinds of resources. The actors, when they are cooperating, are also competing for a larger piece of the resources. For example, some Hollywood actors work together to produce a movie, which should bring ticket office income and reputation. The former is countable, while the latter is not. From a different viewpoint, some languages cooperate to create "oral and written communications" between different kinds of people in an act (region). The number of people who use a certain language can be regarded as a countable resource, which the actor (the language) shares. In order to describe the cooperation sharing or competition gain, Fu *et al.* defined an "actor node weight".[14] Basically it was defined as the part of a countable resource, which the actor shares or gains. However, the actual competition intensity of two actors should depend on the act size, T, which is defined as the number of the actors taking part in the act. If Z_l denotes the total countable resources in act l, and z_{il} denotes the part gained by actor i, the weight of actor i in act l is defined as $W_{li} = T_l z_{il}$. The normalized node total weight (NNTW) is defined as $\omega_i = \left(\sum_l W_{li}\right) / \sum_j \left[\left(\sum_l W_{lj}\right)\right]$.[15,17−19] In Ref. [18] we presented definitions and interpretations of 14 real world bipartite cooperation–competition networks. Some information is presented in Table 1. We empirically found that in all the networks the NNTW distributions obey the so-called "shifted power

*Project supported by the National Natural Science Foundation of China (Grant Nos. 10635040 and 70671089).

†Corresponding author. E-mail: darendo10@yahoo.com.cn

© 2010 Chinese Physical Society and IOP Publishing Ltd

law (SPL) functions".[12,18] The function can be expressed as $P(\omega) \propto (\omega + \alpha)^{-\gamma}$. When $\alpha = 0$, it takes a power law form. On condition that ω is normalized ($0 < \omega_i < 1$ and $\sum_{i=1}^{L} \omega_i = 1$, where L denotes the total number of ω), we can prove[18] that the SPL function turns into an exponential function when $\alpha \to 1$. Therefore an SPL interpolates between a power law and an exponential function. Actually, similar and relevant interpolated distribution functions, such as the generalized Pareto distribution,[21,22] the stretched exponential distribution,[23] the Pareto–Lévy law[24] and the Yule–Simon distribution[25] have been proposed.

Table 1. Definitions and parameters of 14 real world collaboration–competition networks.

real world networks	act	actor	NNTW	$N\alpha$	γ	τ/year
2004 Athens Olympic	sport events	athletes	sport scores	4500	3617	2.28×10^{-4}
university matriculation	batches	universities	lowest matriculation marks	2277	3684	0.016
supermarket	commodities	supermarkets	buyer's marks	6.27	8.21	no data
PC selling	notebook PC types	shops	selling price	4.15	6.063	3
IT product	products	factories	order of evaluation	2.86	4.465	28
undergraduate courses	scientific subjects	general courses	number of students on a course	1.21	3.16	55
university recruitment	university	middle school	number of matriculated students	0.897	2.696	no data
human acupuncture point	diseases	acupuncture points	acupunctured number	0.842	2.552	no data
book borrowing	scientific subjects	book	number of books borrowed	0.035	2.187	no data
Beijing restaurant	dishes	restaurants	grades by clients	0.122	2.118	75
training institution	training courses	institutions	charges of a course	0.174	2.100	7
mixed drink	drink	ingredients	content of an ingredient	0.0766	1.783	340
China mainland movie	actors	movies	number of downloads	0.0497	1.662	26
world language	countries	languages	population of a language	0.0086	1.632	2520

Since both of the NNTW distribution parameters, α and γ describe the heterogeneity of the competition gain distribution,[15,19] after reading Table 1, one may ask why the real world networks show such different heterogeneity. To answer the question, an evolution model of a network property relating to the heterogeneity must be proposed. In this article we present a model of network entropy evolution. It is well known that the Shannon entropy in information theory can be expressed as $S_S = -k_S \sum_{i=1}^{n} p_i \ln p_i$ where p_i denotes the appearance probability of the i-th state, and k_S is the famous Shannon entropy constant. The entropy is a measure of uncertainty, or the averaged information content one needs for knowing the value of the random variable.[26] If p_i=1, and all the p_j ($j \neq i$) =0 (the most heterogeneous circumstance), $S_S = 0$ which means that no further information is needed for the prediction. If $p_i = 1/n$ ($i = 1, \ldots, n$, the most homogeneous circumstance), $S_S = k_S \ln n$, the largest information content is requested. Inspired by this idea, we will propose an entropy description of the NNTW distribution so that an understanding of cooperation–competition network evolution can be obtained. The rest of the present article is organized as follows. In Section 2 we will propose the definition of the entropy. The network evolution model by means of the entropy and the analytic discussion on the relation between the NNTW distribution parameters and the entropy, as well as the relation between entropy and time, will be presented in Section 3. In Section 4 we will compare the analytic results with the empirical results obtained for the aforementioned 14 real world networks. In the last section, Section 5, a summary and discussion will be presented.

2. Competition information entropy

Consider a "one winner game" where N players compete for only one award. Let the game repeat M times, if the i-th player wins m times then $u_i = m/M$ can be considered as his winning probability on condition that M is large enough. In the light of Shannon entropy, we may define $S_u = -\sum_{i=1}^{N} u_i \ln u_i$ as the "competition information entropy" of the system, which interprets the total information content needed for a winner prediction. If the i-th player wins all the M times, i.e. he plays much better than the others and certainly wins every time, we have u_i=1, and all the u_j ($j \neq i$) =0, then S_u=0, one can make the easiest prediction. If $u_i = 1/N$ (each player has the same probability of winning), then $S_u = \ln N$; the prediction is the most difficult to make and the largest infor-

mation content is required. Correspondingly, $u_i \ln u_i$ can be considered as the information content needed for the i-th winner prediction.

For the aforementioned real world cooperation–competition networks, in each empirically observed SPL distribution function, $P(\omega) \propto (\omega + \alpha)^{-\gamma}$, usually the values of ω are different for different values of i, while α remains the same for all the actors. Therefore, we can regard α as a part of the competition gain, which is evenly distributed due to a game rule. We assume that all the systems start to evolve at a common time. The reason for this will be explained in Section 4. With this assumption the total competition gain $\omega + \alpha$ becomes a function of common evolution time. Then we can define the normalized evolutionary competition gain $v_i = (\omega_i + \alpha)/(\sum_j (\omega_j + \alpha))$ as the winning probability of the i-th actor. If the i-th actor gains all the resource at all times, i.e., he always wins, then $v_i(t) = 1$. If he always gets nothing, the $v_i(t)$=0. With a similar consideration, we define $S_v = -\sum_{i=1}^{N} v_i \ln v_i$ as the competition information entropy of the networks, which again denotes the total information content needed for a winner prediction. Also $v_i \ln v_i$ is the information content needed for the i-th winner prediction.

3. Entropy evolution model and analyses

The only important assumption in our entropy evolution model is that the increase of the i-th actor winning probability per time unit is proportional to the decrease of its competition information entropy (the information content needed for the i-th winner prediction) per unit time. This is consistent with the entropy definition, therefore we believe that it should be universal at least for a type of cooperation–competition network. Of course, the assumption must be verified by comparison with the empirical investigations, which will be presented in the next section. The model can then be described by the following equation:

$$\frac{\partial v_i}{\partial t} = c'\left(\frac{v_i \ln v_i}{t - t_\delta} - \frac{v_i \ln v_i}{t}\right), \tag{1}$$

where v_i denotes the winning probability of the i-th actor, t_δ represents a very small positive constant, and c' is a positive proportional constant. We label the values of v from 1 through N in decreasing order, i.e., $0 \geq v_1 \geq v_2 \geq \cdots \geq v_i \geq \cdots \geq v_N \geq 1$ and $\sum_{i=1}^{N} v_i = 1$. Solving the equation, we obtain

$$\ln(-\ln v_i) = \frac{c}{t_\delta} \ln\left(\frac{t - t_\delta}{t}\right) + B_i', \tag{2}$$

where c is a constant relating to c', and B_i' is a constant depending on i. This leads to

$$v_i = B_i^{-[(t-t_\delta/t)]^{c/t_\delta}}, \tag{3}$$

where B_i is a constant relating to B_i'. Consider variable v which locates in $[B_{i\text{ -min}}^{-[(t-t_\delta/t)]^{c/t_\delta}}, B_{i\text{ -max}}^{-[(t-t_\delta/t)]^{c/t_\delta}}]$, where $B_{i\text{ -min}}$ and $B_{i\text{ -max}}$ denote the minimum and the maximum values of B_i, respectively. Letting $v_i < v$, one has

$$p(v_i) = \frac{\partial p(v_i < v)}{\partial v_i} \propto v_i^{-[t/(t-t_\delta)]^{c/t_\delta} - 1} = v_i^{-\gamma(t)}, \tag{4}$$

where

$$\gamma(t) = \left(\frac{t}{t - t_\delta}\right)^{c/t_\delta} + 1 \approx 2 + \frac{c}{t} > 2.$$

Thus we obtain

$$\gamma - 2 \approx \frac{c}{t} \propto t^{-1}. \tag{5}$$

Equation (5) implies that this evolution model is suitable only for real world networks which show NNTW distributions with γ values larger than 2. According to the discussion in Ref. [27] and the references therein, the rank number i also obeys a similar power law distribution. Based on this relationship and the empirical conclusion $P(v_i) \propto v_i^{-\gamma}$, $P(v_i' \geq v) \propto v_i^{1-\gamma}$, we have

$$P(v_i' \geq v) = \frac{i}{N} \propto v_i^{1-\gamma} \tag{6}$$

and

$$v_i \propto i^{-1/(\gamma-1)}. \tag{7}$$

In order to compare the entropy with the normalized v and α, we redefine a normalized competition information entropy as $S_{\rm nv} = -(\sum_{i=1}^{N} v_i \ln v_i)/(\ln N)$. From the definition and Eqs. (5)–(7), one can obtain

$$S_{\rm nv} = \left(-\sum v_i \ln v_i\right)/(\ln N) \approx \left[-\frac{\int i^{-1/\gamma-1} \ln i^{-1/\gamma-1} \mathrm{d}i}{\int i^{-1/\gamma-1} \mathrm{d}i} + \ln\left(\int i^{-1/\gamma-1} \mathrm{d}i\right)\right] / \ln N. \tag{8}$$

This leads to

$$S_{\rm nv}(\gamma) \approx \left\{ \frac{[1/(\gamma-1)]N^{1-1/(\gamma-1)} \ln N}{N^{1-1/(\gamma-1)}-1} + \ln\left(\frac{N^{1-1/(\gamma-1)}-1}{1-1/(\gamma-1)}\right) - \frac{1/(\gamma-1)}{1-1/(\gamma-1)} \right\} / \ln N, \tag{9}$$

and

$$S_{\rm nv}(t) \approx \left\{ \frac{[t/(t+c)]N^{c/(t+c)} \ln N}{N^{c/(t+c)}-1} + \ln\left(\frac{N^{c/(t+c)}-1}{c/(t+c)}\right) - \frac{t}{c} \right\} / \ln N. \tag{10}$$

In Ref. [19] we obtain an analytic and empirical conclusion, $\gamma = 2 + N\alpha$, and by the above-mentioned relation $\gamma \approx 2 + c/t$, we obtain $N\alpha \approx c/t$. From Eq. (10) we have

$$S_{\rm nv}(N\alpha) \approx \left[\frac{N^{N\alpha/(1+N\alpha)} \ln N}{(1+N\alpha)(N^{N\alpha/(1+N\alpha)}-1)} + \ln\left(\frac{N^{N\alpha/(1+N\alpha)}-1}{N\alpha/(1+N\alpha)}\right) - \frac{1}{N\alpha} \right] / \ln N. \tag{11}$$

4. Comparison between analytic and empirical conclusions

In this section we will present a comparison between the analytic and the empirical conclusions about the relation between entropy $S_{\rm nv}$ and evolution time t, as well as the relation between entropy and NNTW parameters $N\alpha$ and γ. We have to say that this job is not easy. Our group investigated 34 real world cooperation–competition networks.[28] However, for 10 of them we were unable to find the data about NNTW and the data number of another 10 was not large enough to obtain reliable statistics, so only 14 real world networks could be used for an $S_{\rm nv}$–γ relation comparison.[18,19] Among them, we could find time evolution information only for 10 networks.[18] It was even worse that only 7 of the 10 networks showed parameter γ values larger than 2[18] and thus could be used for comparing $S_{\rm nv}$–t and $S_{\rm nv}$–$N\alpha$ relations (actually the $S_{\rm nv}$–γ relation deduction does not relate to the time evolution). Since it is so hard to find suitable systems, we really cannot select the systems to suit the analytic conclusion even if we wanted to do so. Therefore the 14 or 7 real world networks may be considered to be obtained through random selection. They can serve as representatives of a large number of cooperation–competition networks. The main information about the 14 networks is shown in Table 1.

There appears another problem in comparing the $S_{\rm nv}$–t relation. Obviously, it is almost impossible to obtain data about the system evolution at all times. What are listed in Table 1 are values of the so-called "system evolution time duration" τ. In most cases, the cooperation–competition between the actors occurs right after the birth of the acts; therefore, we can define the cooperation–competition evolution time duration as the time duration between the act birth and the act termination or data collection. We define the "act birth" as the time when the first actor joins it, and the "act termination" as the time when the cooperation task is accomplished and the actors disband. In the second section, we have assumed that all the networks start to evolve at a common time. If we also assume that every system obeys the entropy evolution model, all the systems progress with a common time. Different systems show different entropies and NNTW distribution parameters only due to their different evolution time durations. The empirically obtained relation between entropy $S_{\rm nv}$ and evolution time duration τ, and the relation between entropy and NNTW parameters $N\alpha$ and γ should approximately obey Eqs. (9)–(11).

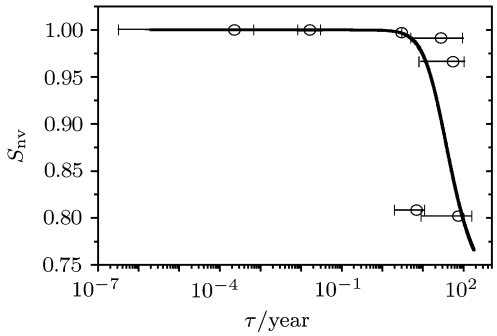


Fig. 1. Relations between evolution time duration and competition information entropy for 7 real world networks. The wider solid line shows the model analytic conclusion, Eq. (10). The circles show the empirical data. The thinner bars show the time duration error ranges of the data.[18]

We emphasize that the description of the cooperation–competition dynamics will certainly be in

a logarithmic time scale because the evolution time durations of different systems must be very different and cross many orders of magnitude.

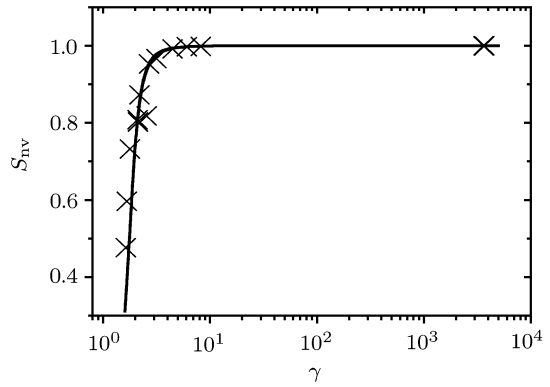


Fig. 2. Relations between competition information entropy and NNTW distribution parameter γ for 14 real world networks. The solid line shows the model analytic conclusion, Eq. (9). The crosses show the empirical data.

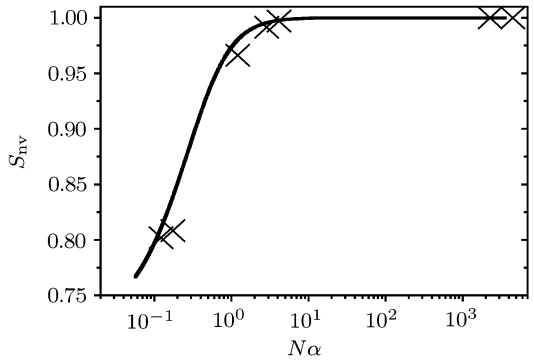


Fig. 3. Relations between competition information entropy and NNTW distribution parameter α for 7 real world networks. The solid line shows the model analytic conclusion, Eq. (11). The crosses show the empirical data.

Therefore, it is enough to obtain reliable data of the duration in order of magnitude. In Ref. [18] we presented the evolution time durations of the 10 systems and explained how the data were obtained and why the data were reliable in order of magnitude. As can be seen in Figs. 1–3, the agreement between analytic and empirical conclusions is fairly good.

5. Summary and discussion

We proposed two quantities: winning probability v and competition information entropy $S_{\rm nv}$. Based on the definitions, a micro-dynamic evolution equation (1) is proposed, which only means that for an actor, the smaller the competition information entropy becomes in the evolution, the larger the winning probability that he can obtain will be. Since we already empirically know that the distribution of v probability obeys a power law, the macro-dynamic relation between system entropy $S_{\rm nv}$ and imaginary evolution time τ, as well as the relations between $S_{\rm nv}$ and competition gain distribution parameters, γ and α, can be analytically deduced. The conclusions then are compared with the empirical results. Although the data number is small, the comparison still shows a good agreement. We thus believe that our model is reasonable at least for some of the cooperation–competition networks.

Now we can answer the question we asked before: why do real world networks show such different heterogeneities? From Figs. 2 and 3 one can see that $S_{\rm nv}$ is monotonically dependent on either γ or α. Since we already know that both γ and α describe the heterogeneity of the competition gain distribution (the smaller either γ or α is, the larger the heterogeneity will be[15,19]) we obtain a conclusion that $S_{\rm nv}$ also monotonically shows the heterogeneity of the competition gain. This is in agreement with our qualitative discussion in Section 2. Figure 1 clearly shows the monotonic decreasing tendency of $S_{\rm nv}$ as the imaginary time develops. This means that in all the real world systems discussed in the current study, the heterogeneity becomes increasingly large with a universal dynamics. The competition gains gradually concentrate onto increasingly fewer actors. Different real world networks show very different heterogeneities of the competition gain distribution only due to the fact that they are confined in very different evolution time durations. This is our main finding from the model.

References

[1] Bengtsson M and Kock S 2000 *Industrial Marketing Management* **29** 411

[2] Wang Y S and Wu H 2004 *Physica* A **339** 609

[3] Law R, Marrow P and Dieckmann U 1997 *Evolutionary Ecology* **11** 485

[4] Guimaraes Jr P R, Machado G, Aguiar M A M, Jordano P, Bascompte J, Pinheiro A and Furtado dos Reis S 2007 *J. Theor. Bio.* **249** 181

[5] Saavedra S, Reed-Tsochas F and Uzzi B 2009 *Nature* **457** 463

[6] Puurtinen M and Mappes T 2009 *Proc. R. Soc.* B **276** 355

[7] Pugliese E, Castellano C, Marsili M and Pietronero L 2009 *Eur. Phys. J.* B **67** 319
[8] Albornoz J M and Parravano A arXiv:0904.2741v2 [nlin.AO] 19 Jan 2010
[9] Scheffran C and Hannon B 2007 *Complexity* **13** 78
[10] Mohanty P K 2007 *Physica* A **384** 75
[11] Zhang P P, Chen K, He Y, Zhou T, Su B B, Jin Y D, Chang H, Zhou Y P, Sun L C, Wang B H and He D R 2006 *Physica* A **360** 599
[12] Chang H, Su B B, Zhou Y P and He D R 2007 *Physica* A **383** 687
[13] Zhou T, Wang B H, Jin Y D, He D R, Zhang P P, He Y, Su B B, Chen K, Zhang Z Z and Liu J G 2007 *Int. J. Mod. Phys. C* **18** 297
[14] Fu C H, Zhang Z P, Chang H, Tao J R, Chen Z H, Dai Y L, Zhang W and He D R 2008 *Physica* A **387** 1411
[15] Fu C H, Xu X L, Liu A F, Wu Y P, Shen D, Liu S J, Qian X, Feng Y C, Wei C L and He D R 2008 *Chin. Phys. Lett.* **25** 4181
[16] Chen Y Z, Fu C H, Chang H, Li N and He D R 2008 *Chin. Phys.* B **17** 3580
[17] Shi J J, Wang Y L and He D R 2009 *Chin. Phys. Lett.* **26** 078902
[18] Xu X L, Fu C H, Shen D, Liu A F and He D R arXiv:0910.2516v2 [physics.soc-ph]
[19] Xu X L, Fu C H, Liu A F and He D R 2010 *Chin. Phys. Lett.* **27** 048901
[20] Chen H B, Fan Y, Fang J Q and Di Z R 2009 *Acta Phys. Sin.* **58** 1383 (in Chinese)
[21] Pickands J 1975 *The Annals of Statistics* **3** 119
[22] Hosking J R M and Wallis J R 1987 *Technometrics* **29** 339
[23] Laherrere J and Sornette D 1998 *Eur. Phys. J.* B **2** 525
[24] Mandelbrot B 1960 *Int. Econo. Rev.* **1**(2) 79
[25] Yan X Y and Wang M S 2010 *Acta Phys. Sin.* **59** 851 (in Chinese)
[26] Shannon C E 1948 *Bell System Technical Journal* **27** pp. 379–423, 623–656
[27] Hu H B and Wang X F 2008 *Physica* A **387** 3769
[28] Liu S J, Qian X, Feng Y C, Fu C H, Xu X L, Liu A F, Wu Y P and He D R 2008 *Review of Science & Technology* **26** 79 (in Chinese)