Journal of Library and Information Studies 18:2 (December 2020) pp.31-41
doi: 10.6182/jlis.202012_18(2).031

Quantifying Interactions between Layers in Duplex Networks Using Interactive Entropy

Ronda J. Zhang[1,2], Fred Y. Ye[1,2]

Abstract

The multiplex network has attracted wide attention because it can better describe the various relationships in the real world, and the duplex network is a special two-layer form of it. The purpose of this research focuses on introducing new entropy-based measures for multiplex network, especially duplex network. The new measure called Interactive entropy in duplex network (IE) is based on information entropy, a key concept of information theory. It can be applied to various types of multiplex networks to quantify the differences between the layers of inside duplex network.

Keywords: Multiplex Network; Duplex Network; Interactive Entropy

1. Introduction

There is a great growth in the studies of complex networks in recent years. While the need to better simulate the real world is increasing faster than ever before, the existing single-layer network model (Albert & Barabási, 2002; Barabási & Albert, 1999; Strogatz, 2001) has not been able to adapt to this change. Therefore, multilayer networks (Boccaletti et al., 2014; de Domenico et al., 2013) have attracted more and more attention, especially a simplified version called multiplex network (Battiston, Nicosia, & Latora, 2014; Solé-Ribalta, de Domenico, Gómez, & Arenas, 2014), where all the nodes are the same in different layers.

A multiplex network is a system formed by N nodes and M layers of interactions where each node belongs to the M layers at the same time. Each layer α is formed by a network G_α. Now we focus on the two-layer multiplex network, which is called as duplex network. This is a reliable and useful way to simplify the whole complex system.

Viewing the network system from the perspective of information and constructing a measure that contains the overall information of the network are the key issues in network information theory (Shi, Chen, Long, Wang, & Pan, 2019). The information theory leaves us with many tools to measure the amount of information, among which the most important one is entropy (Shannon, 1948). It has also been successfully introduced into the field of network science, where many different types of network entropy have been introduced (Anand & Bianconi, 2009;

[1] International Joint Informatics Laboratory (IJIL), Nanjing University – University of Illinois, Nanjing-Champaign, Nanjing, China
[2] Jiangsu Key Laboratory of Data Engineering and Knowledge Service, School of Information Management, Nanjing University, Nanjing, China
* Corresponding Author: Fred Y. Ye, E-mail: yye@nju.edu.cn

De Domenico & Biamonte, 2016; Zhang, Li, & Deng, 2018).

In defining the various concepts of entropy, the most significant thing is to find the right probability distribution. The remaining degree distribution (Solé & Valverde, 2004) is used to form the entropy and mutual information, while the degree distribution (Wang, Tang, Guo, & Xiu, 2006) is used to form another kind of entropy to study the robustness of networks to random failures. There are several common types of network entropy (Cai, Du, & Feldman, 2014; Cai, Du, & Ren, 2011; Wu, Tan, Deng, & Zhu, 2007), and some scholars have conducted in-depth research on their similarities and differences (Cai, Cui, & Stanley, 2017). These entropies are all built to quantify network information based on the distribution of a certain characteristic of the network. Also, it is worth mentioning that the previous scholars are all defined in the context of single layer complex network and are proven effective to capture the uncertainty of networks.

In this research, we focus on some entropy-based measures that are suitable for duplex networks aiming to characterize the structure and functional features of the duplex networks. Initially, we introduce one entropy-based measure based on the degree distribution to reflect the characteristics of duplex network. Then we verify the validity of these measures in various types of duplex networks. Finally, we compare our measure with previous ones and expand the probability set used in it.

2. Methodology

In a single layer network, the degree distribution is

$$p(k) = \frac{N_k}{N} \qquad (1)$$

where N_k is the number of nodes with degree k in the network. The degree distribution here represents the probability of taking a node arbitrarily in a certain network, whose degree happens to be k. It can be an average measure of a network's heterogeneity. It is suitable for single layer networks. Since the probability distribution used here when calculating entropy is degree distribution, and the degree distribution is more suitable for unweighted networks (Zhao, Rousseau, & Ye, 2011; Zhao & Ye, 2012), the entropy of the degree distribution is also more suitable for unweighted networks.

In order to make it more appropriate for multiplex network, we extend the definition via simplifying the multiplex network into several two-layer networks. Follow the definition of the previous work (Zhang & Ye, 2020), a duplex network is defined as a multiplex network composed of two layers, where there are the same N nodes in each layer.

The degree distribution of layer α ($\alpha = 1,2$) is $p(k)^{[\alpha]}$

$$p(k)^{[\alpha]} = \frac{N_k^{[\alpha]}}{N} \qquad (2)$$

where $N_k^{[\alpha]}$ is the number of nodes with degree k in the network.

Definition: Interactive entropy in duplex network (IE) is.

$$IE(\alpha = 1 \| \alpha = 2) = \sum_{k=1}^{N-1} p\left(k\right)^{[1]} \bullet \log \frac{p\left(k\right)^{[1]}}{p\left(k\right)^{[2]}} \qquad (3)$$

In information theory, relative entropy is a measure of the difference between two probability distributions (Kullback & Leibler, 1951). We use the same concept to define the interactive entropy in a duplex network.

When presenting the results later, we choose 2 as the base of the above logarithmic function, where the base can also be *e* or 10. When the base is 2, the value of the distribution entropy can be regarded as the minimum bits required to encode the information in each layer.

Some properties of IE can be derived. First, if the degree distribution of the two layers in the duplex network is exactly the same, then the interactive entropy is equal to zero. Second, as shown in Eq.(3), IE is asymmetric. The same conclusion can be drawn from the concept of relative entropy in information theory, i.e., the value of interactive entropy equals to the information loss generated when one layer is used to fit another layer in the duplex network.

In order to better explain the concept of IE, we introduce a simple example. In Figure 1, Layer A and B form a duplex network when combined. The number on the node indicates the degree of the node. It can be easily observed from the figure that there are two types of nodes in layer A, namely nodes whose degree are1 and 2, while there are three types of nodes in layer B, which are nodes with degree 1, 2, 3.

It is also worth noting that when calculating interactive entropy, if any $p(k)^{[1]} = 0$ or $p(k)^{[2]} = 0$, then a very small amount ε should be assigned to these $p(k)^{[1]}$, $p(k)^{[2]}$, such as ε = 1e-5. The sum of the processed degree distributions is not equal to 1, so we normalize the newly obtained distribution to obtain $p'(k)^{[1]}$ and $p'(k)^{[2]}$. In this way, the interactive entropy of the two can be easily calculated. The same example could be used to clarify this. In Figure 1, $p(k)^{[A]} = \{p(1) = 0.5, p(2) = 0.5, p(3) = 0\}$ and $p(k)^{[B]} = \{p(1) = 0.25, p(2) = 0.5, p(3) = 0.25\}$. After the above progress, $p'(k)^{[A]} = \{p'(1) = 0.5/(1+\varepsilon), p'(2) = 0.5/(1+\varepsilon), p'(3) = \varepsilon/(1+\varepsilon)\}$. Interactive entropy, which means that the difference between network A and B can be calculated, and $IE(\alpha=A||\alpha=B)$ is 0.4998.

3. Studies

The interactive entropy can be utilized in a large variety of networks. These networks can be either weighted or unweighted from the perspective of network weight, and can be either information networks or social networks from the perspective of network type. Here are a few examples to demonstrate the effectiveness of

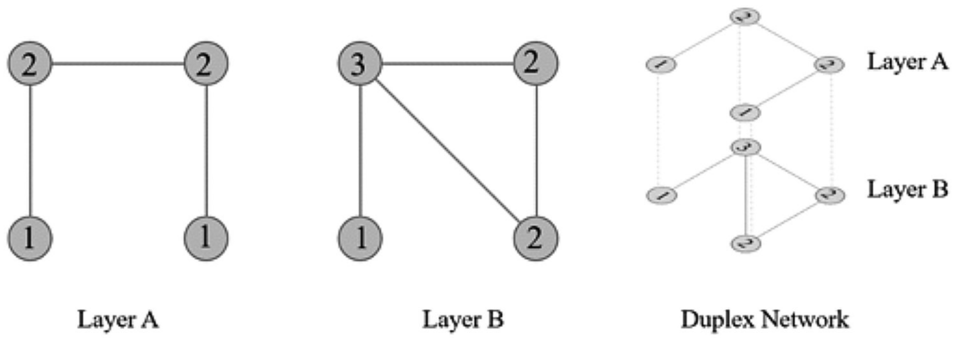

Figure 1. An Example of IE

interactive entropy in distinguishing the difference in degree-based distribution between different layers of duplex network.

3.1 Data

3.1.1 The information network of co-word in information science and library science (ICW)

ICW is a three-layer multiplex network and can be regarded as three duplex networks when matching two of the three layers together. The layers represent the co-word networks of the extracted from the field of information science and library science in 2012, 2015 and 2018. The nodes represent the keywords in the publications and the edges represent the number of times that two keywords occurring at the same time in one publication. By definition, the nodes of each layer in a multiplex network should be consistent. Therefore, fifty of the most frequently-used words in three years are selected from the network required in this research.

3.1.2 Star war social duplex network (SWS)

SWS is a duplex network drawn from the classic movie "Star war." Its nodes represent movie characters, and edges represent their interactive relationships. The interaction relations referred to by the edges in the two layers are interactions and mentions. The former represents the number of conversations between the two characters, and the latter represents the number of times the two characters are jointly mentioned.

3.1.3 Kapferer tailor shop duplex network (KTS)

KTS is a duplex network firstly observed by Kapferer (1972). The nodes represent different people, and the edges represent relationships between them. The network is unweighted, and an edge means that such a work- and assistance-related relationship exists between two people. This relationship was recorded in the Kapferer's two visits during which there was a strike. We call these two layers "before" and "after."

After the networks are simply introduced, their main network parameters and indicators are computed as shown in Table 1.

3.2 Results and Analysis of IE

Interactive entropy is characterized by the degree distribution between the two layers of a duplex network to measure the difference of the network. In order to see the degree distribution of

Table 1. Basic Parameters of Three Networks

Network	Type	Weight	Layer	Density	Diameter	Average degree	Average weighted degree
ICW	information network	weighted	2012	0.261	3	12.8	41.0
			2015	0.382	3	18.7	79.6
			2018	0.370	3	18.1	71.2
SWS	social network	weighted	interaction	0.063	6	7.0	26.0
			mention	0.129	5	14.5	97.9
KTS	social network	unweighted	before	0.213	4	8.1	16.2
			after	0.301	3	11.4	22.9

each duplex/multiplex network and its difference more clearly, Figure 2 is proposed, in which the horizontal axis of the graph is degree count, and the vertical axis is the probability of a certain degree count.

In the ICW network, the interactive entropy values are as follows: IE(12||15) = 3.69, IE(15||18) = 2.97, IE(12||18) = 3.04. It can be seen that in the two-layer network formed by the three-layer network in pairs, the degree distribution difference between the two-layer network formed in 12 and 15 years is the largest, and the difference between 15 and 18 years is the smallest. This is consistent with the difference in degree distribution between layers that can be directly observed in Figure 2.

As for the SWS network, the degree distribution curve of the interaction layer is concentrated on the left side of the figure, and that of the mentioned layer is more biased to the right side. In the SWS network, the value of interactive entropy is as follows: IE(mention||interaction) = 2.70.

In the KTS network, the degree distribution curve of the "before" period is more to the right, and that of the "after" period is more to the left. Intuitively, the difference between the two is huge, so the corresponding interactive entropy should also be relatively big. The value of interactive entropy is IE(before||after) = 4.43.

4. Discussion

Now we discuss more issues and extend a comparison on the differences between IE and existing entropy-type measures.

4.1 Entropy using other distribution

As mentioned earlier, when calculating interactive entropy, it is crucial to choose a reasonable distribution that can represent the overall properties of the network. Since the degree of a node is only related to the number of neighbors it has, without considering the weight of the edge, now the weighted network has become more and more commonly used when simulating various relationships in the real world. So we also consider improving it by introducing another degree distribution-*h*-degree distribution. *H*-degree is first proposed by Zhao et al. (2011) in the network, it is an efficacious method to measure the importance of nodes in weighted network. The *h*-degree of a node is the number d_h if this node has at least d_h links with other nodes and the strength of each of these links is greater than or equal to d_h. Correspondingly, the above formula could be modified as below.

The *h*-degree distribution of layer α (α = 1,2) is $p(k)_H^{[\alpha]}$

$$p(k)_H^{[\alpha]} = \frac{N_k^{[\alpha]}}{N} \tag{4}$$

where $N_k^{[\alpha]}$ is the number of nodes with *h*-degree *k* in the network.

Interactive entropy using *h*-degree distribution in duplex network is

$$IE_H(\alpha = 1 \,||\, \alpha = 2) = \sum_{k=1}^{N-1} p(k)_H^{[1]} \cdot \log \frac{p(k)_H^{[1]}}{p(k)_H^{[2]}} \tag{5}$$

It can be calculated that the *h*-degree distribution entropy of the three layers is 1.37, 1.90, 1.86, and the interactive entropy using *h*-degree distribution is IE_H(12||15) = 0.86, IE_H(15||18) = 0.16, IE_H(12||18) = 0.54. While the degree distribution entropy of the three layers is 3.92, 4.35 and 4.09, and the interactive entropy is *IE*(12||15) = 3.69, *IE*(15||18) = 2.97, *IE*(12||18) = 3.04. The trends are the same.

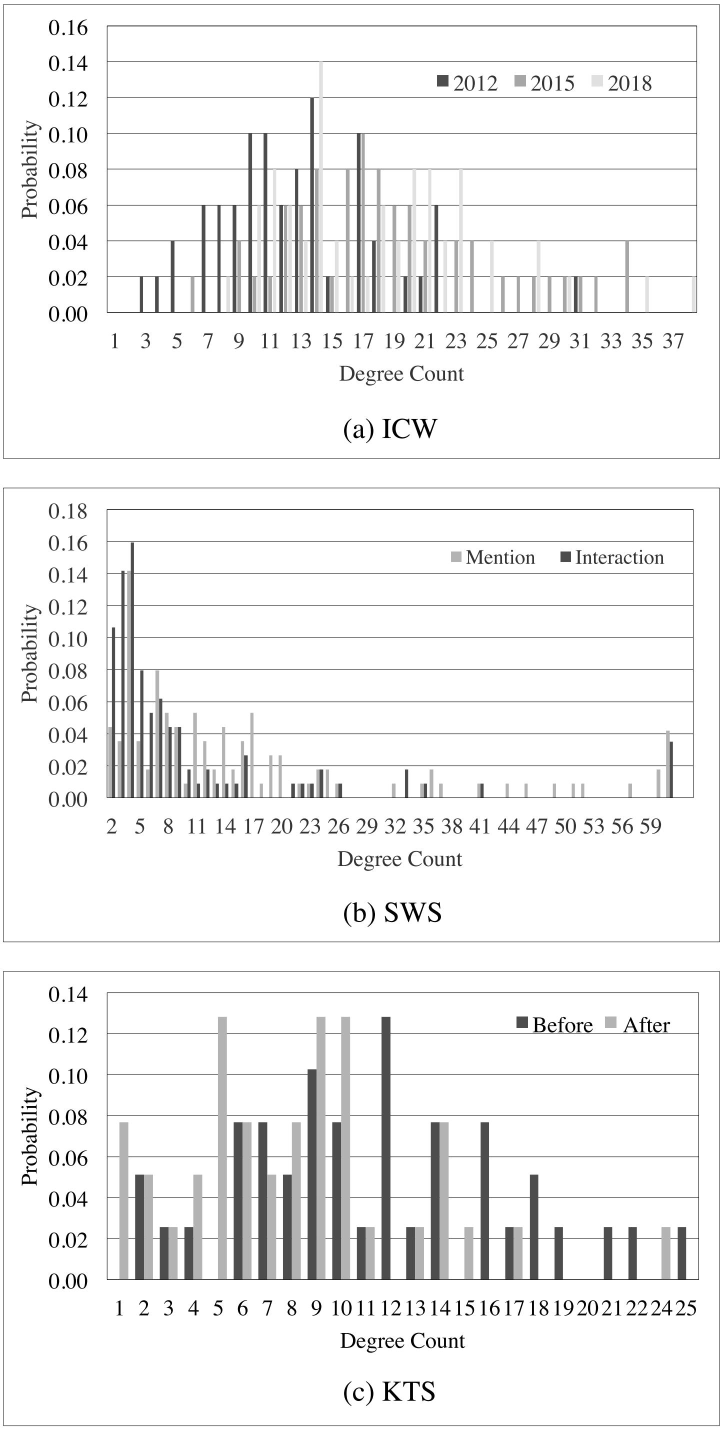

Figure 2. Degree Distribution Diagram of the Example Networks

The above demonstrated the use of other distributions in our definition of interactive entropy. We can also examine other entropy ideas and what distribution they use. Comparing with the entropy proposed by Battiston, Nicosia, and Latora (2014), the IE belongs to the interactive entropy, which means it pays more attention to overall interaction between layers, rather than different states of nodes or edges between layers. In their definition, the entropy of the multiplex degree is a measure of node properties. It uses the distribution of the degree of node i among the various layers. In this distribution, the denominator is the sum of the degree of the node in each layer, and the numerator is the degree of the node in a certain layer. Therefore, the direction of the result is whether the degree of a certain node is evenly distributed among the layers. If it is uniform, the entropy takes the maximum value; if not and the distribution is extremely uneven, for example, there is only one non-zero value in a certain layer, this entropy takes the minimum value. This idea is also a good way to use entropy to measure multiplex networks, but it is different from our implementation path. We measure the overall interaction relationship using the distribution that can better represent the entire network, such as the degree distribution and h-degree distribution, and other entropy may be more specific to specific nodes and edges using the degree of a certain node/edge among layers.

4.2 Layer similarity and interactive entropy in duplex network

For comparing layers in multiplex networks, especially in duplex networks, there are already some indicators, such as the layer similarity (Zhang & Ye, 2020). We can compare the interactive entropy and the layer similarity. In its definition, for each layer, there is an adjacency matrix $A^{[\alpha]}=a_{ij}$, where $a_{ij}^{[\alpha]}$ represents the weight of the edge connecting node i and node j.

The relationship that node i has with the other nodes can be represented by a vector

$$\boldsymbol{k}_i^{[\alpha]} = [\, a_{i1}^{[\alpha]},\ a_{i2}^{[\alpha]},\ a_{i3}^{[\alpha]},\ \ldots\ldots,\ a_{in}^{[\alpha]}] \tag{6}$$

The similarity of a certain node i could be:

$$NSim_i^{[1,2]} = \frac{\boldsymbol{k}_i^{[1]}\boldsymbol{k}_i^{[2]}}{|\boldsymbol{k}_i^{[1]}||\boldsymbol{k}_i^{[2]}|} \tag{7}$$

The layer similarity could be

$$LSim^{[1,2]} = \frac{\sum_{i=1}^{n} NSim_i^{[1,2]}}{n} \tag{8}$$

It can be concluded directly from the formula that the two indicators have a certain consistency. Yet, they have their own emphasis. The layer similarity starts from the difference between different layers of the same node in the duplex network, and the average of all node similarity is the layer similarity. Interactive entropy calculates the difference between the two-degree distributions of the two layers. It can be concluded that the former is from individual to the whole, while the latter is always focusing on the whole. In clarifying the difference between the duplex network, there are both differences and connections.

In the ICW network, the results of layer similarity are as follows: *LSim* [12,15] = 0.5311, *LSim* [15,18] = 0.6070, *LSim* [12,18] = 0.4570. And the interactive entropy is: IE(12||15) = 3.69, IE(15||18) = 2.97, IE(12||18) = 3.04. From the numerical values and its corresponding Figure 3, it can be seen that the trends of the two indicators are also slightly different. IE and LSim both show the highest similarity of the duplex network

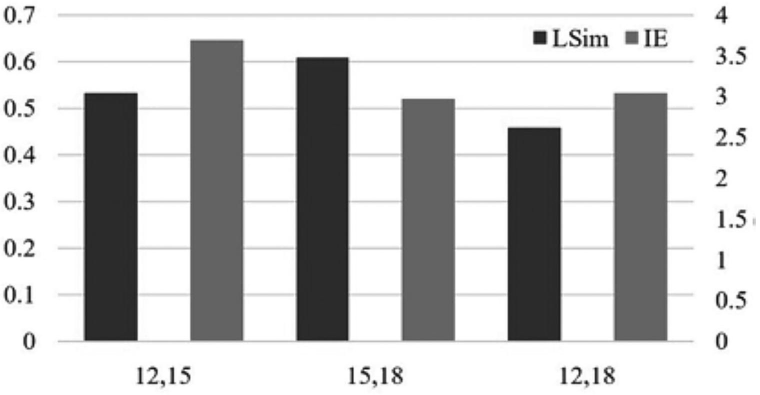

Figure 3. The Comparation between LSim and IE

of 15 and 18, but LSim shows the similarity of 12 and 15 is in the middle, while IE shows the similarity of 12 and 18 is in the middle. The conclusion drawn by the example is the same as the conclusion directly drawn by the formula.

There is one thing worth noting is that the values of these two indicators are opposite, which means the larger the interactive entropy, the greater the difference between the layers; the greater the similarity between the layers, the smaller the difference between the layers.

5. Conclusion

In this paper, based on the degree distribution of network, a new measure called interactive entropy (IE) of duplex networks is proposed. It can analyze the difference in degree distribution between layers, which is a form of interaction between layers. Subsequently, empirical studies reveal and verify the feasibility of this method. The types of test data set are rich and diverse, including not only weighted networks but also unweighted ones, and not only information networks but also social networks. Finally, some discussions are made to extend the degree distribution used in the original context to the *h*-degree distribution, and to compare the similarities and differences between the interactive entropy and the existing measures.

At present, IE is only limited to duplex networks. Their variations for multiplex networks could be explored and extended in the future.

Acknowledgement

The authors acknowledge National Natural Scientist Foundation of China Grant No. 71673131 for financial support.

Author contributions

R. J. Z. collected and processed the data and wrote the paper, and F. Y. Y initiated the idea and wrote the paper.

References

Albert, R., & Barabási, A. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47-97. doi: 10.1103/

RevModPhys.74.47

Anand, K., & Bianconi, G. (2009). Entropy measures for networks: Toward an information theory of complex topologies. *Physical Review E, 80*(4), 45-102. doi: 10.1103/PhysRevE.80.045102

Barabási, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509-512. doi: 10.1126/science.286.5439.509

Battiston, F., Nicosia, V., & Latora, V. (2014). Structural measures for multiplex networks. *Physical Review E, 89*(3), 032804. doi: 10.1103/PhysRevE.89.032804

Boccaletti, S., Bianconi, G., Criado, R., del Genio, C. I., Gómez-Gardeñes, J., Romance, M., ... Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports, 544*(1), 1-122. doi: 10.1016/j.physrep.2014.07.001

Cai, M., Cui, Y., & Stanley, H. E. (2017). Analysis and evaluation of the entropy indices of a static network structure. *Scientific Reports, 7*(9340). doi: 10.1038/s41598-017-09475-9

Cai, M., Du, H. F., & Feldman, W. M. (2014). A new network structure entropy based on maximum flow. *Acta Physica Sinica, 63*(6), 060504. doi: 10.7498/aps.63.060504

Cai, M., Du, H. F., & Ren, Y. K. (2011). A new network structure entropy based node difference and edge difference. *Acta Physica Sinica, 60*(11), 110513. doi: 10.7498/aps.60.110513

De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., ... Arenas, A. (2013). Mathematical formulation of multilayer networks. *Physical Review X, 3*(4), 041022. doi:10.1103/PhysRevX.3.041022

De Domenico, M., & Biamonte, J. (2016). Spectral entropies as information-theoretic tools for complex network comparison. *Physical Review X, 6*(4), 041062. doi: 10.1103/PhysRevX.6.041062

Kapferer, B. (1972). *Strategy and transaction in an African factory: African wokers and Indian management in a Zambian town*. Manchester, England: Manchester University Press.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*(1), 79-86. doi: 10.1214/aoms/1177729694

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(4), 623-656. doi: 10.1002/j.1538-7305.1948.tb00917.x

Shi, D., Chen, D., Long, H., Wang, C., & Pan, G. (2019). Quantifying complex network information based on communicability sequence entropy. *Scientia Sinica Physica, Mechanica & Astronomica, 49*(7), 070502. doi: 10.1360/SSPMA-2019-0029

Solé-Ribalta, A., de Domenico, M., Gómez, S., & Arenas, A. (2014). Centrality rankings in multiplex networks. In F. Menczer, J. Hendler, W. Duton, M. Strohmaier, E. T. Meyer, & C. Cattuto (Chairs). *Proceedings of the 2014 ACM Conference on Web Science* (pp. 149-155). New York, NY:

Association for Computing Machinery. doi: 10.1145/2615569.2615687

Solé, R. V., & Valverde, S. (2004). Information theory of complex networks: On evolution and architectural constraints. In E. Ben-Naim, H. Frauenfelder, & Z. Toroczkai (Eds.), *Complex networks* (pp. 189-207). Heidelberg, Germany: Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-540-44485-5_9

Strogatz, S. (2001). Exploring complex networks. *Nature, 410*(6825), 268-276. doi: 10.1038/35065725

Wang, B., Tang, H., Guo, C., & Xiu, Z. (2006). Entropy optimization of scale-free networks' robustness to random failures. *Physica A: Statistical Mechanics and its Applications, 363*(2), 591-596. doi: 10.1016/j.physa.2005.08.025

Wu, J., Tan, Y.-J., Deng, H.-Z., & Zhu, D.-Z. (2007). Heterogeneity of scale-free networks. *Systems Engineering-Theory & Practice, 27*(5), 101-105. doi: 10.1016/S1874-8651(08)60036-8

Zhang, Q., Li, M., & Deng, Y. (2018). Measure the structure similarity of nodes in complex networks based on relative entropy. *Physica A: Statistical Mechanics and its Applications, 491*, 749-763. doi: 10.1016/j.physa.2017.09.042

Zhang, R. J., & Ye, F. Y. (2020). Measuring similarity for clarifying layer difference in multiplex ad hoc duplex information networks. *Journal of Informetrics, 14*(1), 100987. doi: 10.1016/j.joi.2019.100987

Zhao, S. X., Rousseau, R., & Ye, F. Y. (2011). h-Degree as a basic measure in weighted networks. *Journal of Informetrics, 5*(4), 668-677. doi: 10.1016/j.joi.2011.06.005

Zhao, S. X., & Ye, F. Y. (2012). Exploring the directed h-degree in directed weighted networks. *Journal of Informetrics, 6*(4), 619-630. doi: 10.1016/j.joi.2012.06.007

(Received: 2020/8/27; Accepted: 2020/11/20)

圖書資訊學刊　第18卷第2期　（2020.12）　頁31-41
doi: 10.6182/jlis.202012_18(2).031

交互作用熵：量化雙層網絡交互作用的新測度

Quantifying Interactions between Layers in Duplex Networks Using Interactive Entropy

張家榕[1,2]　葉　鷹[1,2]
Ronda J. Zhang[1,2], Fred Y. Ye[1,2]

摘　要

多層網絡因其能夠很好地描述現實世界中的各種關係而廣受關注，而雙層網絡是多層網絡的特殊形式。本研究的目的在於引入一種新的基於熵的針對多層網絡，特別是雙層網絡的測度。這種新的測度指標基於資訊論中的關鍵概念——資訊熵，稱作交互作用熵，可應用於各種類型的多層網絡以量化其中雙層之間的差異。

關鍵字：多層網絡、雙層網絡、交互作用熵

1 中國南京大學–伊利諾大學國際信息學聯合實驗室（南京）
International Joint Informatics Laboratory (IJIL), Nanjing University – University of Illinois, Nanjing-Champaign, Nanjing, China
2 中國南京大學信息管理學院江蘇省數據工程和知識服務重點實驗室（南京）
Jiangsu Key Laboratory of Data Engineering and Knowledge Service, School of Information Management, Nanjing University, Nanjing, China
* 通訊作者Corresponding Author: 葉鷹Fred Y. Ye, E-mail: yye@nju.edu.cn

註：本中文摘要由作者提供。
以APA格式引用本文：Zhang, R. J., & Ye, F. Y. (2020). Quantifying interactions between layers in duplex networks using interactive entropy. *Journal of Library and Information Studies, 18*(2), 31-41. doi: 10.6182/jlis.202012_18(2).031
以Chicago格式引用本文：Ronda J. Zhang and Fred Y. Ye. “Quantifying interactions between layers in duplex networks using interactive entropy.” *Journal of Library and Information Studies* 18, no. 2 (2020): 31-41. doi: 10.6182/jlis.202012_18(2).031