A betweenness structure entropy of complex networks

Qi Zhang\textsuperscript{a}, Meizhu Li\textsuperscript{a}, Yong Deng\textsuperscript{a,c,b,*}

\textsuperscript{a}School of Computer and Information Science, Southwest University, Chongqing, 400715, China
\textsuperscript{b}School of Automation, Northwestern Polytechnical University, Xian, Shaanxi 710072, China
\textsuperscript{c}School of Engineering, Vanderbilt University, Nashville, TN, 37235, USA

Abstract

The structure entropy is an important index to illuminate the structure property of the complex network. Most of the existing structure entropies are based on the degree distribution of the complex network. But the structure entropy based on the degree can not illustrate the structure property of the weighted networks. In order to study the structure property of the weighted networks, a new structure entropy of the complex networks based on the betweenness is proposed in this paper. Comparing with the existing structure entropy, the proposed method is more reasonable to describe the structure property of the complex weighted networks.

Keywords: complex networks, structure entropy, betweenness, weighted network

\*Corresponding author: Yong Deng, School of Computer and Information Science, Southwest University, Chongqing, 400715, China.

Email address: ydeng@swu.edu.cn, prof.deng@hotmail.com (Yong Deng)
1. Introduction

The complex networks is a graph with non-trivial topological features, the feature that do not occur in simple networks but often occur in real networks. Many real networks are the complex networks, such as the social networks, information networks, technological networks and biological networks [1]. Recently, many researcher have been interested to explore the complex networks. In 1998, Watts and Strogatz proposed the principle of 'Small-world' for the complex networks on Nature [2]. Then the 'Scale-free networks' is proposed by some researchers [3]. Then the statistical theory is introduced in the complex networks [4, 5]. Those researches have revealed that the structure property is important to research the complex networks.

Many of the existing structure entropies are based on the degree distribution of the complex networks. But the degree of the complex networks is a local measure to some degree, ignoring the influence of the edge’s weighted in the structure property. As a results, the structure entropy based on the degree can not describe the structure property of those complex weighted networks, especially for those networks with a uniform degree distribution and different weighted of the edges. To describe the structure property of those complex weighted networks, we need to find a new method to represent the structure entropy.

Compared with the degree measure of the complex networks, the betweenness is a global measure of complex networks. It is defined based on the shortest path of the networks. It can be used to describe the structure
property of the complex networks from the global view.

In this paper, we proposed a new structure entropy of the complex networks which is based on the betweenness of the complex networks and the information theory. The results of our research have revealed that the structure entropy based on the betweenness is a useful method to describe the structure property of the complex networks.

The rest of this paper is organised as follows. Section 2 introduces some preliminaries of this work. In section 3, a new structure entropy of the complex networks based on the betweenness is proposed. The application of the proposed method is illustrated in section 4. Conclusion is given in Section 5.

2. Preliminaries

2.1. Betweenness

The betweenness is an important index which can be used to illuminate the importance of the nodes. It is defined based on the shortest path of the network [6].
The betweenness of the complex networks is defined as follows [6]:

$$bet(i) = \frac{\nu(i)}{\sum_{s \neq i \neq t} \sigma_{st}} (s \neq i \neq t)$$  \hspace{1cm} (1)

In the Eq. (1), the $\sigma_{st}$ is the number of the shortest path from vertex $s$ to vertex $t$, $\nu(i)$ is the number of the shortest path which have go to through the vertex $i$ [6].

2.2. Existing structure entropy

The structure entropy of the complex networks is based on the information entropy [7] and the statistic characteristics of the complex networks. It can be used to describe the structure property of the complex networks.

The information entropy is a conception of information theory which is proposed by Shannon [7]. Shannon defined the information as "the reduction
of entropy”, ”the reduction of uncertainty of a system”, and firstly proposed
the quantitative description method for information.

Suppose $X = \{x_1, x_2, x_3, \ldots, x_n\}$ is a discrete random variable, the
appearance probability of information source given by $X$ is denoted as $p_i =
p(x_i), i = 1, 2, \ldots, n$, and $\sum_{i=1}^n p_i = 1$. Then the information entropy is defined
as follows:

$$H = -k \sum_{i=1}^n p_i \log p_i$$

(2)

Where $k$ is equal to 1, $n$ is the number of the probabilities.

Many researchers have proposed the methods to calculate the structure
entropy of the complex networks, such as the structure entropy based on the
degree distribution [8], the structure entropy based on the automorphism
partition of the network [9] and the structure entropy based on the degree
dependence matrices [10]. Most of those structure entropies are based on the
degree of the nodes, defined as follows [8]:

$$H_{\text{deg}} = -k \sum_{j=1}^N p_j \log p_j$$

(3)

Where the $p_j$ is defined as follows:

$$p_j = \frac{\text{Degree}(j)}{\sum_{j=1}^N \text{Degree}(j)}$$

(4)

Where the $\text{Degree}(j)$ represent the $j$th vertex’s degree and $N$ is the total
number of the nodes in the network.
Figure 2: In this graph, we can partition the network based on the degree. The degree partition \( D = \{\{1, 9\}, \{3, 4, 5, 10, 11, 13, 14, 15\}, \{2, 7, 8, 12\}\} \) is a coarser automorphism partition of the networks. In the cell \( \{1, 9\} \) of degree partition, all vertices have degree 1. In this network \( P = D \), and \( V_1 = \{1, 9\}, V_2 = \{3, 4, 5, 10, 11, 13, 14, 15\}, V_3 = \{2, 7, 8, 12\} \). It is clearly that \( p_1 = 2/15, p_2 = 9/15 \), \( p_3 = 4/15 \). The structure entropy based on the degree partition of this network \( H_{\text{partition}} = 0.9276 \).

The structure entropy based on the automorphism partition of the network is defined as follows [9]:

\[
H_{\text{partition}} = - \sum_{p=1}^{\vert P \vert} p_p \log p_p 
\]  
(5)

Where \( P \) is the automorphism partition of the network, \( p_p \) is the probability that a vertex belongs to the cell \( V_i \) of the \( P \). Note that given a network’s automorphism partition \( P = \{V_1, V_2, V_3, \ldots, V_k\} \), the \( p_p \) is calculated as:

\[
p_p = \frac{|V_p|}{\sum_{p=1}^{k} |V_p|} = \frac{|V_p|}{N}
\]  
(6)

Where the \( k \) is the cell’s mounts of the \( P \). The Fig. 2 shows an example about how to calculate the structure entropy based on the automorphism partition of the network.
2.3. The shortcoming of degree-based structure entropy

A weighted network is shown in Fig. 3.

![Figure 3: The network A](image)

The details of the network A is shown in Table 1.

| Node label | degree | betweenness | The number of path across the vertex |
|------------|--------|-------------|-------------------------------------|
| vertex 1   | 3      | 0.035       | 20                                  |
| vertex 2   | 3      | 0.035       | 20                                  |
| vertex 3   | 3      | 0.3385      | 98                                  |
| vertex 4   | 3      | 0.035       | 20                                  |
| vertex 5   | 3      | 0.2101      | 65                                  |
| vertex 6   | 3      | 0.1518      | 50                                  |
| vertex 7   | 3      | 0.035       | 20                                  |
| vertex 8   | 3      | 0.1206      | 42                                  |
| vertex 9   | 3      | 0.0195      | 16                                  |
| vertex 10  | 3      | 0.0195      | 16                                  |
The network A has 10 nodes and 15 edges. Each node’s degree is 3, which means that change the value of the edge’s weighted, the degree-based structure entropy of the network A is invariable.

3. Proposed structure entropy

To address the issue in Fig 3, we proposed a new structure entropy based on the betweenness of the complex networks. It is defined as follows:

\[
H_{\text{bet}} = -\sum_{i=1}^{n} p_i \log p_i
\]  
(7)

Where \( p_i \) is defined as follows:

\[
p_i = \frac{v(i)}{\sum_{i=1}^{n} v(i)}
\]  
(8)

Where \( v(i) \) is the betweenness which is defined in section 2.1.

To show the necessity of the proposed method, we have calculated the information loss of the network A with the existing structure entropy and the proposed structure entropy. The results are shown in Table 2.
Table 2: The information loss test for the network A

| Loss Vertex | $H_{bet}$ | $H_{bet}^{loss}$ | $H_{deg}$ | $H_{deg}^{loss}$ | $H_{partition}$ | $H_{partition}^{loss}$ |
|-------------|----------|-----------------|-----------|-----------------|----------------|------------------------|
| The network A | 1.8585   | 2.3026          | 0         |                 |                |                        |
| Vertex 1    | 1.9641   | -0.1055         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 2    | 1.9589   | -0.1004         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 3    | 2.0481   | -0.1895         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 4    | 1.892    | -0.0335         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 5    | 2.1407   | -0.2821         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 6    | 1.7638   | 0.0948          | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 7    | 1.7531   | 0.1055          | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 8    | 1.8165   | 0.0420          | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 9    | 1.9774   | -0.1189         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |
| Vertex 10   | 1.9774   | -0.1189         | 2.1808    | 0.1218          | 0.6365         | -0.6365                |

Where the $H_{x}^{loss}$ represents the information loss of the network A which is calculated with the existing structure entropy and the proposed structure entropy.

The results show that the proposed structure entropy can illuminate the difference of the information loss of the nodes in the network A.

In order to prove the reasonability of the proposed method, the information loss of the Zachary’s Karate Club network [11] is calculated. The results are shown in Table 3, Table 4 and Fig. 5.
Figure 4: The Zachary’s Karate Club network

Table 3: The details of the Zachary’s Karate Club network

| Network | Nodes | edges | $C$  | $L$     |
|---------|-------|-------|------|---------|
| Karate  | 34    | 78    | 0.4726 | 2.8966 |
| Loss | Vertex | betweenness | degree | $H_{bet}$ | $H_{bet}^{loss}$ | $H_{deg}$ | $H_{deg}^{loss}$ |
|------|--------|-------------|--------|----------|-----------------|----------|-----------------|
| Networls | 2.8857 | 0 | 3.2609 | 0 | |
| vertex1 | 0.1513 | 16 | 3.1404 | -0.2547 | 3.1970 | 0.0639 |
| vertex2 | 0.0241 | 9 | 2.7765 | 0.1092 | 3.2031 | 0.0577 |
| vertex3 | 0.1339 | 10 | 2.6654 | 0.2202 | 3.2198 | 0.0411 |
| vertex4 | 0.0092 | 6 | 2.8019 | 0.0838 | 3.2194 | 0.0415 |
| vertex5 | 0.0092 | 3 | 2.8246 | 0.0610 | 3.2247 | 0.0361 |
| vertex6 | 0.0266 | 4 | 2.8243 | 0.0613 | 3.2117 | 0.0491 |
| vertex7 | 0.0098 | 4 | 2.8243 | 0.0613 | 3.2117 | 0.0491 |
| vertex8 | 0.0092 | 4 | 2.8001 | 0.0856 | 3.2357 | 0.0252 |
| vertex9 | 0.0174 | 5 | 2.7861 | 0.0996 | 3.2401 | 0.0207 |
| vertex10 | 0.1157 | 2 | 2.7884 | 0.0972 | 3.2433 | 0.0175 |
| vertex11 | 0.0092 | 3 | 2.8850 | 0.0007 | 3.2247 | 0.0361 |
| vertex12 | 0.0092 | 1 | 2.8867 | -0.0010 | 3.2490 | 0.0118 |
| vertex13 | 0.0092 | 2 | 2.8861 | -0.0005 | 3.2393 | 0.0215 |
| vertex14 | 0.0092 | 5 | 2.8755 | 0.0101 | 3.2411 | 0.0197 |
| vertex15 | 0.0120 | 2 | 2.8769 | 0.0087 | 3.2446 | 0.0163 |
| vertex16 | 0.0098 | 2 | 2.8633 | 0.0224 | 3.2446 | 0.0163 |
| vertex17 | 0.0098 | 2 | 2.8588 | 0.0269 | 3.2265 | 0.0343 |
| vertex18 | 0.0126 | 2 | 2.8843 | 0.0014 | 3.2422 | 0.0187 |
| vertex19 | 0.0104 | 2 | 2.8624 | 0.0232 | 3.2446 | 0.0163 |
| vertex20 | 0.0868 | 3 | 2.9885 | -0.1029 | 3.2433 | 0.0176 |
| vertex21 | 0.0092 | 2 | 2.8815 | 0.0042 | 3.2446 | 0.0163 |
| vertex22 | 0.0092 | 2 | 2.8759 | 0.0097 | 3.2422 | 0.0187 |
| vertex23 | 0.0092 | 2 | 2.8626 | 0.0231 | 3.2446 | 0.0163 |
| vertex24 | 0.0098 | 5 | 2.8577 | 0.0280 | 3.2207 | 0.0401 |
| vertex25 | 0.0207 | 3 | 2.9116 | -0.0259 | 3.2178 | 0.0431 |
| vertex26 | 0.0031 | 3 | 2.9357 | -0.0501 | 3.2195 | 0.0414 |
| vertex27 | 0.0092 | 2 | 2.7801 | 0.1055 | 3.2367 | 0.0241 |
| vertex28 | 0.0126 | 4 | 2.7510 | 0.1347 | 3.2264 | 0.0344 |
| vertex29 | 0.0174 | 3 | 2.7973 | 0.0884 | 3.2371 | 0.0238 |
| vertex30 | 0.0092 | 4 | 2.7687 | 0.1170 | 3.2242 | 0.0367 |
| vertex31 | 0.0106 | 4 | 2.7605 | 0.1251 | 3.2361 | 0.0248 |
| vertex32 | 0.0311 | 6 | 2.8206 | 0.0651 | 3.2225 | 0.0384 |
| vertex33 | 0.0308 | 12 | 2.6796 | 0.2061 | 3.1929 | 0.0680 |
| vertex34 | 0.1325 | 17 | 2.6244 | 0.2613 | 3.1893 | 0.0716 |

The results show that the vertex 33, vertex 34, vertex 1 and vertex 3 are important to the network which is the same as the degree-based structure entropy.
4. Application

In this section, the proposed method is used to calculate the structure entropy of the other real networks, namely, the US-airport network \[12\], Email networks \[12\], the Germany highway networks \[13\], the US power grid and the protein-protein interaction network in budding yeast \[12\]. The results are shown in Table \[\text{12}\].
Table 5: The structure entropy of the real networks

| Network          | Nodes | Edges | \( H_{\text{deg}} \) | \( H_{\text{bet}} \) | \( H_{\text{partition}} \) |
|------------------|-------|-------|-----------------------|-----------------------|-----------------------------|
| US Airport       | 500   | 5962  | 5.025                 | 4.7338                | 3.1263                      |
| Email            | 1133  | 10902 | 6.631                 | 5.5021                | 3.1780                      |
| Yeast            | 2375  | 23386 | 7.0539                | 6.0931                | 3.0345                      |
| US power grid    | 4941  | 13188 | 8.3208                | 5.7191                | 1.7018                      |
| Germany highway  | 1168  | 2486  | 6.9947                | 5.6383                | 0.6909                      |

The \( H_{\text{deg}} \) represents the structure entropy which is based on the degree. The \( H_{\text{partition}} \) represents the structure entropy which is based on the degree partition. The \( H_{\text{deg}} \) represents the structure entropy which is proposed in the paper.

![Figure 6: The US airport network](image-url)
Figure 7: The Email network

Figure 8: The protein-protein interaction network in budding yeast
Figure 9: The US power grid

Figure 10: The Germany highway network

The calculate process of the degree-based structure entropy and the proposed structure entropy are shown in Fig. 6, Fig. 7, Fig. 8, Fig. 9 and Fig.
5. Conclusion

The results of our research reveal that compared with the existing structure entropy the proposed structure entropy is more effective to describe the structure property of the weighted networks. It is a new method to explore the structure property of the complex networks.

Acknowledgments

The work is partially supported by National Natural Science Foundation of China (Grant No. 61174022), Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20131102130002), R&D Program of China (2012BAH07B01), National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013801), the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.BUAA-VR-14KF-02).

References

[1] M. E. Newman, The structure and function of complex networks, SIAM review 45 (2) (2003) 167–256.

[2] D. J. Watts, S. H. Strogatz, Collective dynamics of ’small-world’ networks, nature 393 (6684) (1998) 440–442.
[3] A.-L. Barabási, R. Albert, H. Jeong, Scale-free characteristics of random networks: the topology of the world-wide web, Physica A: Statistical Mechanics and its Applications 281 (1) (2000) 69–77.

[4] J. Berg, M. Lässig, Correlated random networks, Physical review letters 89 (22) (2002) 228701.

[5] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Reviews of modern physics 74 (1) (2002) 47.

[6] M. Barthelemy, Betweenness centrality in large complex networks, The European Physical Journal B-Condensed Matter and Complex Systems 38 (2) (2004) 163–168.

[7] C. E. Shannon, A mathematical theory of communication, ACM SIGMOBILE Mobile Computing and Communications Review 5 (1) (2001) 3–55.

[8] B. Wang, H. Tang, C. Guo, Z. Xiu, Entropy optimization of scale-free networks robustness to random failures, Physica A: Statistical Mechanics and its Applications 363 (2) (2006) 591–596.

[9] Y.-H. Xiao, W.-T. Wu, H. Wang, M. Xiong, W. Wang, Symmetry-based structure entropy of complex networks, Physica A: Statistical Mechanics and its Applications 387 (11) (2008) 2611–2619.

[10] X.-L. Xu, X.-F. Hu, X.-Y. He, Degree dependence entropy descriptor for complex networks, Advances in Manufacturing 1 (3) (2013) 284–287.
[11] Uci network data repository, http://networkdata.ics.uci.edu/data.php?id=105 (2014).

[12] Pajek datasets, http://vlado.fmf.uni-lj.si/pub/networks/data/ (2014).

[13] Tore Opsahl, http://toreopsahl.com/datasets/ (2014).