IEICE TRANS. INF. & SYST., VOL.E94–D, NO.11 NOVEMBER 2011

LETTER

Strength-Strength and Strength-Degree Correlation Measures for Directed Weighted Complex Network Analysis

Shi-Ze GUO†, **Zhe-Ming LU**††a), **Zhe CHEN**†, ***Nonmembers*, *and* Hao LUO**††, ***Member***

SUMMARY This Letter defines thirteen useful correlation measures for directed weighted complex network analysis. First, in-strength and out-strength are defined for each node in the directed weighted network. Then, one node-based strength-strength correlation measure and four arc-based strength-strength correlation measures are defined. In addition, considering that each node is associated with in-degree, out-degree, in-strength and out-strength, four node-based strength-degree correlation measures and four arc-based strength-degree correlation measures are defined. Finally, we use these measures to analyze the world trade network and the food web. The results demonstrate the effectiveness of the proposed measures for directed weighted networks.

key words: *complex network, directed network, weighted network, directed weighted network, strength-strength correlation, strength-degree correlation*

1. Introduction

Complex network science is a young and active research area inspired largely by the empirical study of real-world networks that exhibit substantial non-trivial topological features. Scale-free networks [1] and small-world networks [2] are two kinds of typical complex networks. The former possess power-law degree distributions, while the latter have short path lengths and high clustering coefficients [3]. The study of complex networks has attracted researchers from various areas. A network $\boldsymbol{G}$ is composed of a set of nodes $\boldsymbol{V} = \{v_1, v_2, \ldots, v_N\}$ joined by edges $\boldsymbol{E} = \{e_{ij}\}$, where e_{ij} denotes the edge from node v_i to node v_j. In many real-world networks, edges are often associated with weights $\boldsymbol{W} = \{w_{ij}\}$ that differentiate them in terms of their strength, intensity or capacity [4]. We call these networks weighted networks [2], [5]–[7]. On the other hand, many real-life networks are directed networks composed of directed edges (i.e., arcs) [8], [9]. If a network is both weighted and directed, we call it a directed weighted network [6], [9], [10]. For undirected unweighted networks, many measures have been used to characterize their features. Here, we focus on the generalization of the degree-degree correlation for directed and weighted networks. Inspired by existing results for undirected networks reported in [11], Zhou *et al.* [12] have introduced arc-based degree-degree correlation measures for directed networks. Researchers have found a very interesting empirical phenomenon in weighted networks, i.e., the power-law strength-degree correlation [13]. However, there are relatively few research reports on strength-strength or strength-degree correlation measures for directed weighted networks. Recently, only one arc-based strength-degree correlation measure has been introduced [14] for the weighted directed network on worldwide tourism. The aim of this letter is to define in detail all possible strength-degree and strength-strength correlations for directed weighted complex networks, and then apply these measures to analyze the world trade network and the food web.

Manuscript received January 14, 2011.
Manuscript revised May 21, 2011.

†The authors are with North Electronic Systems Engineering Corporation, Beijing 100083, China.
††The authors are with School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China.
a) E-mail: zheminglu@zju.edu.cn (Corresponding Author)

DOI: 10.1587/transinf.E94.D.2284

2. Strength-Strength Correlation Measures for Directed Weighted Networks

2.1 In-Strength and Out-Strength

For a directed weighted simple graph $\boldsymbol{G}$ on N nodes, its adjacency matrix $\boldsymbol{A}$ is an $N \times N$ matrix where the non-diagonal entry a_{ij} satisfies: if there exists an arc from the head node v_i to the tail node v_j, then $a_{ij} = 1$; otherwise, $a_{ij} = 0$. For the diagonal entry, we have $a_{ii} = 0$. The number of head nodes directly adjacent to Node v_i is called the in-degree of v_i, while the number of tail nodes directly adjacent to it is its out-degree. The in-degree is denoted as k_i^{in} and the out-degree as k_i^{out}. Furthermore, the sum of arc weights inbound to Node v_i is called the in-strength of v_i and the sum of arc weights outbound from it is its out-strength. The in-strength is denoted as s_i^{in} and the out-strength as s_i^{out}. Based on above definitions, we have

$$\begin{aligned} k_i^{\text{in}} &= \sum_{j=1}^{N} a_{ji}, \qquad k_i^{\text{out}} = \sum_{j=1}^{N} a_{ij} \\ s_i^{\text{in}} &= \sum_{j=1}^{N} a_{ji} w_{ji}, \quad s_i^{\text{out}} = \sum_{j=1}^{N} a_{ij} w_{ij} \end{aligned} \tag{1}$$

2.2 Node-Based Strength-Strength Correlation Measures

There are two types of definition for strength-strength correlations, one is node-based and the other is arc-based. This subsection introduces the first type of definition. This definition focuses on the correlation between the in-strength and out-strength for each single node. We denote this correlation as a function $s_{\text{nn}}(s_{\text{in}})$, i.e., the average out-strength over all nodes with the same in-strength s_{in}. Thus, we can define

$s_{nn}(s_{in})$ as

$$s_{nn}(s_{in}) = \frac{\sum_{i:\, s_i^{in}=s_{in}} s_i^{out}}{N \cdot P_{in}(s_{in})} \tag{2}$$

where $P_{in}(s_{in})$ is the nodal in-strength distribution and N is the number of nodes. If the curve $s_{nn}(s_{in})$ is with a positive slope, the nodal in-strength and out-strength are positively correlated; If the curve $s_{nn}(s_{in})$ is with a negative slope, the nodal in-strength and out-strength are negatively correlated; If the slope equals 0, the nodal in-strength and out-strength are not correlated.

2.3 Arc-Based Strength-Strength Correlation Measures

This subsection focuses on the correlations between the in(out)-strength of the head node and the in(out)-strength of the tail node for each arc. Obviously, there are four combinations, i.e., in-in, out-in, in-out and out-out correlations, which can be denoted as functions $s_{\text{in-in}}(s_{in})$, $s_{\text{out-in}}(s_{in})$, $s_{\text{in-out}}(s_{out})$ and $s_{\text{out-out}}(s_{out})$, respectively. In this letter, we simply define them as

$$s_{\text{in(out)-in}}(s_{in}) = \frac{\sum_{i:\, s_i^{in}=s_{in}} \left[\frac{1}{s_i^{in}} \sum_{j=1}^{N} a_{ji} w_{ji} s_j^{\text{in(out)}} \right]}{N \cdot P_{in}(s_{in})} \tag{3}$$

$$s_{\text{in(out)-out}}(s_{out}) = \frac{\sum_{i:\, s_i^{out}=s_{out}} \left[\frac{1}{s_i^{out}} \sum_{j=1}^{N} a_{ij} w_{ij} s_j^{\text{in(out)}} \right]}{N \cdot P_{out}(s_{out})} \tag{4}$$

where $P_{in}(s_{in})$ and $P_{out}(s_{out})$ are the nodal in-strength and out-strength distributions respectively, and N is the number of nodes.

3. Strength-Degree Correlation Measures for Weighted Directed Networks

3.1 Node-Based Strength-Degree Correlation Measures

Similarly, there are two types of definition for strength-degree correlations, one is node-based and the other is arc-based. This subsection introduces the first type of definition, i.e., the correlation between the in(out)-strength and in(out)-degree for each single node. Obviously, there are four combinations, i.e., in-in, out-in, in-out and out-out correlations, which can be denoted as functions $s_{\text{in-in}}(k_{in})$, $s_{\text{out-in}}(k_{in})$, $s_{\text{in-out}}(k_{out})$ and $s_{\text{out-out}}(k_{out})$, respectively. Assume $c, d \in \{\text{in}, \text{out}\}$, we can define them as follows

$$s_{c-d}(k_d) = \frac{\sum_{i:\, k_i^{d}=k_d} s_i^{c}}{N \cdot P_d(k_d)} \tag{5}$$

where $P_{in}(k_{in})$ and $P_{out}(k_{out})$ are the nodal in-degree and out-degree distributions respectively.

3.2 Arc-Based Strength-Degree Correlation Measures

In fact, the arc-based strength-degree correlations mean the weighted degree-degree correlations between the in(out)-degree of the head node and the in(out)-degree of the tail node for each arc. Obviously, there are four combinations, i.e., in-in, out-in, in-out and out-out correlations, which can be denoted as functions $k_{w\text{\_in-in}}(k_{in})$, $k_{w\text{\_out-in}}(k_{in})$, $k_{w\text{\_in-out}}(k_{out})$ and $k_{w\text{\_out-out}}(k_{out})$, respectively. We can define them as

$$k_{w\text{\_in(out)-in}}(k_{in}) = \frac{\sum_{i:\, k_i^{in}=k_{in}} \left[\frac{1}{s_i^{in}} \sum_{j=1}^{N} a_{ji} w_{ji} k_j^{\text{in(out)}} \right]}{N \cdot P_{in}(k_{in})} \tag{6}$$

$$k_{w\text{\_in(out)-out}}(k_{out}) = \frac{\sum_{i:\, k_i^{out}=k_{out}} \left[\frac{1}{s_i^{out}} \sum_{j=1}^{N} a_{ij} w_{ij} k_j^{\text{in(out)}} \right]}{N \cdot P_{out}(k_{out})} \tag{7}$$

4. Analysis of the World Trade Network and the Food Web

To understand our five strength-strength correlation measures and eight strength-degree correlation measures, we first apply these thirteen measures to analyze a real weighted directed network, i.e., the world trade network (WTN), based on the same dataset used by Subramanian and Wei [15]. In the world trade network, we can view countries as nodes of the network and the existence of trade flows as arcs in a simple directed graph and the logarithmic trade values as weights of arcs. For each node, the sum of logarithmic import trade values can be counted as the in-strength, while the out-strength would be the sum of logarithmic export trade values. Serrano and Boguna [10] have shown that the world trade network in the year 2000 displays the typical properties of a complex network, thus we extract the data in the year 2000 from the dataset [15] as our test database. Table 1 lists some indices of this database, we can see that the world trade network has a relatively high density, which indicates that trade ties between countries have become tighter and tighter. The high clustering coefficient indicates that, if we randomly select three countries, the possibility that each country trades with the other two countries is very high.

Table 1 Some indices of the world trade network for the year 2000.

Item	Value
No. Nodes (Countries), N	157
No. Arcs (Trade Flows), M	11938
Density, $M/[N(N\text{-}1)]$	0.487
Average in-degree $<k_{in}>$	76.04
Average out-degree $<k_{out}>$	76.04
Average in-strength $<s_{in}>$	798.11
Average out-strength $<s_{out}>$	798.11
Clustering coefficient C	0.801

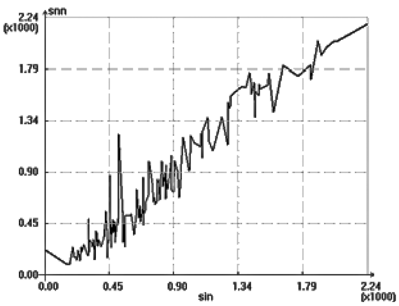

Fig. 1 The correlation between in-strengths and out-strengths in WTN.

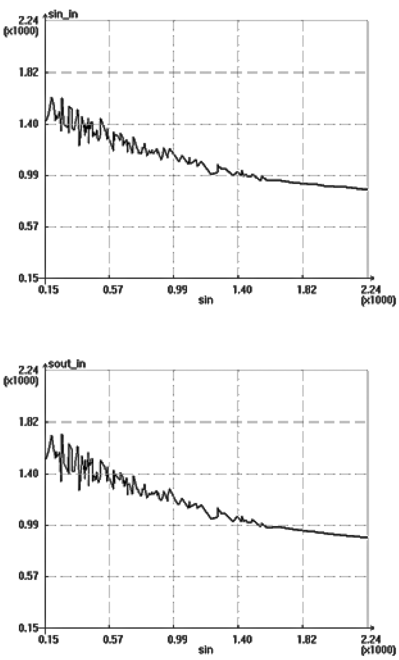

Fig. 2 Arc-based strength-strength correlations in WTN.

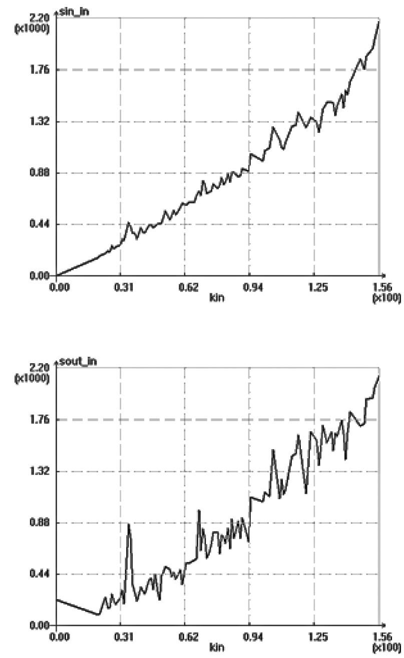

Fig. 3 Node-based strength-degree correlations in WTN.

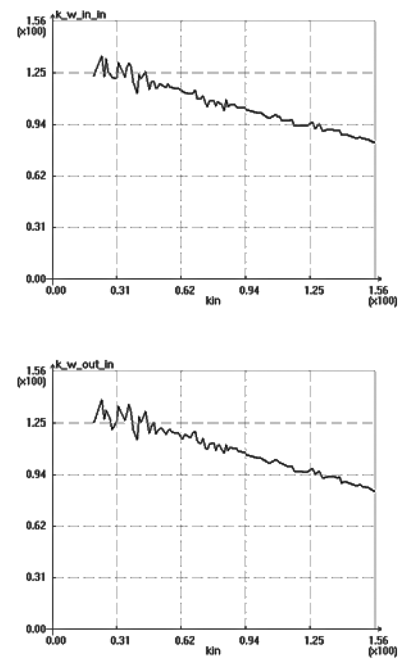

Fig. 4 Arc-based strength-degree correlations in WTN.

Figure 1 shows the node-based strength-strength correlation $s_{\mathrm{nn}}(s_{\mathrm{in}})$ for WTN, we can see that the $s_{\mathrm{nn}}(s_{\mathrm{in}})$ curve is with a positive slope around 1.0, which means that a country has more import trade values will also have more export trade values. For example, the country USA in WTN has the largest import trade value 2237.98, while its export trade value 2178.58 is also the largest one. Figure 2 shows two examples of four arc-based strength-strength correlations, $s_{\text{in-in}}(s_{\mathrm{in}})$ and $s_{\text{out-in}}(s_{\mathrm{in}})$ (similar results can be obtained for $s_{\text{in-out}}(s_{\mathrm{out}})$ and $s_{\text{out-out}}(s_{\mathrm{out}})$). From Fig. 2, we can see that the arc-based strength-strength correlation curves are with the negative slope between -1 and 0, which means that the country owning more trade values tends to trade with countries in possession of a bit less trade values. For example, the country USA in WTN has the largest import trade value 2237.98, while the average import value of its partners is the smallest value 869.54. Figure 3 shows two examples of four node-based strength-degree correlations, $s_{\text{in-in}}(k_{\mathrm{in}})$ and $s_{\text{out-in}}(k_{\mathrm{in}})$ (similar results can be obtained for $s_{\text{in-out}}(k_{\mathrm{out}})$ and $s_{\text{out-out}}(k_{\mathrm{out}})$). From Fig. 3, we can see that all curves are with a positive slope, which means that a country has more trade flows will also have more trade values. For example, the country USA in WTN has the largest number of importing partners 156, while its average import value 2171.67 is also the largest one. Figure 4 shows two examples of four arc-based strength-degree correlations, $k_{w\text{\_in-in}}(k_{\mathrm{in}})$ and $k_{w\text{\_out-in}}(k_{\mathrm{in}})$ (similar results can be obtained for $k_{w\text{\_in-out}}(k_{\mathrm{out}})$ and $k_{w\text{\_out-out}}(k_{\mathrm{out}})$). From Fig. 4, we can see that the arc-based strength-degree correlation curves are with the neg-

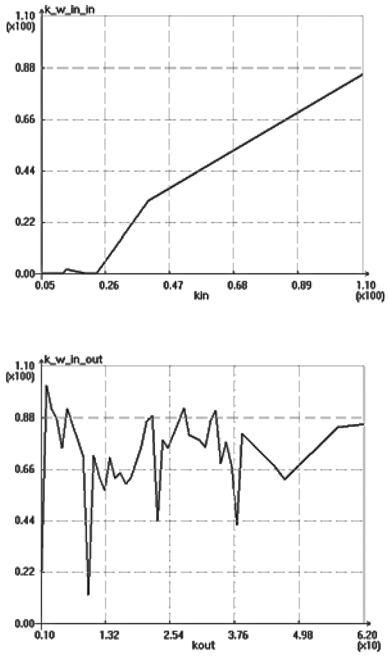

Fig. 5 Arc-based strength-degree correlations in FW.

ative slope between -1 and 0, which means that the country owning more trade flows tends to trade with countries in possession of a bit less weighted trade flows. For example, the country USA in WTN has the largest number of importing partners 156, while the average number of weighted trade flows of its partners is only 82. We also apply our proposed correlations to characterize another directed weighted network, i.e., the food web (FW) "Florida.paj" with 128 vertices and 2106 arcs [16]. Based on the simulation results, we find that there are two positive correlations in the food web, i.e., its node-based strength-strength correlation and its arc-based in strength-in degree correlation as shown in Fig. 5 (above). However, the remainder eleven correlations are all uncertain, e.g., the arc-based in strength-out degree correlation as shown in Fig. 5 (below).

5. Conclusions

This Letter proposes thirteen effective measures for directed complex networks. They are derived from nodal in-strengths, out-strengths, in-degrees and out-degrees. They are simple and easy to calculate. The application in analyzing the world trade network and the food web has shown their feasibility and usability. However, we should note that some of proposed correlations may not be applicable to networks with bipolar weights.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under grant 61070208.

References

[1] A.L. Barabási and R. Albert, "Emergence of scaling in random networks," Science, vol.286, no.5439, pp.509–512, Oct. 1999.

[2] D.J. Watts and S.H. Strogatz, "Collective dynamics of 'small-world' networks," Nature, vol.393, no.6684, pp.440–442, June 1998.

[3] M.E.J. Newman, "The structure and function of complex networks," SIAM Review, vol.45, no.2, pp.167–256, April 2003.

[4] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks," Proc. National Academy of Sciences, vol.101, no.11, pp.3747–3752, Nov. 2004.

[5] M. Granovetter, "The strength of weak ties," American Journal of Sociology, vol.78, no.6, pp.1360–1380, June 1973.

[6] J.J. Luczkowich, S.P. Borgatti, J.C. Johnson, and M.G. Everett, "Defining and measuring trophic role similarity in food webs using regular equivalence," Journal of Theoretical Biology, vol.220, pp.303–321, 2003.

[7] D.J. Watts and S.H. Strogatz, "Collective dynamics of 'small-world' networks," Nature, vol.393, no.6684, pp.440–442, June 1998.

[8] B. Tadic, "Dynamics of directed graphs: The world-wide Web," Physica A: Statistical Mechanics and its Applications, vol.293, pp.273–284, 2001.

[9] M. Girvan and M.E.J. Newman, "Community structure in social and biological networks," Proc. National Academy of Sciences of the United States of America, vol.99, no.12, pp.7821–7826, 2002.

[10] M.Á. Serrano and M. Boguñá, "Topology of the world trade web," Physical Review E, vol.68, no.1, pp.015101.1–015101.4, 2003.

[11] A. Pusch, S. Weber, and M. Porto, "Generating random networks with given degree-degree correlations and degree-dependent clustering," Physical Review E, vol.77, pp.017101.1–017101.4, 2008.

[12] J. Zhou, G. Xiao, L. Wong, X. Fuc, S. Mad, and T.H. Cheng, "Generation of arbitrary two-point correlated directed networks with given modularity," Physics Letters A, vol.374, pp.3129–3135, 2010.

[13] Q. Ou, Y.-D. Jin, T. Zhou, B.-H. Wang, and B.-Q. Yin, "Power-law strength-degree correlation from resource-allocation dynamics on weighted networks," Phys. Rev. E, vol.75, no.2, pp.021102.1–021102.5, 2007.

[14] J.I.L. Miguéns and J.F.F. Mendes, "Weighted and directed network on traveling patterns," BIOWIRE 2007, (eds. P. Liò et al.), LNCS 5151, pp.145–154, 2008.

[15] A. Subramanian and S.J. Wei, The WTO promotes trade, strongly but unevenly, http://www.nber.org/˜wei/data.html, 2005.

[16] Food Webs. Pajek datasets. http://vlado.fmf.uni-lj.si/pub/networks/data/bio/foodweb/foodweb.htm.