A Letters Journal Exploring the Frontiers of Physics

December 2014

EPL, **108** (2014) 58001
www.epljournal.org
doi: 10.1209/0295-5075/108/58001

# Dynamic motifs in socio-economic networks

Xin Zhang[1], Shuai Shao[2], H. Eugene Stanley[2] and Shlomo Havlin[3]

[1] *College of Transport and Communication, Shanghai Maritime University - Shanghai, 201306, China*
[2] *Center for Polymer Studies and Department of Physics, Boston University - Boston, MA 02215, USA*
[3] *Department of Physics, Bar-Ilan University - 52900 Ramat-Gan, Israel*

received 7 August 2014; accepted in final form 7 November 2014
published online 2 December 2014

PACS 89.75.-k – Complex systems
PACS 89.65.Ef – Social organizations; anthropology
PACS 64.60.aq – Networks

**Abstract** – Socio-economic networks are of central importance in economic life. We develop a method of identifying and studying motifs in socio-economic networks by focusing on "dynamic motifs," *i.e.*, evolutionary connection patterns that, because of "node acquaintances" in the network, occur much more frequently than random patterns. We examine two evolving bi-partite networks: i) the world-wide commercial ship chartering market and ii) the ship build-to-order market. We find similar dynamic motifs in both bipartite networks, even though they describe different economic activities. We also find that "influence" and "persistence" are strong factors in the interaction behavior of organizations. When two companies are doing business with the same customer, it is highly probable that another customer who currently only has business relationship with one of these two companies, will become customer of the second in the future. This is the effect of influence. Persistence means that companies with close business ties to customers tend to maintain their relationships over a long period of time.



**Introduction.** – Many complex physical, biological, and social systems can be modeled and better understood as complex networks [1–11]. Socio-economic research is a multidisciplinary research area in which relationships between economic activities and their social environment are used to constitute socio-economic networks. Understanding patterns of economic organization interactions is essential if we are to uncover the mechanism and the structure of the socio-economic environment [12–17].

Network motif analysis is a sub-graph mining method proposed by Milo *et al.* [18]. Motifs are small (usually from three to seven nodes in size) connected sub-graphs within a given structure that appear in the network more frequently than they would if the network were completely random. This idea was first presented in 2002 by Uri Alon and his group, who discovered network motifs in the gene regulation network of the bacteria *E. coli* and in a large number of natural networks [18–20]. Following their seminal work, many studies have been conducted on this subject. Network motifs have been found in systems such as biological networks [19,21], electronic circuit networks [22], transport networks [23], and individual online affiliation networks [24]. Although some researchers have used network motifs to understand how social relationships influence individual interactions [25], little research has examined the impact of the socio-economic network itself on interactions among the organizations of the network [26]. To date most research has focused on static network motifs that are either a single-time snapshot of a phenomenon being investigated or an aggregate view over time [27]. Because most complex phenomena are time-varying, researchers are beginning to consider dynamic networks that evolve over time, such as how the network structure evolves in time with changes across physiological states [28–30]; however there is less research on time evolution of real networks in economic systems.

Both network topology and its time evolution must be considered if we are to understand the dynamics of a complex network. Our approach to this problem is to develop a method that allows us to analyze statistically the evolution of socio-economic motifs. We define dynamic network motifs as statistically significant sub-graph patterns that evolve in a network. By tracking the occurrence of dynamic motifs in a network that models organizational socio-economic interactions, we can observe the evolution of local configurations. Using this model we can then evaluate how a socio-economic network influences a company's decisions.

**Model.** – Studies of organizational interaction usually focus on buyer-seller business relationships as the primary source of competitive advantage —because they directly reflect market structure and competition [31]. However, to understand the mechanisms of business organization behavior we must also take into consideration social network influences. Figure 1 shows how social patterns affect organizational economic behavior in terms of a bipartite undirected network that contains two types of nodes. The $S$ set denotes nodes representing seller companies that provide products or services, and the $B$ set denotes nodes representing buyer companies that seek the products or services. A link exists between nodes $s_i$ and $b_j$ when seller $s_i$ signs at least one contract with company $b_j$. The network is bi-partite since there are no direct links between seller nodes or between buyer nodes. In many economic networks links and nodes appear and disappear and rewiring of links occurs frequently. Thus, the network shown in fig. 1 is time dependent and evolves over time. We thus define dynamic network motifs as statistically significant recurring evolutionary sub-graph patterns in a network over a specific time period. Suppose $G = \langle V, E\rangle$ is an arbitrary sub-graph in a given undirected network $N$ where an edge $e_{j,k} = \langle v_j, v_k\rangle$ connects $v_j$ and $v_k$, $e_{j,k} \in E$. Since the network evolves over time, $N_i$ represents the given network at $t_i$ and $G_i$ denotes the corresponding sub-graph at $t_i$.

For two time points $t_i < t_{i+1}$, let $G_i = \langle V, E_i\rangle$ and $G_{i+1} = \langle V, E_{i+1}\rangle$ denote the sub-graph of $N_i$ and $N_{i+1}$ with a timestamp at $t_i$ and $t_{i+1}$, respectively, and $N_i$ and $N_{i+1}$ represent the entire network at $t_i$ and $t_{i+1}$, respectively.

Each evolutionary type $j$ is regarded as a dynamic network motif candidate and its appearance numbers $C_j^{E_i-E_{i+1}}$ is recorded by scanning all possible evolutionary types from $N_i$ to $N_{i+1}$. Note that $J$ is the set containing the numbers of all possible evolutionary sub-graphs, so $j \in J$. We then use an "occurrence concentration" metric to evaluate the fraction of a certain evolutionary type occurrences of all possible evolutionary types to form a specific sub-graph [32]. Using $P_j^{E_i-E_{i+1}}$ to denote the occurrence concentration of evolutionary type $j$, we have

$$P_j^{E_i-E_{i+1}} = \frac{C_j^{E_i-E_{i+1}}}{\sum_{j\in J} C_j^{E_i-E_{i+1}}}. \quad (1)$$

In a way similar to static motif analysis, we generate random networks $R$ according to the real network $N$ to compute the statistical significance of an evolutionary motif. For a strict comparison, a randomized algorithm was used to form a random version to replicate all the evolutionary characteristics of real network $N$ from $t_i$ to $t_{i+1}$. At each time step, we assume that three distinct processes act to change the network.

i) The activity of sellers or buyers in the market. Here we only consider the sellers and buyers that maintain their relationships from $t_i$ to $t_{i+1}$. We thus use $m_s$ to represent the fraction of links remaining unchanged in $N$ from $t_i$ to $t_{i+1}$.

ii) The entry of new firms into the market. We use $n_b$ to denote the proportion of new buyer nodes entering and we use $n_s$ to denote the fraction of new seller nodes entering network $N$ from $t_i$ to $t_{i+1}$.

iii) The rewiring of surviving firms and new entering firms. Newcomers in the network either choose randomly to connect to other new nodes, or to link to surviving nodes. So $1 - m_s$ of links in real network to $N$ in $t_{i+1}$ represent the proportion of new links generated in $N$ from $t_i$ to $t_{i+1}$.

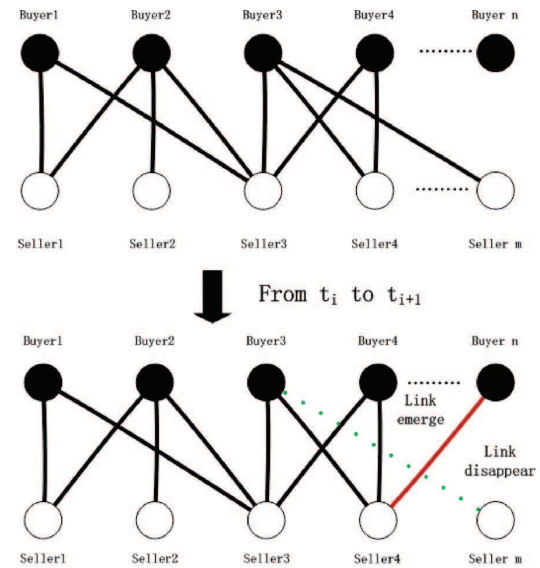


Fig. 1: (Colour on-line) Time evolution of the socio-economic network. A bi-partite network is shown with white nodes representing sellers and black nodes denoting buyers. Black lines between nodes mean that a contract exists between two companies at $t_i$. The red line represents a new link added to the network between $t_i$ and $t_{i+1}$ while the green dotted line means that a link disappears between $t_i$ and $t_{i+1}$.

Figure 2 shows the specific process of the randomization algorithm. We first scan possible sub-graph evolution patterns from real networks $N_i$ to $N_{i+1}$ and calculate the occurrence concentration of each motif studied. The same evolution pattern scan and occurrence concentration calculation are carried on from real network $N_i$ to random control network $R_{i+1}$ to test whether the occurrence of a certain sub-graph evolution pattern is significantly greater than its occurrence in random network. To eliminate random factors, we apply the single sample student test to determine whether the dynamic motif concentrations in the real network and the randomized network differ significantly. Here we focus on four-node sub-graphs. There

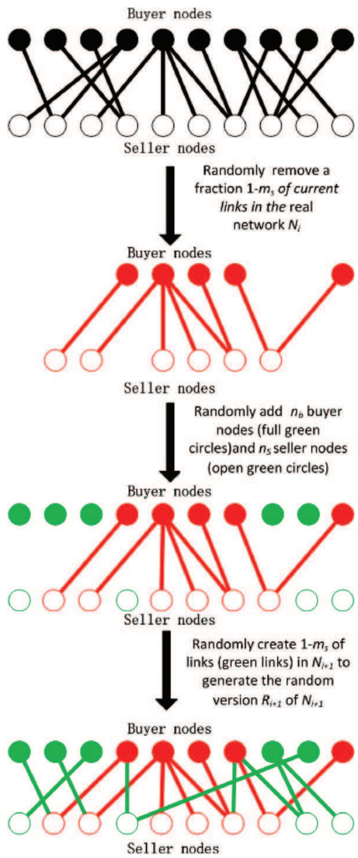


Fig. 2: (Colour on-line) Randomization algorithm to create a random network $R_{i+1}$ that captures the evolutionary characteristic of the real network evolving process from $N_i$ to $N_{i+1}$.

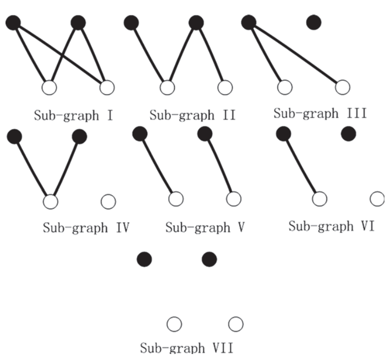


Fig. 3: Seven sub-graphs of four nodes (2 buyers and 2 sellers) in the bi-partite network.

are seven different four-node sub-graphs (see fig. 3) and seven sub-graph evolution patterns (see fig. 4).

**Data.** – We apply our model to two real business data sets: commercial ship chartering and build-to-order ships. Both data sets are from Clarkson Sin, a world shipping industry consulting firm [33]. Ship chartering data was recorded monthly from 1 January 2009 to 31 December 2012. Build-to-order ship data was recorded monthly from 1 January 2010 to 31 December 2012. Table 1 and table 2 show the summary of network characteristics.

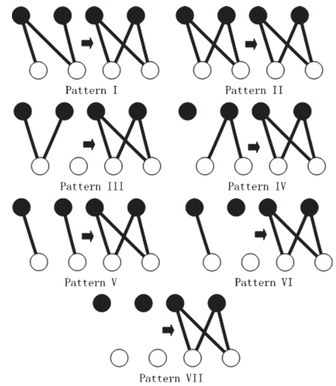


Fig. 4: Seven sub-graph evolution patterns are analyzed.

Table 1: Network characteristic summary of ship chartering data over 2009 to 2012.

| Year | Number of charters /owners | Number of links | Average degree | Charter/owners average degree |
|---|---|---|---|---|
| 2009 | 251/417 | 4845 | 14.5 | 19.3/11.6 |
| 2010 | 249/429 | 4942 | 14.6 | 19.8/11.5 |
| 2011 | 252/442 | 5272 | 15.2 | 20.9/11.9 |
| 2012 | 245/435 | 5142 | 15.1 | 21.0/11.8 |

Table 2: Network characteristic summary of ship order-to-build data over 2010 to 2012.

| Year | Number of shipyards /owners | Number of links | Average degree | Shipyard/owners average degree |
|---|---|---|---|---|
| 2009 | 797/1710 | 2563 | 2.04 | 3.22/1.5 |
| 2010 | 691/1591 | 2303 | 2.02 | 3.33/1.45 |
| 2011 | 635/1403 | 2051 | 2.92 | 3.23/1.46 |

Because nodes often enter or leave the market, we examine the year-to-year ratio between companies leaving and companies entering the market. Table 3 and table 4 show the leaving and entering rates of companies in the network in one-year intervals at the end of each sample period.

In the ship chartering market, approximately 17% of charter customers and 13% of ship owners leave the market after one year and approximately 16% of new charter customers and 14% of new ship owners enter the market. In

Table 3: Fraction of companies leaving from and entering into the charter market and owner market from 2009 to 2012 based on one-year interval comparisons.

| Time | Charter leaving rate | Charter new entering rate | Owner leaving rate | Owner new entering rate |
|---|---|---|---|---|
| 2009 to 2010 | 0.2 | 0.19 | 0.13 | 0.16 |
| 2010 to 2011 | 0.15 | 0.16 | 0.13 | 0.16 |
| 2011 to 2012 | 0.17 | 0.15 | 0.13 | 0.12 |

Table 4: Fraction of companies leaving from and entering into shipyard market and ship owner market from 2010 to 2012 based on one-year interval comparisons.

| Time | Shipyard leaving rate | Shipyard new entering rate | Owner leaving rate | Owner new entering rate |
|---|---|---|---|---|
| 2010 to 2011 | 0.32 | 0.18 | 0.60 | 0.52 |
| 2011 to 2012 | 0.32 | 0.24 | 0.60 | 0.48 |

the build-to-order ship market, the rates for leaving and entering are much higher for both shipyards and customer firms. Each year more than 30% of shipyards leave and 22% new shipyards enter the market, and 60% of the customer firms leave and 50% enter.

Since the rates of leaving and entering are very close, the total number of market players in both markets tend to be stable over time. Because firms frequently gain and lose customers, we track the changes of the links month by month. Figure 5 shows that number of network links fluctuates greatly over time. In the ship chartering network approximately one-quarter of the links disappears over a single month, *i.e.*, only 75% of the business relationship patterns among companies will be maintained in the following month. In the build-to-order ship market the network link fluctuation is much stronger. Only approximately 50% of the links remain unchanged, *i.e.*, when a customer firm orders a new ship from a shipyard, there is only a 50% probability that it will order another ship from the same shipyard in the following month.

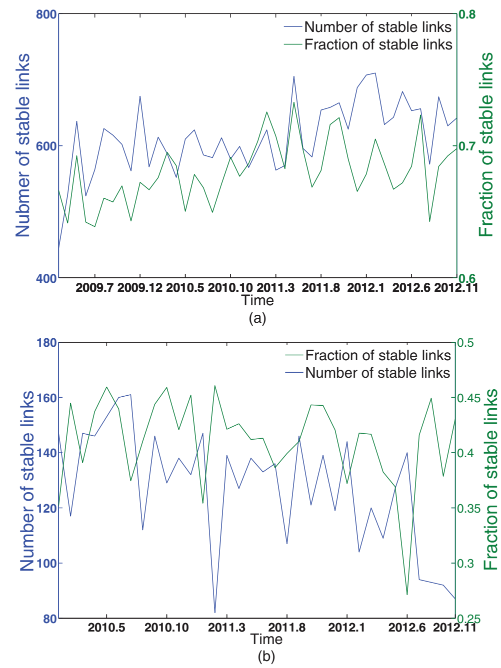


Fig. 5: (Colour on-line) Each link is traced month by month to see if it remains unchanged in next month. The number and fraction of stable links in both (a) ship chartering network and (b) new-building ship order-delivery network are shown.

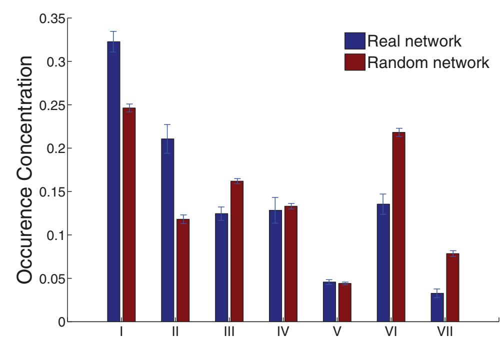


Fig. 6: (Colour on-line) Mean values are shown of the occurrence concentrations of each sub-graph evolution pattern in a real network represented by blue bars (ship chartering data from 2009 to 2012) and in the corresponding random control network represented by red bars. Error bars indicate one standard deviation.

**Results.** – We first apply our dynamic motif model to study the organization of the interconnections in the commercial ship chartering market (see fig. 6). Note that all seven evolution patterns occur statistically differently from that of the random version, since their $t$-test results are all significant (see table 5).

Moreover, the evolution patterns I and II appear much more frequently than the randomized version and the other five evolution patterns occur less frequently in the real network than in the random one. The evolution pattern I whose occurrence percentage is the highest, appears approximately 32% whereas in the random network it is around 25%. This means that in real economic-social networks two companies with the same customer have a much higher chance to have another common customer who already has relationship with one of these two companies. Notice that pattern II occurs 22% compared to the random network which only happens 12%. This means that in a real network it is almost twice more likely than

Table 5: $t$-test summary of the evolution patterns in the ship chartering network. The mean value of the occurrence in the real network is calculated as the average value of the occurrence of three periods that are: 2009 to 2010; 2010 to 2011; 2011 to 2012. The standard deviation of the occurrence in the real network is calculated as the average value of the differences of the occurrence of each period from the mean value.

| Evolution pattern | Mean value of occurrence concentration /std | | $P$-value |
|---|---|---|---|
| | Real network | Random network | |
| Pattern I | 0.32/0.01 | 0.24/0.004 | $2 \times 10^{-12}$ |
| Pattern II | 0.21/0.006 | 0.12/0.005 | $1 \times 10^{-12}$ |
| Pattern III | 0.12/0.005 | 0.16 /0.003 | $3 \times 10^{-23}$ |
| Pattern IV | 0.14/0.01 | 0.13/0.001 | $4 \times 10^{-7}$ |
| Pattern V | 0.05/0.001 | 0.04/0.001 | $6 \times 10^{-10}$ |
| Pattern VI | 0.13/0.01 | 0.23/0.002 | $5 \times 10^{-23}$ |
| Pattern VII | 0.033/0.002 | 0.08/0.002 | $2 \times 10^{-55}$ |

in a random network that four companies which have close connections tend to keep their current relationship in the future. The evolution patterns III and IV represent how much two companies that have one same customer tend to do business with another company which had no relationship with both of them before. These two patterns occur in real networks much less frequently than in random ones. In addition, patterns V, VI and VII representing the occurrence of two nodes initially without any mutual common neighbors will in the future develop a fully connected local structure. These three evolution patterns occur in socio-economic networks less than in random ones. Thus, only patterns I and II appear significantly more than random ones, meaning that a node that is linked to the motif has a high chance i) to create another link in the motif or ii) to stay in the motif. Attracting new nodes to the motif is less probable than in random networks, since people prefer to connect to someone that is connected to his friends, see also the analogous case in regular social networks [25].

Next, we analyze the same dynamic motif evolution model in fig. 4, in the second data set of ship order-to-build market. Although ship order-to-build market has a totally different business behavior from the chartering market and its network is much sparser than the chartering network, we still find similar evolution patterns as those seen in fig. 7 and table 6.

These results support our conclusions from the ship chartering market. The evolution pattern I whose occurrence percentage is the highest, appears approximately 29%. This value is very close to the occurrence (32%) of this motif in the ship chartering market. Occurrence concentration is more than twice that of the random network which is only 12%. This further supports our previous conclusion that if two nodes A and B share the same neighbor C, and A has another neighbor D, it is highly possible that B will be connected to D in the future. The

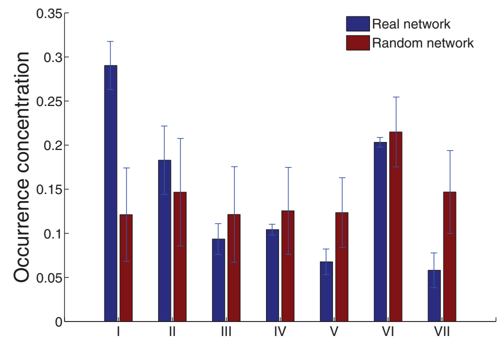


Fig. 7: (Colour on-line) Mean values are shown of the occurrence concentrations of each sub-graph evolution pattern in a real network (ship order-to-delivery data from 2010 to 2012) and in the corresponding random network. Error bars indicate one standard deviation.

Table 6: $t$-test summary of the evolution patterns in the ship build-to-order network. The mean value of the occurrence in the real network is calculated as the average value of the occurrence of two periods that are 2010 to 2011 and 2011 to 2012. The standard deviation of the occurrence in the real network is calculated as the average value of the differences of the occurrence of each period from the mean value.

| Evolution pattern | Mean value of occurrence concentration /std | | $P$-value |
|---|---|---|---|
| | Real network | Random network | |
| Pattern I | 0.29/0.03 | 0.12/0.05 | $5 \times 10^{-30}$ |
| Pattern II | 0.18/0.04 | 0.14/0.07 | $2 \times 10^{-3}$ |
| Pattern III | 0.09/0.02 | 0.12/0.06 | $2 \times 10^{-21}$ |
| Pattern IV | 0.10/0.01 | 0.13/0.05 | $5 \times 10^{-34}$ |
| Pattern V | 0.07/0.03 | 0.12/0.05 | $10 \times 10^{-5}$ |
| Pattern VI | 0.21/0.01 | 0.22/0.05 | $2 \times 10^{-4}$ |
| Pattern VII | 0.06/0.02 | 0.15/0.05 | $9 \times 10^{-12}$ |

evolution pattern II appears approximately 18% in contrast to the random network which is around 14%. This supports our second conclusion that nodes forming a fully connected sub-graph tend to maintain their past edges. Moreover, the other five evolution patterns occur in a real network significantly less than in a random network.

Our results show that patterns I and II can be regarded as dynamic motifs, since they appear significantly more than the random ones. Moreover, dynamic motifs I and II exist in both networks that we analyzed. All other sub-graphs appear less than the random ones since the total appearance must be one. No matter how different these two networks are, such as the average degree and network density, the occurrence of these two dynamic motifs appears similar. This suggests the existence of a strong fundamental social influence. High occurrence concentration of motif I reveals that companies with common customers tend to make similar decisions. Indeed, companies with common customers have relatively more opportunities to interact with each other to share information than those without mutual customers so that interactions among organizations will influence their business decision. Motif II reveals that companies embedded in very close business ties with each other prefer to keep the current business relationship for a long time. For example, if both two buyers buy from two sellers, it is highly possible they will keep this relationship in the future. Organizational interactions and social relationships play an essential role in the dynamics of the economic society. We find that the organization interaction behavior can be characterized by influence and persistence. Note also that patterns III and IV occur almost in the same frequency in both systems, probably due to their symmetric nature.

**Conclusion.** – We have empirically characterized dynamic organization interactions in the socio-economic environment. We propose a dynamic motif model that incorporates features of social influence of organizations' economic behavior. By detecting dynamic motifs from firms' buyer-seller transaction data, we found two motifs which present similar organizational interaction patterns in two different business networks. This suggests that we can extract important social effects of organizational interaction in the socio-economic behavior. These findings provide a valuable insight into the relationship between the economic function and the social network structure.

* * *

We wish to thank ONR (Grant N00014-09-1-0380, Grant N00014-12-1-0548, Grant N62909-14-1-N019), DTRA (Grant HDTRA-1-10-1-0014, Grant HDTRA-1-09-1-0035), NSF (Grant CMMI 1125290), the European, MULTIPLEX, CONGAS and LINC projects, DFG, the Next Generation Infrastructure (Bsik) and the Israel Science Foundation for financial support. We also thank the FOC program of the European Union for support. XZ acknowledges the support of the Ministry of Education of the People's Republic of China via Grant 10YJC630394.

REFERENCES

[1] Albert R., Jeong H. and Barabási A. L., *Nature (London)*, **401** (1999) 130.
[2] Cohen R., Erez K., Ben-Avraham D. and Havlin S., *Phys. Rev. Lett.*, **85** (2000) 4626.
[3] Newman M., *Networks: An Introduction* (Oxford University Press, Oxford) 2010, p. 720.
[4] Cohen R. and Havlin S., *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge) 2010.
[5] Fujiwara Y. and Aoyama H., *Eur. Phys. J. B*, **77** (2010) 565.
[6] Newman M. E. J., *Proc. Natl. Acad. Sci. U.S.A.*, **98** (2001) 404.
[7] Liben-Nowell D. and Kleinberg J., *J. Am. Soc. Inf. Sci. Technol.*, **58** (2007) 1019.
[8] Helbing D., *Nature (London)*, **497** (2013) 497.
[9] Jiang Z. Q., Xie W. J., Li M. X., Podobnik B., Zhou W. X. and Stanley H. E., *Proc. Natl. Acad. Sci. U.S.A.*, **110** (2013) 1600.
[10] Mason W. and Watts D. J., *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012) 764.
[11] Zhang X., Podobnik B., Kenett D. Y. and Stanley H. Eugene, *Physica A*, **415** (2014) 43.
[12] Gulati R., Sytch M. and Tatarynowicz A., *Organ. Sci.*, **23** (2012) 449.
[13] Kilduff M. and Brass D. J., *Acad. Manag. Ann.*, **4** (2010) 317.
[14] Uzzi B., *Am. Sociol. Rev.*, **64** (1999) 481.
[15] Rauch J. E., *J. Econ. Lit.*, **39** (2001) 1177.
[16] Borgatti S. P. and Foster P. C., *J. Manag.*, **29** (2003) 991.
[17] Tsai W., *Acad. Manag. J.*, **44** (2001) 996.
[18] Milo R., Shen-Orr S., Itzkovitz S., Kashtan N., Chklovskii D. and Alon U., *Science*, **298** (2002) 824.
[19] Kashtan N., Itzkovitz S., Milo R. and Alon U., *Phys. Rev. E*, **70** (2004) 031909.
[20] Kashtan N. and Alon U., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 13773.
[21] Kittisopikul M. and Süel G. M., *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010) 13300.
[22] Valverde S. and Solé R. V., *Phys. Rev. E*, **72** (2005) 026107.
[23] Kaluza P., Kölzsch A., Gastner M. T. and Blasius B., *J. R. Soc. Interface*, **7** (2010) 1093.
[24] Onnela J. P., Saramäki J., Hyvönen J., Szabó G., De Menezes M. A., Kaski K., Barabási A. L. and Kertész J., *New. J. Phys.*, **9** (2007) 179.
[25] Gallos L. K., Rybski D., Liljeros F., Havlin S. and Makse H. A., *Phys. Rev. X*, **2** (2012) 031014.
[26] Ohnishi T., Takayasu H. and Takayasu M., *J. Econ. Interact. Coord.*, **5** (2010) 171.
[27] Kovanen L., Karsai M., Kaski K., Kertész J. and Saramäki J., *J. Stat. Mech.* (2011) P11005.
[28] Bashan A., Bartsch R. P., Kantelhardt J. W., Havlin S. and Ivanov P. C., *Nat. Commun.*, **3** (2012) 702.
[29] Bartsch R. P. and Ivanov P. C., *Networks of Networks: The Last Frontier of Complexity* (Springer International Publishing, Switzerland) 2014, p. 203.
[30] Bartsch R. P. and Ivanov P. C., *Nonlinear Dynamics of Electronic Systems* (Springer) 2014, p. 702.
[31] Atalay E., Hortaçsu A., Roberts J. and Syverson C., *Proc. Natl. Acad. Sci. U.S.A.*, **108** (2011) 5199.
[32] Kabutoya Y., Nishida K. and Fujimura K., in *Web Technologies and Applications*, edited by Du Xiaoyong, Fan Wenfei, Wang Jianmin, Peng Zhiyong and Sharaf Mohamed A., Vol. **6612** (Springer) 2011, p. 321.
[33] Clarksons Ltd., `http://www.clarksons.net`.