Motifs in weighted networks and their Hirsch subgraphs

Simon S. Li [a, b], Ronald Rousseau [c, d] and Fred Y. Ye [a, b]
[a] School of Information Management,
Nanjing University, Nanjing 210023, CHINA
[b] Jiangsu Key Laboratory of Data Engineering and Knowledge Service,
Nanjing 210023, CHINA
[c] KU Leuven, Dept. Mathematics, Celestijnenlaan 200B,
B-3000 Leuven (Heverlee), BELGIUM
[d] Universiteit Antwerpen, B-2000 Antwerpen, BELGIUM
e-mail: chqlee87@gmail.com; ronald.rousseau@kuleuven.be;
yye@nju.edu.cn (corresponding author)

ABSTRACT

Motifs are patterns of interactions occurring in complex networks and characterized by the fact that they occur significantly more than expected. Network indicators derived from motifs are introduced in this contribution. They provide yet another numerical view on network structures. Based on the notion of a motif's h-index the motif's Hirsch subgraph is constructed. This subgraph is a new characteristic structure in weighted networks. Use of these notions has been illustrated in two case studies: one involving a co-authorship and one involving a co-keyword network.

Keywords: Motif; Motif's H-Index; Hirsch subgraph; Information network; Weighted network.

INTRODUCTION

Since the end of the 20th century the study of networks has become a hot topic. This renewed interest in networks was largely due to the ubiquity of the Internet and consequent work by, e.g., Rousseau (1997), Watts and Strogatz (1998) Barabási and Albert (1999) and Albert and Barabási (2002). Terms such as "small world networks", "scale-free networks" and "complex networks" became popular in many fields including molecular biology, genetics, medical informatics and neuroscience (Jeong et al. 2000; Yarfitz and Ketchell 2000; Girvan and Newman 2002; Raeymaekers 2002; Newman 2003, 2012; Barabási and Oltvai 2004; Mittler et al. 2004; Boccaletti et al. 2006). Also information scientists recognized the importance of network theory (Kretschmer 1997; Chen et al. 2002; Otte and Rousseau 2002; Börner et al. 2007). In this contribution we will especially focus on network motifs.

Motifs are patterns of interactions occurring in complex networks and characterized by the fact that they occur significantly more than expected (Milo et al. 2002). To be precise: a motif M is a pattern that re-occurs in a network. Mathematically a motif can be said to be the name of the equivalence class of a set of isomorphic subgraphs. We refer to these subgraphs as the subgraphs corresponding to motif M.

Motifs occur in transcription networks, signaling and neuronal networks, economics, information and social networks and other types of networks (Sporns and Kötter 2004; Krumov et al. 2011; Zhang et al. 2014). In all these networks motifs serve as basic building blocks. Barabási and Oltvai (2004) pointed out that the occurrence of motif clusters seems to be a general property of all real networks. In Alon (2007) the author points out that one of the most common composite motifs is a negative feedback loop between two proteins. An example is shown in Box 1 of (Alon 2007). Krumov et al. (2011) analyzed co-authorship networks and found that numbers of received citations of group of authors were correlated with the occurrence of certain motifs. In particular they observed that the box motif, i.e. a closed chains of four authors, had the highest average citation frequency per link. Although these authors stated that their results were robust across ways of mapping citation frequencies onto the co-author graph, Klosik et al. (2014) pointed out that the results obtained by Krumov et al., (2011) were highly sensitive to the exact implementation of an author disambiguation procedure. Choobdar et al. (2012) proposed a new method to incorporate edge weight information in motif mining. Further, Han et al. (2013) stated that the motif-based node degree and edge degree could be applied to measure the importance of nodes and edges in a network. For further use we recall that the strength of a connected subgraph in a weighted network is simply the sum of the weights of the links constituting this subgraph.

The h-index idea as introduced by Hirsch (2005) for individual scientists has been applied in many other cases, including network analysis (Korn et al. 2009; Schubert et al. 2009). We make a distinction between three cases: h-indices of networks, h-indices of nodes and h-indices of links.

A network, weighted or not, has node h-index h_N if h_N is the largest natural number such that h_N of its nodes have a degree at least equal to h_N (Schubert et al. 2009). Similarly, a weighted network has h-strength h_S if h_S is the largest natural number such that h_S of its link weights (link strengths) have strength at least equal to h_S (Zhao et al. 2014). Zhao, Rousseau and Ye (2011) introduced the h-degree of a node in a weighted undirected network as follows: The h-degree (d_h) of node n in a undirected weighted network is equal to $d_h(n)$, if $d_h(n)$ is the largest natural number such that n has $d_h(n)$ links each with strength at least equal to $d_h(n)$. This notion was extended to the case of directed networks in (Zhao et al. 2012). A weighted network's h-core, or h-core in short, is defined as the set of all nodes with h-degree at least equal to h and their links, if they exist.

In this application we extract simple motifs using the FANMOD software (Wernicke and Rasche 2006). Next we combine the ideas of motifs and h-type indices. This leads to the notion of a motif's Hirsch subgraph. In this article we focus on a simple motif, namely a 3-node triangular motif, but point out that the method can be extended to other multi-node motifs.

METHOD

Characteristic indicators of motifs: the MSIKCA requirements

According to (Milo et al. 2002) network motifs are connected subgraphs (sub-networks) of a given network that meet the following three criteria:

(a) The probability that this subgraph appears in a randomized network an equal or greater number of times than in the real network under study is smaller than a cut-off value p_c. In (Milo et al., 2002) the value for p_c was taken equal to 0.01 and was estimated from 1000 randomized networks. These randomized networks had the same number of nodes and the same single node properties as the network under study. Concretely: each node in the randomized network had the same number of incoming and outgoing edges as the original one in the real network. Furthermore, the randomized networks used to calculate the significance of n-node subgraphs were generated so as to preserve the same number of all (n-1)-node subgraphs as in the real network.

(b) The number of times the motif appears in the real network, and this with distinct sets of nodes, is at least four.

(c) The number of appearances in the real network (N_{real}) is significantly larger than in the randomized networks (N_{rand}). Concretely, the following criterion was used: $N_{real} - N_{rand} > 0.1(N_{rand})$. This extra requirement was added to avoid detecting as motifs some common subgraphs that have only a slight difference between N_{rand} and N_{real} but have a narrow distribution in the randomized networks.

As these requirements were proposed by Milo, Shen-Orr, Itzkovitz, Kashtan, Chklovskii and Alon (Milo et al. 2002) we refer to them as the MSIKCA requirements. These concrete, somewhat overlapping, restrictions determining which subgraphs are considered to be motifs in a network, are described in the Supplementary Materials of (Milo et al. 2002). We note though that these requirements are ad hoc and for this reason do not provide a mathematically precise definition of a motif.

The importance of motifs is illustrated by the following example. Milo et al. (2002) show that there are 13 different three-node subgraphs in a directed network. Yet only one of them, the so-called feed-forward loop occurs in the genetic network of E. coli. This network has also a second motif, which is a 4-node subnetwork, known as the bi-fan. For details and drawings of these motifs we refer the reader to (Milo et al. 2002).

The procedure for finding motifs, especially in anything but a trivially small network clearly needs dedicated software. Such a program, designed by Wernicke (2005) and called FANMOD is freely available at http://theinf1.informatik.uni-jena.de/motifs/. Note though that counting motifs or subnetworks in general is not trivial as different methods are feasible. Indeed, there exist different ways to count motif frequency referred to as F1, F2 and F3 (Schreiber and Schwöbbermeyer 2005). Frequency concept F1 has no restrictions and considers all matches, even if elements of the target graph have to be used several times. Concept F2 allows the sharing of vertices but not of edges and therefore calculates the number of instances of a motif that have disjoint edges. Finally, concept F3 only considers subnetworks to be different if their vertices and edges are disjoint and hence, they can be seen as non-overlapping clusters. This third frequency is known as the uniqueness value of a motif (Schwöbbermeyer 2008). Concept F1 is the common method and is the one used by FANMOD. FANMOD can find motifs of 3 to 8 nodes in a given network and can calculate the following three indicators (Milo et al. 2002; Wernicke and Rasche 2006; Han et al. 2013).

Indicator 1: Relative subgraph and motif frequency

Given a connected subgraph V with n nodes, if V appears N_{real} times in a real network X and if there exist N_{sub} connected subgraphs with n nodes in X, then the relative frequency of subgraph V in a given network is defined as

$$f(V) = \frac{N_{real}}{N_{sub}} \tag{1}$$

If V is a motif of the network, then f(V) is called the relative motif frequency.

Indicator 2: motif p-value

Let M be a motif, or generally an equivalence class of connected subgraphs, which occurs N_{real} times in a real network. Now one considers *n* randomized networks as described above and for each i = 1, … , n one compares N_{rand}^{i}, the number of times M occurs in the i-th random network, with N_{real}. If $N_{rand}^{i}/ N_{real} \geq 1$ then $p_i = 1$, otherwise $p_i = 0$. Finally the p-value of M is:

$$p = \frac{\sum_{i=1}^{n} p_i}{n} \tag{2}$$

If p is smaller than the given cut-off value p_c, M is a motif. The smaller p, the more important the motif.

Indicator 3: z-score

For a motif, let N_{real} be its number of appearances in the real network and N_{rand}^{i} its number of appearances in the i-th randomized network. Denoting the mean of all N_{rand}^{i} as $\langle N_{rand} \rangle$ and the corresponding standard error as σ_{rand}, we define the z-score of the real network by:

$$z = \frac{N_{real} - \langle N_{rand} \rangle}{\sigma_{rand}} \tag{3}$$

The three indicators characterize a motif in a network. Next we will apply the ideas explained above in the context van h-indices in weighted networks.

Motifs and h-indices

Given a k-node ($k \geq 3$) motif in a weighted network we define the following two notions: the motif's h-index and the motif's Hirsch subgraph. We recall that the corresponding subgraphs of a motif are subgraphs in a weighted network and hence each of them has a strength as defined earlier. If weights are natural numbers then the strength of a k-motif is at least equal to k-1.

Definition 1. If M is a k-node motif in a given weighted network, then its h-index, denoted as h_M, is defined as the largest natural number such that there are h_M corresponding subgraphs with strength at least equal to h_M.

Definition 2. If M is a k-node motif in a given weighted network then M's Hirsch subgraph is the union of all corresponding subgraphs with strength at least h_M.

We remark that a motif's Hirsch subgraph is not necessarily connected. Moreover, a Hirsch subgraph may be the union of more than h_M subgraphs. Indeed, this may happen if the subgraph at rank h_M has weight h_M and one or more subgraphs (at rank h_M +1, etc.) also have weight h_M.

By the above definitions, the algorithm for finding a motif's Hirsch subgraph in a given weighted network is constructed as follows. If a motif is already known one may start at step 1.

Step 0: Obtaining all k-node motifs ($k \geq 3$) in a weighted network X, using the FANMOD software (Wernicke & Rasche, 2006) or any other software that is capable of finding motifs;

Step 1. Choose a motif M;

Step 2: Extracting all corresponding subgraphs in X, using the ESU (Enumerate SUbgraphs) algorithm (Wernicke, 2005; Wernicke & Rasche, 2006);

Step 3: Computing the strengths of the corresponding subgraphs and ranking them from large to small;

Step 4: Based on step 3, determine h_M;

Step 5: Construction of the Hirsch subgraph of M in network X.

We provide an example of this algorithm, based on Figure 1.

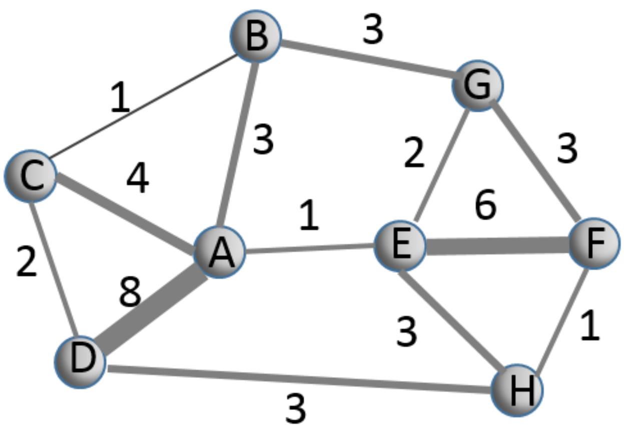
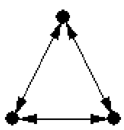


Figure 1: Weighted undirected network U

Using FANMOD, we computed the main parameters of the 3-node triangle motif (which has ID 238 in FANMOD). Results are shown in Table 1.

Table 1: The main parameters of the 3-node motif M in U

ID	motif	Frequency (original)	mean-frequency (random)	standard-error (random)	z-score	p-value
238		18.18%	0.10%	0.01088	16.621	0

The concrete computation consists of the following steps:

Step 0-1: We have chosen the triangle (ID 238) and verified, see Table 1, that it is indeed a motif for the network U shown in Figure 1. The calculation showing that the frequency of this motif in network U is 18.18 percent is given in the Appendix.

Step 2: Using the ESU algorithm, we find four corresponding subgraphs: {A,B,C}, {A,C,D}, {E,F,G}, {E,F,H}.

Step 3: Computing the strengths of each 3-node subgraph, we obtain the values 14, 11, 10, 8 when ranked from largest to smallest.

Step 4: Clearly $h_M = 4$.

Step 5: Bringing these four subgraphs together leads to M's Hirsch subgraph, shown in Figure 2.

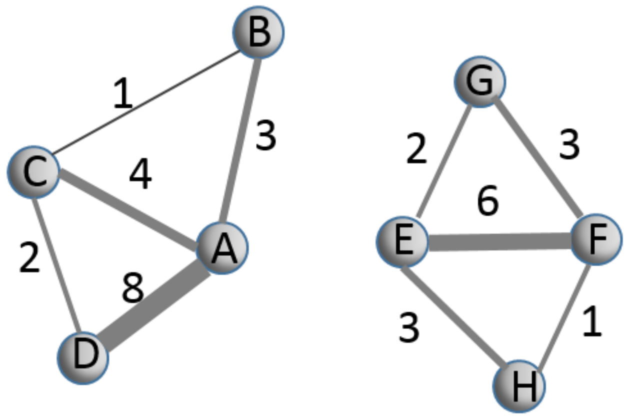


Figure 2: Hirsch' subgraph of motif M in graph U

As a further illustration we worked out two real world case studies.

MATERIALS AND METHOD

Two datasets are considered as case studies. Dataset 1 was retrieved from the Web of Science (WoS) databases (SCI, SSCI and A&HCI) on June 8, 2015, using the following search strategy:
TS = ((“information retrieval”) or (“information search”)) and PY = 2010-2014.

Dataset 2 was retrieved from the Web of Science (WoS) databases (SCI, SSCI and A&HCI) on June 15, 2015, using the following search strategy:
TS = (bibliometric* or informetric* or scientometric* or webmetric* or cybermetric*) and PY = 2005-2012.

These datasets contain, respectively, 3,411 and 3,058 items. The first one is used to construct a co-author network, while the second one is used to construct a co-keyword network. Keywords are retrieved from the DE field (Authors Keywords) and the ID field (Keywords Plus).

RESULTS

As a first step we construct weighted networks based on these datasets. For dataset 1, we construct a co-author network in which authors are connected if they have co-authored at least one article and weights represent the number of actual collaborations between authors. For dataset 2, we construct a co-keyword network. Words are connected if they co-occur in the Authors or Keywords Plus fields, weights represent the number of times words co-occur. Basic network parameters are shown in Table 2.

Table 2: The basic network parameters of datasets 1 and 2

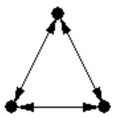
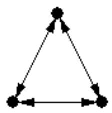
	Subnetwork Parameters	**Complete co-author network**	**Network consisting of all motifs**	**h-core**	**h-subnet**	**M's Hirsch subgraph**
Dataset1	Nodes	7856	6611	33	14	26
	Edges	15398	14613	53	9	30
	Average node degree	3.9201	4.42	3.21	1.29	2.31
	Average of edge strength	1.08	1.08	3.42	6.11	4.57
	Network density	0.0005	0.0007	0.1004	0.0989	0.0923
	Average clustering coefficient	0.78	0.93	0.85	0.17	0.98
	Average of degree centrality	0.0005	0.0007	0.1004	0.0989	0.0923
	Parameters	**Complete co-keyword network**	**Network consisting of all motifs**	**h-core**	**h-subnet**	**M's Hirsch subgraph**
Dataset2	Nodes	4882	4840	7	15	126
	Edges	21949	21505	22	19	250
	Average of node degree	8.9918	8.8864	6.2857	2.5333	3.9683
	Average of edge strength	1.1677	1.1591	19.5909	27.7368	4.94
	Network density	0.0018	0.0018	1.0476	0.1810	0.0317
	Average clustering coefficient	0.8800	0.8864	1.0	0.2273	0.9844
	Average of degree centrality	0.0018	0.0018	1.0476	0.1810	0.0317

For the definitions of the network indicators shown in Table 2 we refer the reader to

(Wasserman and Faust 1994; Otte and Rousseau 2002).

Choosing the triangle subgraph we obtained the results shown in Table 3. This shows that the triangle is indeed a motif for these two networks. The basic parameters of this motif's Hirsch subgraph are shown in the last column of Table 2.

Table 3: Main parameters of the triangle motif in two real networks

Dataset Type	**Motif**	**Frequency [original]**	**Mean Frequency [random]**	**Standard Error [random]**	**z - score**	**p - value**
1 Co-author		50.822%	0.0004 %	1.2 e-005	42.42	0
2 Co-keyword		2.2668%	0.0046 %	0.00016	137.8	0

Although the main parameters of this motif in the co-authorship network and in the co-keyword network are quite different, their p-values are so small that they can only be reported as zero. This means that in both networks the triangle motif is a very important structure.

The Hirsch subgraphs of the triangle motif in datasets 1 and 2 are shown in Figures 3 and 4 respectively, drawn using NetDraw. We recall that the nodes belonging to these subgraphs are not necessarily high degree nodes. Nodes in a triangle Hirsch subgraph have at least degree two, but nodes with exactly degree two may occur often. Figure 3 has many nodes with degree two. Node ScienceWatch.com (middle left) in the co-keyword network also has only degree two.

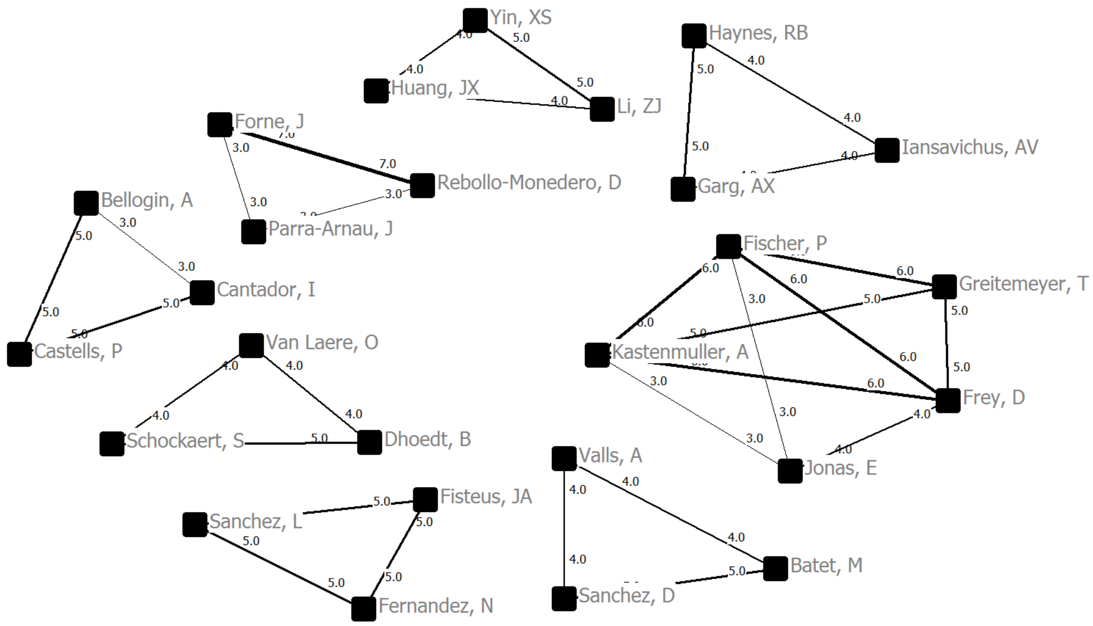


Figure 3: The triangle's Hirsch subgraph of the co-authorship network (h_M =13)

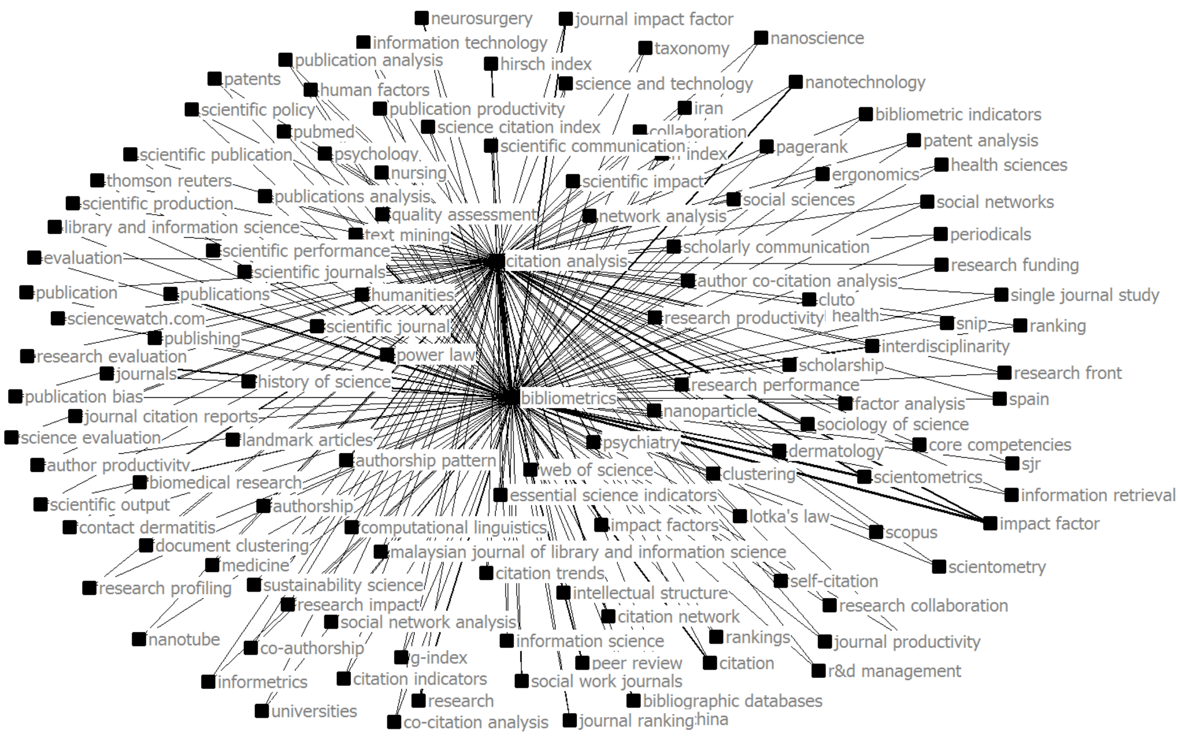


Figure 4: The triangle's Hirsch subgraph of the co-keyword network (h_M =95)

The Hirsch graphs shown in Figures 3 and 4 are very different, illustrating that their characteristics depend on the concrete weighted information network.

ANALYSIS AND DISCUSSION

In this section, we first formulate some simple mathematical results.

If M is a k-motif ($k \geq 3$) not necessarily the triangle motif, and a link l_1 has strength larger than or equal to $h_M - 1$, then each connected subgraph corresponding to this motif M and build on this link, i.e. the link l_1 is one edge of this subgraph, is part of this motif's Hirsch graph. This is obvious as there are at least two edges in a 3-motif (k-1 links in general) and a link has at least weight 1, so that a subgraph including l_1 has at least strength $(h_M\text{-}1) + 1 = h_M$ and hence is part of this motif's Hirsch network.

The following inequality holds for the triangle motif M

$$3 \leq h_M \leq N_{real} \qquad (4)$$

Proof. By definition a motif occurs at least four times. Moreover, a triangle has a least strength 3. This proves the inequality.

Using the same arguments we have the following inequality for any k-motif (k≥3)

$$min(k-1,4) \leq h_M \leq N_{real} \qquad (5)$$

Motif strengths follow a power-law. Furthermore, checking if, at least for our examples, motif strengths follow a power-law distribution, we find this to hold approximately as illustrated in Figure 5.

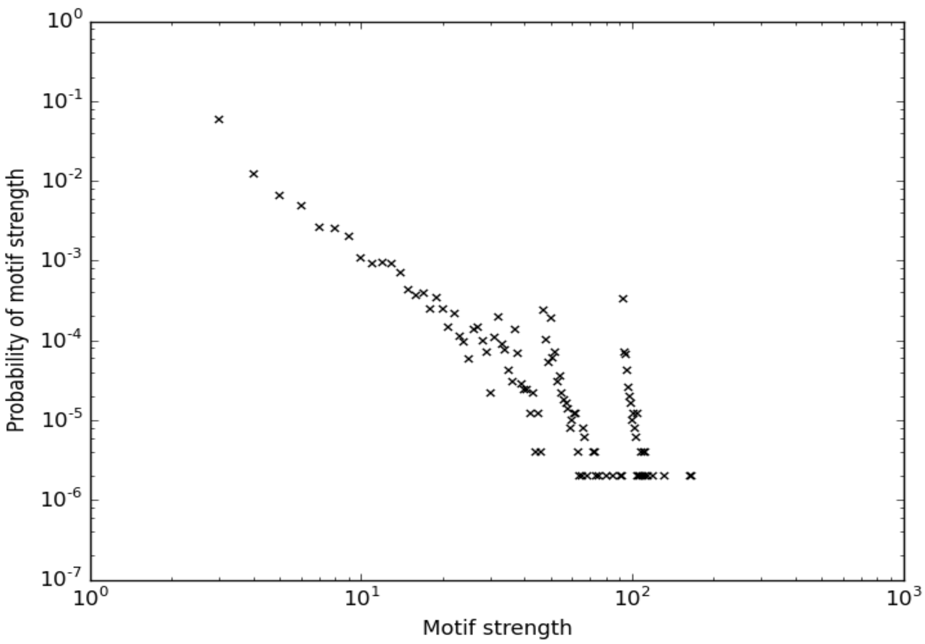


Figure 5: Distribution of motif-strength of dataset 2 (co-keyword network)

The power-law package developed by Alstott et al. (2014, available at https://pypi.python.org/pypi/powerlaw), finds a best-fitting power-law parameter α = 2.58.

CONCLUSION

Motifs are known to be significant structures in many types of networks. By considering the strength of motifs in weighted networks we introduced an h-index of a motif and constructed the motif's Hirsch subgraph. These motif derived network indicators provide yet another numerical view on network structures. A motif's Hirsch subgraph is a new characteristic structure in such networks. Use of these notions has been illustrated in two case studies: one involving a co-authorship network and one involving a co-keyword network.

Our work has a number of limitations. First we realize that as a motif's Hirsch subgraph is a union of local structures, it might be of less importance for the network as a whole. Moreover, we only processed one 3-node motif. This leaves many k-node motifs to be studied. However, when k ≥ 4, the computational complexity increases quickly. At present, we do not know a software tool that can handle the computation when k > 8.

We consider the motif's Hirsch subgraph which combines the h-index idea with the notion of motifs in weighted networks to be a new and interesting cluster-structure. As such we express the hope that our work will stimulate further investigations involving motifs.

ACKNOWLEDGEMENTS

Simon S. Li and Fred Y. Ye acknowledge the financial support from NSFC Grants No. 71173187 and 71673131 and the Jiangsu Key Laboratory Fund for financial supports. Ronald Rousseau is supported by NSFC Grant number 71573225.

REFERENCES

Albert, R. and Barabási, A.-L. 2002. Statistical mechanics of complex networks. *Review of Modern Physics*, Vol. 74, no. 1: 47–97.

Alon, U. 2007. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, Vol. 8, no. 6: 450-461.

Alstott, J., Bullmore, E., and Plenz, D. 2014. Powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE*, Vol. 9, no. 1: e85777.

Barabási, A.- L., and Albert, R. 1999. Emergence of scaling in random networks. *Science*, Vol. 286, no. 5439: 509-512.

Barabási, A. L., and Oltvai, Z. N. 2004. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, Vol. 5, no. 2: 101-113.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.U. 2006. Complex networks: Structure and dynamics. *Physics Reports*, Vol. 424, no. 4: 175-308.

Börner, K., Sanyal, S., and Vespignani, A. 2007. Network science. *Annual Review of Information Science and Technology*, Vol. 41, no. 1: 537-607.

Chen, CM., Cribbin, T., Macredie, R., and Morar, S. 2002. Visualizing and tracking the growth of competing paradigms: Two case studies. *Journal of the American Society for Information & Technology*, Vol. 53, no. 8: 678-689.

Choobdar, S., Ribeiro, P., and Silva, F. 2012. Motif mining in weighted networks. *12th IEEE International Conference on Data Mining (ICDM)*, Brussels: IEEE. pp. 210-217.

Girvan, M., and Newman, M.E.J. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the USA*, Vol. 99, no. 12: 7821-7826.

Han, H., Liu W.L., and Wu L.Y. 2013. The measurement of complex network based on motif. *Acta Physica Sinica*, Vol. 62, no. 16: 168904.

Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 102, no. 46: 16569-16572.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabási, A.L. 2000. The large-scale organization of metabolic networks. *Nature*, Vol. 407, no. 6804: 651-654.

Klosik, D. F., Bornholdt, S., and Hütt, M. T. 2014. Motif-based success scores in coauthorship networks are highly sensitive to author name disambiguation. *Physical Review E*, Vol. 90, no. 3: 032811.

Korn, A., Schubert, A., and Telcs, A. 2009. Lobby index in networks. *Physica A*, Vol. 388, no. 11: 2221–2226.

Kretschmer, H. 1997. Patterns of behaviour in coauthorship networks of invisible colleges. *Scientometrics*, Vol. 40, no. 3: 579-591.

Krumov, L., Fretter, C., Müller-Hannemann, M., Weihe, K., and Hütt, M. T. 2011. Motifs in co-authorship networks and their relation to the impact of scientific publications. *The European Physical Journal B*, Vol. 84, no. 4: 535-540.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. 2002. Network motifs: simple building blocks of complex networks. *Science,* Vol. 298, no. 5594: 824–827.

Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. 2004. Reactive oxygen gene network of plants. *Trends in Plant Science*, Vol. 9, no. 10: 490-498.

Newman, M.E.J. 2003. The structure and function of complex networks. *SIAM Review,* Vol. 45, no. 2: 167–256.

Newman, M.E.J. 2012. Communities, modules and large-scale structure in networks. *Nature Physics,* Vol. 8, no. 1: 25-31.

Otte, E., and Rousseau, R. 2002. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, Vol. 28, no. 6: 441- 453.

Raeymaekers, L. 2002. Dynamics of Boolean networks controlled by biologically meaningful

functions. *Journal of Theoretical Biology*, Vol. 218, no. 3: 331-341.
Rousseau, R. 1997. Sitations: an exploratory study. *Cybermetrics*, 1(1), Available at : http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html
Schreiber, F., and Schwöbbermeyer, H. 2005. Frequency concepts and pattern detection for the analysis of motifs in networks. *Transactions on Computational Systems Biology* III (LNBI 3737). Springer-Verlag: Berlin, Heidelberg: 89–104.
Schubert, A., Korn, A., and Telcs, A. 2009. Hirsch-type indices for characterizing networks. *Scientometrics*, Vol. 78, no. 2: 375-382.
Schwöbbermeyer, H. 2008. Network motifs. *Analysis of Biological Networks* (B.H. Junker & F. Schreiber, eds.). Wiley, NJ.
Sporns, O. and Kötter, R. 2004. Motifs in brain networks. *PLoS Biolology,* Vol. 2, no. 11: 369.
Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications.* New York and Cambridge (UK): Cambridge University Press.
Watts, D.J., and Strogatz, S.H. 1998. Collective dynamics of 'small-world' networks. *Nature*, Vol. 393, no. 6684: 440-442.
Wernicke, S. 2005. A faster algorithm for detecting network motifs. *Algorithms in Bioinformatics*. Springer-Verlag: Berlin, Heidelberg: 165-177.
Wernicke, S., and Rasche, F. 2006. FANMOD: a tool for fast network motif detection. *Bioinformatics*, Vol. 22, no. 9: 1152-1153.
Yarfitz, S. and Ketchell, D.S. 2000. A library-based bioinformatics services program. *Bulletin of the Medical Library Association*, Vol. 88, no. 1: 36-48.
Zhang, X., Shao, S., Stanley, H. E., and Havlin, S. 2014. Dynamic motifs in socio-economic networks. *Europhysics Letters (EPL)*, Vol. 108, no. 5: 58001.
Zhao, S.X., Rousseau, R. and Ye, F.Y. 2011. h-Degree as a basic measure in weighted networks. *Journal of Informetrics*, Vol. 5, no. 4: 668-677.
Zhao, S.X., and Ye, F.Y. 2012. Exploring the directed h-degree in directed weighted networks. *Journal of Informetrics*, Vol. 6, no. 4: 619-630.
Zhao, S.X., Zhang, P.L., Li, J., Tan, A.M., and Ye, F.Y. 2014. Abstracting core subnet of weighted networks based on link strengths. *Journal of the Association for Information Science and Technology*, Vol. 65, no. 5: 984-994.

APPENDIX

Network U has eight nodes. The maximum possible number of connected 3-graphs is $\binom{8}{3}$ = 56. Yet, not all of them are connected subgraphs in U: only 22 are, namely {A, B, C}, {A, B, D}, {A, B, E}, {A, B, G}, {A, C, D}, {A, C, E}, {A, D, E}, {A, D, H}, {A, E, F}, {A, E, G}, {A, E, H}, {B, C, D}, {B, C, G}, {B, E, G}, {B, F, G}, {C, D, H}, {D, E, H}, {D, F, H}, {E, F, G}, {E, F, H}, {E, G, H}, {F, G, H}. Among these 22 four are triangles. Hence N_{real} = 4/22 = 0.1818.