The Eurasia Proceedings of Educational & Social Sciences (EPESS)
ISSN: 2587-1730

The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2016

Volume 4, Pages 329-335

ICEMST 2016: International Conference on Education in Mathematics, Science & Technology

SCIENTIFIC COLLABORATION NETWORK OF ACADEMICIANS IN METU

İlker TÜRKER
Karabuk University, Department of Computer Engineering

Fatih GÖKÇE
M.Sc. Student in Karabuk University

Serhat Orkun TAN
Karabuk University, Vocational School

ABSTRACT: Scientific collaboration networks (SCNs) are web-like structures generated by collaborating patterns between scientists. Every co-authoring activity corresponds to a link between authors in such a network. Being successful prototypes of evolving complex networks, SCNs display the generic properties of self-organizing structures including social networks, in an aspect mirroring the scientific activities of the authors also. Collecting the scientific collaboration data of Middle East Technical University (METU) from Web of Science, we constructed a SCN spanning the years 1980 to 2015. Performing the network analysis procedures, we calculated the network parameters like average separation, average degree, degree distribution, average clustering coefficient etc. We outlined that the SCN of METU shows small-world and scale-free properties, also having high clustering between scientists.

Key words: Complex networks, scientific collaboration networks, scale-free networks, small-world.

INTRODUCTION

Studying scientific collaboration actions as a complex network provides broader understanding to the interactions between scientists (Newman, 2004). These networks are also good prototypes of self-organizing systems, with high resemblance in underlying organizing principles (Barabasi et al., 2002). The main motivations of scientific collaboration network (SCN) studies are both uncovering the evolving self-organizing principles in time, and also uncovering the bibliographic relations between scientists (Cavusoglu & Turker, 2013). By this view, the tools of network science empower the studies in bibliographic evolution (Wagner & Leydesdorff, 2005).

The advantages of studying SCNs as a complex network is first, they are evolving systems that are expanded by the addition of new authors, and also the establishment of new links between the existing authors. This evolution in time is successfully captured by the publication based storage of scientific papers in the databases like ISI Web of Science, Scopus etc. The second engaging property of these networks are that they are governed by the preferences of the nodes (authors), to publish a paper with another one, purely with their own choices.

Complex systems like SCNs are converted into networks by considering the authors as nodes, while the collaborations of these authors in a particular research paper define the links (Barabasi & Albert, 1999). SCNs display the universal properties like being *small-world* or *scale-free*. Small world property means, despite having numerous nodes, a short path from one node to another can be found and this situation is valid for the majority of the network (Watts & Strogatz, 1998). Scale-free stands for the network topology that the degree distribution is consistent with power-law decay (Albert & Barabasi, 2002; Amaral, Scala, Barthelemy, & Stanley, 2000; Clauset, Shalizi, & Newman, 2009; Virkar & Clauset, 2014). The degree of a node means the number of distinct neighbors it has. The distribution of node degrees in the whole network define the degree distribution, defining the interconnection characteristics of the network.

- Selection and peer-review under responsibility of the Organizing Committee of the conference
*Corresponding author: İlker TÜRKER-icemstoffice@gmail.com

Together with these universal properties observed in many real networks like biological systems, neural networks, computer networks, power grid networks, linguistic networks, social networks etc., SCNs display resemblance in the other network parameters like clustering and preferential attachment. Clustering means that your neighbors (or collaborators in a SCN) are also neighbors of each other. Preferential attachment means a node just attached to the network prefers to attach the more popular nodes (i.e. nodes having more neighbors, more degree), resulting power-law degree distribution.

The common properties mentioned above emerge in not only the SCN studies in distinct disciplines like engineering, mathematics, physics, surgery etc., but also in interdisciplinary studies (Barabasi & Albert, 1999; Barabasi et al., 2002; Cavusoglu & Turker, 2013, 2014; Luzar, Levnajic, Povh, & Perc, 2014; Newman, 2001a, 2001b, 2001c, 2004). Nationwide or international analysis are also performed by several scientists (Cavusoglu & Turker, 2013; Ferligoj, Kronegger, Mali, Snijders, & Doreian, 2015; Hoekman, Frenken, & Tijssen, 2010; Luzar et al., 2014; Ma, Fang, Pang, & Li, 2014; Perc, 2010).

This study mainly focuses on uncovering the network characteristics of scientists in Middle East Technical University (METU), as a leading university of Turkey in scientific productivity. For the timespan we investigate, METU is the second leading university with 21,663 publications, just coming after the most productive Hacettepe University with 24,093 publications. The 3, 4 and 5th most productive universities are İstanbul Technical University (19,305 publications), Ankara University (19,039 publications) and Gazi University (18,026 publications) respectively. On the other hand, METU can be evaluated as the most productive, among the universities that do not have department of medical sciences.

METHODS and RESULTS

We constructed the co-authorship network of METU researchers using ISI Web of Science Data, collected from the online search interface. We used a filtering constraint to achieve the publications addressed to Turkey, starting from the year 1980 to 2015. We also filtered the results in basis of institution to achieve the publications collaborated by METU researchers. We used the cumulative downloading utility of Web of Science to achieve the data in sets of 500 publications' data in each bin.

We constructed the *nodes* (composed of authors) and *edges* (composed of collaboration links) tables where each collaboration pair in a scientific paper results an undirected link between the two authors. We performed the network analysis in Gephi, a tool for complex network analysis and visualization (Bastian, Heymann, & Jacomy, 2009). The results of network analysis are presented and discussed in this section.

This study mainly focuses on uncovering the network characteristics of scientists in Middle East Technical University (METU), as a leading university of Turkey in scientific productivity. For the time span we investigate, METU is the second leading university with 21,663 publications, just coming after the most productive Hacettepe University with 24,093 publications. The 3, 4 and 5th most productive universities are İstanbul Technical University (19,305 publications), Ankara University (19,039 publications) and Gazi University (18,026 publications) respectively.

The linking procedure works as follows: If authors A, B and C write a paper together, we define links between A-B, A-C and also B-C. By the way, we achieved 15,413 authors (nodes) and 101,139 links between these authors. We also captured the time the paper is published, so we had the opportunity to study the time evolution of the network parameters.

We start with presenting the most productive researchers of our network in Table 1. The leading academicians have stunning degrees indicating their productive role in Turkish science literature.

Table 1. Top 30 most productive researchers of METU SCN.

Name	Degree	Weighted Degree
Toppare, Levent	351	1296
Turan, Rasit	277	588
Demir, Ayhan	273	516
Oguz, Temel	257	351
Hasirci, Vasif	250	486
Ozkar, Saim	207	470
Hasirci, Nesrin	206	388
Zeyrek, M.	192	349
Severcan, Feride	179	357
Gunduz, Ufuk	171	444

Korkusuz, Feza	170	296
Yilmaz, Akif	162	255
Sever, Ramazan	162	341
Kiziloglu, U.	161	383
Beklioglu, Meryem	158	237
Sahin, Ertan	156	262
Ozsoy, Emin	154	210
Molnar, J.	153	268
Yalciner, Ahmet Cevdet	145	200
Balci, Metin	144	365
Turker, Lemi	140	331
Weber, Gerhard-Wilhelm	138	215
Tuncel, Gurdal	138	265
Kence, Aykut	136	186
Van den Bleeken, Dieter	133	171
Goncuoglu, M. Cemal	124	203
Yucel, M.	124	349
Ataman, O. Yavuz	124	234
Lin, Shangchao	122	321
Kideys, Ahmet	121	189

We also evaluated the time evolving characteristics of the basic network metrics as in Table 2. These metrics are separately visualized in Figures 1 to 4.

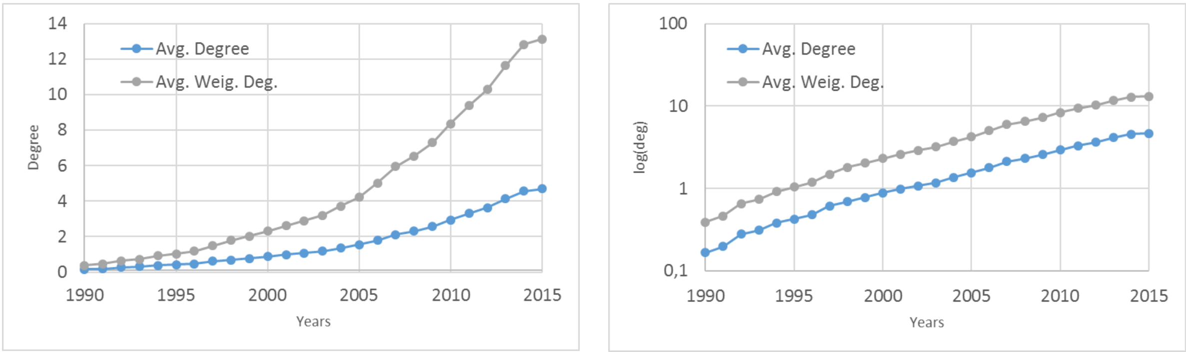
Table 2. Time Evolution of Basic Network Parameters.

	Avg. Degree	**Avg. Weig.Deg.**	**Diameter**	**Modularity**	**Avg. Clust.Coef.**	**Avg. Path Length**
1990	0,166	0,389	24	0,88	0,805	8,388
1991	0,197	0,464	25	0,887	0,801	7,719
1992	0,28	0,655	21	0,893	0,808	7,671
1993	0,311	0,74	21	0,896	0,8	7,159
1994	0,382	0,916	23	0,905	0,794	7,119
1995	0,427	1,044	21	0,911	0,8	6,922
1996	0,481	1,184	21	0,915	0,798	6,888
1997	0,615	1,487	23	0,91	0,798	7,014
1998	0,691	1,803	21	0,9	0,796	6,487
1999	0,775	2,031	21	0,901	0,799	6,454
2000	0,886	2,316	21	0,906	0,802	6,441
2001	0,99	2,617	22	0,909	0,805	6,422
2002	1,077	2,901	21	0,909	0,805	6,354
2003	1,177	3,192	20	0,904	0,805	6,145
2004	1,37	3,731	19	0,901	0,808	6,046
2005	1,554	4,232	18	0,901	0,811	5,88
2006	1,789	5,024	18	0,899	0,811	5,773
2007	2,125	5,955	19	0,896	0,811	5,654
2008	2,322	6,542	17	0,893	0,809	5,533
2009	2,58	7,294	16	0,89	0,808	5,476
2010	2,949	8,364	16	0,89	0,809	5,391
2011	3,315	9,38	16	0,886	0,81	5,297
2012	3,63	10,3	15	0,882	0,811	5,227
2013	4,143	11,645	15	0,881	0,813	5,154
2014	4,577	12,83	15	0,877	0,811	5,084
2015	4,689	13,12	15	0,876	0,811	5,066

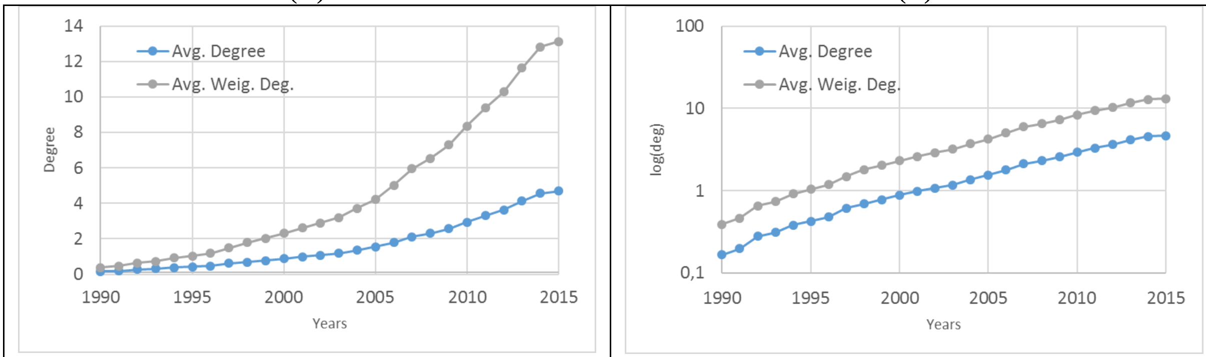
Average Degree

An evolving network grows with the addition of new nodes and links in time. Average degree ‹k› is a quantity that measures the number of links per author (Barabasi et al., 2002). It is a measure of how many collaborators the authors have in average, indicating the networks interconnectedness also.

We present the time dependency of average degree ‹k› in Fig.1, together with the weighted degree values that is effected by the duplicating connections to existing authors. The weighted degree values are 3 times the unweighted degrees, indicating that the existing links in the network are ~3 times repeated in average. The increasing trend of the average degrees together with the weighted degrees in Fig.1A indicate that the growth rate is faster than linear. The right side figure (Fig.1B) is the "semilogy" plot of the same data, having linear x and logarithmic y axis. A straight line in this axis combination indicates that the growth rate is exponential.



(A) (B)



(A) (B)

Figure 1. Average Degree and Weighted Degree in Years, (A) Linear Scale, (B) Log-Linear Scale

Average Clustering Coefficient and Modularity

The clustering coefficient evaluates how much a node's collaborators are willing to collaborate with each other. representing the probability that two of its collaborators wrote a paper together (Barabasi et al., 2002). Modularity is a measure for detecting the community structure in networks, indicating the rate of densely interconnected groups of nodes, having sparse connections to the other groups (Leicht & Newman, 2008). Both metrics have values in a range of 0 to 1. High clustering and modularity is an expected output of real networks.

Our scientific collaboration network displays high clustering and modularity features as presented in Fig. 2. The clustering seems to be nearly constant around 0.8. High clustering is an expected output of social based networks like SCNs (Newman, 2001c). Similarly, the real networks display modular structures, being divided into modules that are strongly interconnected but having rare connections to the other modules. The modularity values around 0.9 indicate that the network is highly modular. This indicates that the researchers tend to collaborate within their cliques with high modularity and clustering.

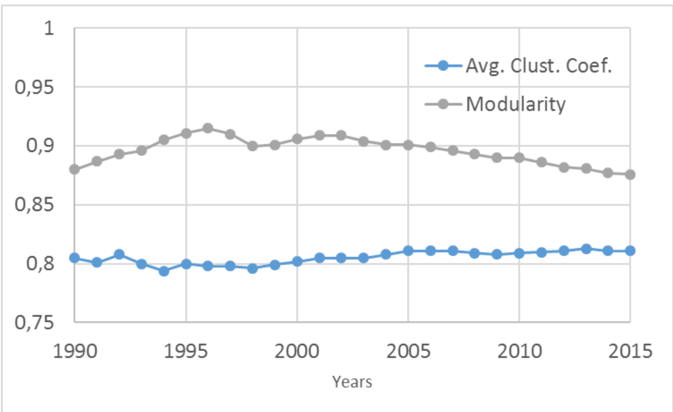


Figure 2. Clustering Coefficient and Modularity in Years

Average Path Length and Diameter

Average path length for a natwork is the mean number of edges along the shortest paths connecting the node pairs. Real networks display relatively short average path length values, that is also known as "small-world phenomenon" in the literature (Albert & Barabasi, 2002). The diameter of the network is the longest of these shortest paths, indicating that the maximum number of links from the most distant edges of the network. The evolution of these metrics are presented in Fig.3.

The average path length seems to converge about 5, consistent with the "six degrees of separation" phenomenon that is used for the "small-world" networks with small distances between nodes (Milgram, 1967; Watts & Strogatz, 1998). The resulting diameter 15 is also relatively small for a large-scale network having ~15 thousand nodes. All these outputs show that the network displays small-world properties.

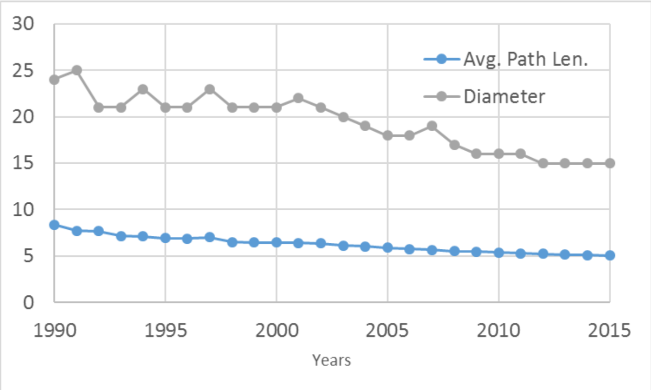
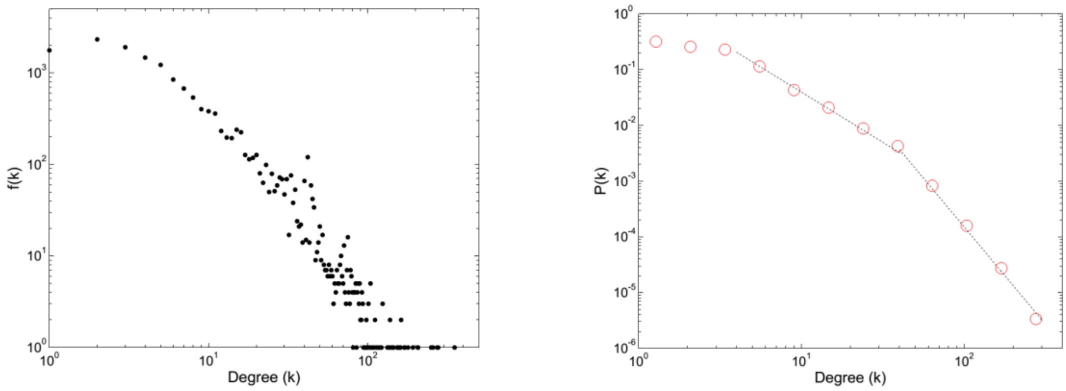


Figure 3. Average Path Length and Diameter in Years

Degree Distribution

Degree distribution *p(k)* is a probability distribution function of a randomly selected node to have k links. Networks having *p(k)* consistent with a power-law tail, are labelled as scale-free networks (Barabasi & Albert, 1999; Barabasi et al., 2002; Newman, 2003). Real networks are generally stated out to be scale-free. The degree distribution graph for the METU SCN is presented in Fig.4. The left side plot shows the frequency values for the degree occurrences, while the right side plot is log-binned and normalized probability distribution achieved from the same data.

The degree distribution in Fig.4B displays two discrete power-law consistent regimes having slopes of -1.8 and -3.5 respectively. These values are also consistent with the previous SCN studies (Barabasi et al., 2002; Cavusoglu & Turker, 2013, 2014; Newman, 2001a). The power-law consistency of degree distribution labels the network as scale-free. A scale-free network is generated with a critical ingredient, preferential attachment. It means that a node tends to connect with the other nodes with higher degrees more likely than the low degree ones (Vazquez, 2003). As a result, we can conclude that preferential attachment is the ingredient of the METU SCN also.



A B

Figure 4. (A) Degree Occurrence Frequencies (B) Degree Distribution plots for the METU SCN.

CONCLUSION

The METU SCN displays the generic properties of the SCNs analyzed in the recent studies. It is small-world, having small node-to-node distances compared to a random network. It is scale-free, having a power-law consistent degree distribution. It shows very high clustering and modularity metrics indicating that scientists tend to form micro cliques within their first neighbors, and also mid-sized cliques within their research groups. The growing rate is observed as exponential, with regard to the linear increasing trend of average degrees in log-linear plot.

The SCN of METU has promoted some super-nodes having degrees of 2 or 3 hundreds, as a generic ingredient of scale-free networks. These hub-like scientists also take the role of percolating the network and providing a superior robustness (Barabási, 2016).

According to the METU official website, the university employs about 791 faculty members (professors, associates professors etc.), 225 academic instructors and 1.273 research assistants (METU, 2016). According to this data, there are 2289 academicians in METU. The METU SCN consists of 15,413 authors, approximately 6.7 times the number of academicians. Taking the past academicians of METU into account, we can say that this rate should be ~5 times the academicians that have been employed in METU up to 2015. If validated with the real data of past academicians count, this number can be considered as the attractiveness coefficient of the METU researchers for the non-METU researchers that are introduced to the METU SCN by collaborating with them.

REFERENCES

Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47-97. doi: 10.1103/RevModPhys.74.47

Amaral, L. A. N., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America, 97*(21), 11149-11152. doi: 10.1073/pnas.200327197

Barabási, A. L. (2016). *Network Science*. Cambridge: Cambridge University Press.

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

Barabasi, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica a-Statistical Mechanics and Its Applications, 311*(3-4), 590-614. doi: 10.1016/s0378-4371(02)00736-7

Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: an open source software for exploring and manipulating networks*. Paper presented at the International AAAI Conference on Weblogs and Social Media, San Jose, California.

Cavusoglu, A., & Turker, I. (2013). Scientific collaboration network of Turkey. *Chaos Solitons & Fractals, 57*, 9-18. doi: 10.1016/j.chaos.2013.07.022

Cavusoglu, A., & Turker, I. (2014). Patterns of collaboration in four scientific disciplines of the Turkish collaboration network. *Physica a-Statistical Mechanics and Its Applications, 413*, 220-229. doi: 10.1016/j.physa.2014.06.069

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *Siam Review, 51*(4), 661-703. doi: 10.1137/070710111

Ferligoj, A., Kronegger, L., Mali, F., Snijders, T. A. B., & Doreian, P. (2015). Scientific collaboration dynamics in a national scientific system. *Scientometrics, 104*(3), 985-1012. doi: 10.1007/s11192-015-1585-7

Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy, 39*(5), 662-673. doi: 10.1016/j.respol.2010.01.012

Leicht, E. A., & Newman, M. E. J. (2008). Community structure in directed networks. *Physical Review Letters, 100*(11). doi: 10.1103/PhysRevLett.100.118703

Luzar, B., Levnajic, Z., Povh, J., & Perc, M. (2014). Community Structure and the Evolution of Interdisciplinarity in Slovenia's Scientific Collaboration Network. *Plos One, 9*(4), 5. doi: 10.1371/journal.pone.0094429

Ma, H. T., Fang, C. L., Pang, B., & Li, G. D. (2014). The Effect of Geographical Proximity on Scientific Cooperation among Chinese Cities from 1990 to 2010. *Plos One, 9*(11), 11. doi: 10.1371/journal.pone.0111705

METU. (2016). About METU. from http://www.metu.edu.tr/history

Milgram, S. (1967). The small-world problem. *Psychology Today, 1*(1), 61-67.

Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E, 64*(1), 8.

Newman, M. E. J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, 64*(1), 7.

Newman, M. E. J. (2001c). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2), 404-409. doi: 10.1073/pnas.021544898

Newman, M. E. J. (2003). The structure and function of complex networks. *Siam Review*, *45*(2), 167-256. doi: 10.1137/s003614450342480

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 5200-5205. doi: 10.1073/pnas.0307545100

Perc, M. (2010). Growth and structure of Slovenia's scientific collaboration network. *Journal of Informetrics, 4*(4), 475-482. doi: 10.1016/j.joi.2010.04.003

Vazquez, A. (2003). Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Physical Review E*, *67*(5), 15. doi: 10.1103/PhysRevE.67.056104

Virkar, Y., & Clauset, A. (2014). Power-Law Distributions In Binned Empirical Data. *Annals of Applied Statistics*, *8*(1), 89-119. doi: 10.1214/13-aoas710

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy, 34*(10), 1608-1618. doi: 10.1016/j.respol.2005.08.002

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440-442. doi: 10.1038/30918