The Eurasia Proceedings of Science, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2018

Volume 2, Pages 209-215

ICRES 2018: International Conference on Research in Education and Science

Modeling Clustered Scale-free Networks by Applying Various Preferential Attachment Patterns

Gokhan KUTLUANA
Bartın University

Ilker TURKER
Karabük University

Abstract: Preferential attachment phenomenon is a key factor providing scale-free behavior in complex networks. In this study, we introduced various preferential attachment patterns applied in a growing Barabasi-Albert network, denoted by a factor α. We first generated networks under constant preferential attachment levels from 0 to 2, where 1 stands for linear preferential attachment. Then we performed network simulations under uniformly distributed random α condition, within the interval [0,2]. Although mean α is 1 for this setup, generated networks displayed greater clustering together with lower modularity and separation values compared to the setup with α=1. We also performed similar network generation procedures with various distribution functions applied for α, each resulting random levels of preferential attachment. We achieved networks with power-law consistent degree distributions with γ coefficients between 2 and 3, together with improved clustering coefficients up to ~0.3. As a result, scale-free network topologies featuring greater clustering levels compared to pure Barabasi-Albert model are achieved.

Keywords: Complex network modeling, Preferential attachment, Scale-free networks, Clustering coefficient

Introduction

Real networks diverge from random networks with their degree sequences and clustering property. Most real networks have power-law consistent degree distributions labeling them as *scale-free* (Albert & Barabási, 2002; Barabási & Albert, 1999). Scale-free networks promote generation of a small number of hubs together with numerous low-degree nodes, where the degree sequences of all nodes are consistent with a linear decaying character in a log-log scale of degree distribution (Clauset, Shalizi, & Newman, 2009; Newman, 2003). This distribution is mostly governed by the preferential attachment phenomenon, a key ingredient of most current network models that highly connected nodes increase their connectivity faster than their less connected peers (Jeong, Néda, & Barabási, 2003). In networks with preferential attachment, new nodes prefer connecting to more connected nodes instead of less connected ones. This generic mechanism has significant roles in real networks and have been subject of many studies of network analysis and modeling (Abbasi, Hossain, & Leydesdorff, 2012; Dereich & Mörters, 2009; Johnson, Faraj, & Kudaravalli, 2014; Milojević, 2010; Poncela, Gómez-Gardenes, Floría, Sánchez, & Moreno, 2008).

Preferential attachment is said to be linear, if the connectivity of a node is linearly dependent to its degree. In some occasions, nodes in real networks display attachment levels with non-linear dependence to node degree. These behaviors are labeled as sub-linear or super-linear preferential attachment. The critical level of this attachment level is abstracted with an α parameter that is equal to 1 for linear preferential attachment, less than 1 for sub-linear and greater than 1 for super-linear attachment levels (Barabási, 2016). Networks with sub-linear preferential attachment levels display stretched exponential degree distribution, meaning fewer and smaller hubs compared to a scale-free network, and a concave degree distribution in a log-log scale. On the other hand, networks with super-linear attachment display convex log-log degree distributions, which lead to a hub-and-spoke topology as a result of “a winner-takes-all” dynamics (Barabási, 2016; Dereich & Mörters, 2011).

Clustering is another key ingredient of real networks, captured with a coefficient indicating in what rate the neighbors of a given node are also neighbors of each other (Barabási, 2016; Newman, 2003). This coefficient is averaged over all nodes to present the level of average clustering of a network. Many modeling studies are conducted to capture both preferential attachment and clustering properties of real networks. These studies employ several mechanisms like triad-formation (Alstott, Klymko, Pyzza, & Radcliffe, 2016; Holme & Kim, 2002; Kim & Diesner, 2017), improving clustering of networks with arbitrary degree distributions (Bansal, Khandelwal, & Meyers, 2009; Colomer-de-Simon & Boguná, 2012; Herrero, 2015) and promoting connections between spatially close nodes (Manna & Sen, 2002; Türker, 2018; Xie, Ouyang, & Li, 2016; Xulvi-Brunet & Sokolov, 2002).

In this study, we aimed to achieve clustered scale-free network topologies by applying various preferential attachment levels including α parameters of constant values below and above 1, uniform distributed random α values, α values generated from some mathematical expressions like sinusoidal or sigmoid functions etc. The results of the corresponding models are presented in the next section. Although no strategy is employed for tuning clustering property, the results demonstrate that scale-free networks with improved clustering according to Barabasi-Albert network are achieved.

Method

We generated growing network models, based on Barabasi-Albert (BA) scale-free model (Albert & Barabási, 2002). The computational fashion followed to realize a BA network is illustrated in Fig. 1. For a new node just joining the network, connecting probability to an old node is directly proportional with the number of occurrences of the node ID in the edges array. This provides realization of linear preferential attachment phenomenon in a numerical fashion.

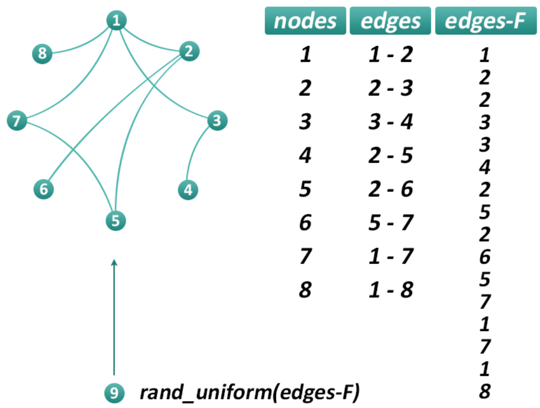


Figure 1. The computational method followed to achieve a BA network is illustrated. The new node (9) just joining the network chooses a node to connect from the flattened edges array (edges-F), with an index randomly generated from a uniform distribution. By the way, nodes have attachment probabilities proportional with their current degrees

Tuning Preferential Attachment

In the above-mentioned setup, the likelihood to connect to a node depends on that node's degree k. The functional form of preferential attachment can be approximated with Eq. 1.

$$\Pi(k) \sim k^{\alpha} \tag{1}$$

For networks with linear preferential attachment $\alpha = 1$, corresponding to linear dependence of Π with k. For any $\alpha > 0$, new nodes tend to make connections to more connected nodes over less connected ones. For $\alpha < 1$ (sub-linear preferential attachment) this bias is weak and insufficient to provide a pure scale-free degree distribution, rather resulting a stretched exponential distribution. On the other hand, for $\alpha > 1$, connecting to more connected nodes is promoted, resulting a structure that rich get richer than in a network with linear preferential attachment (Barabási, 2016). This behavior, labeled as super-linear attachment, exhibits a hub-and-spoke topology and a convex degree distribution in log-log scale.

We tuned the generated network models with this α parameter such that, the occurrence count of the ID of a node with degree k in the edges-F array is equal to k^{α}. For linear attachment, the ID of a node with degree 2 is repeated $2^1 = 2$ times in this array, corresponding to linear attachment as illustrated in Fig. 1. For α values resulting non-integer repeat counts, the result is rounded to the nearest integer. For instance, if $\alpha = 1.4$ and the degree of a node is 2, this count will be $2^{1.4} \sim 2.63$, therefore the node ID will be repeated 3 times in the edges-F array.

Results and Discussion

We first generated networks by setting constant α parameters within the interval [0-2] by incrementing 0.2 for each network setup. This procedure is applied to two sets of networks where the first is constructed with 1 links assigned for each new node and the second is constructed with 5 links assigned for each new node. This count of edges corresponds to the parameter *m* described in BA model in Ref. (Albert & Barabási, 2002).

We present the basic network parameters in Table 1 and 2 as follows. As seen from the tables, average path length displays monotone decaying trend for both setup. Modularity measure remains almost constant for α values up to 1.4, showing a steep decay for greater α values. This indicates that super-linear preferential attachment, resulting a hub-and-spoke topology avoids nodes to organize into modules. For the similar α interval, average clustering coefficient increases dramatically. As a result, networks emerge to get more clustered but less modular for super-linear attachment region.

Table 1. Results for networks generated with *m*=1 links per step and different preferential attachment levels (α). The last column corresponds to random α applied at each step of the growing network, from a uniform distribution between 0 and 2. Clustering coefficient results 0 for m=1 setup.

α	**0**	**0,2**	**0,4**	**0,6**	**0,8**	**1**	**1,2**	**1,4**	**1,6**	**1,8**	**2**	**Rnd**
Avg. Clustering C.	-	-	-	-	-	**-**	-	-	-	-	-	**-**
Avg. Path Length	10,76	10,56	9,36	8,52	8,13	**7,21**	5,66	3,05	2,22	2,03	2,02	**6,35**
Modularity	0,934	0,934	0,934	0,933	0,929	**0,925**	0,904	0,534	0,155	0,031	0,024	**0,892**

Table 2. Results for networks generated with *m*=5 links per step and different preferential attachment levels (α). The last column corresponds to random α applied at each step of the growing network, from a uniform distribution between 0 and 2.

α	**0**	**0,2**	**0,4**	**0,6**	**0,8**	**1**	**1,2**	**1,4**	**1,6**	**1,8**	**2**	**Rnd**
Avg. Clustering C.	0,015	0,017	0,018	0,019	0,026	**0,043**	0,085	0,211	0,603	0,759	0,724	**0,077**
Avg. Path Length	3,24	3,19	3,19	3,15	3,087	**2,964**	2,77	2,5	2,01	2	1,99	**2,88**
Modularity	0,282	0,285	0,29	0,288	0,278	**0,28**	0,263	0,25	0,193	0,152	0,106	**0,272**

A noteworthy output of these simulations is that, selecting α parameter from a uniform random distribution between 0 and 2 (with a mean of 1) results networks with greater clustering (~2 times) compared to networks with linear preferential attachment ($\alpha = 1$). Although they have the same expected α parameter, picking random α at each step emerges as a key factor for improving clustering in scale-free networks. These results also indicate that networks with improved clustering are observed for α values typically greater than 1.

We also present the degree distributions for these networks in Fig. 2 to consult the power-law consistencies. These plots indicate that, degree exponent decreases with increasing α parameter, since the super-linear preferential attachment results convex degree distributions. Visual inspection says that power-law fitting is not a suitable choice for α values greater than 1.4. On the other hand, degree distributions for both α=1 and α= rand(0-2) scenarios result in power-law consistent distributions with exponents close to 2.7, whereas the network with randomized α setup with mean 1 exhibits approximately twice clustering coefficient compared to fixed α=1. This result emerges as a phenomenon that random levels of preferential attachment plays a key role in improving clustering.

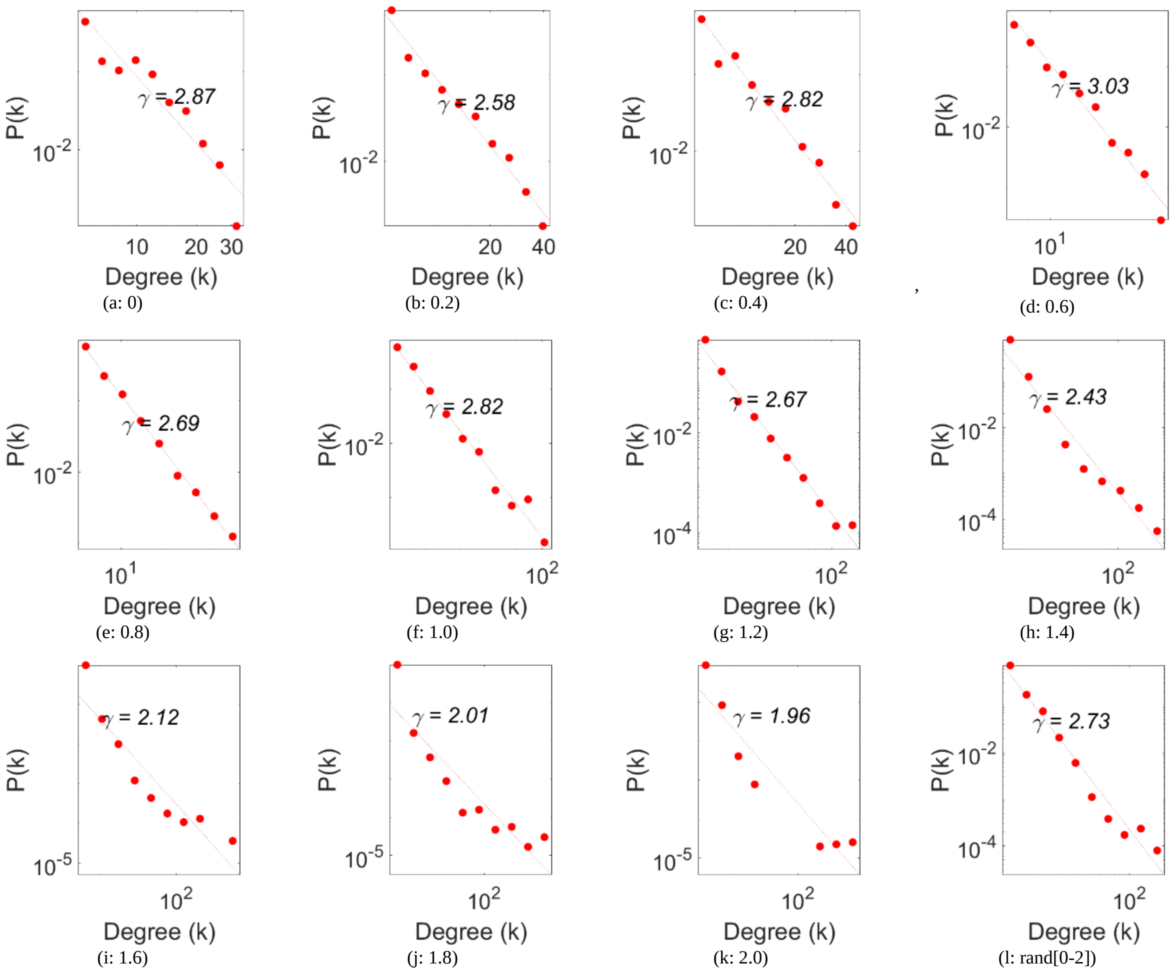
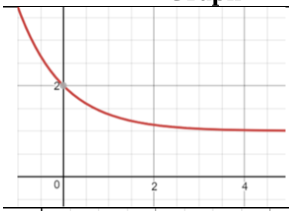
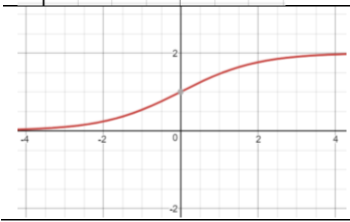


Figure 2. Log binned degree distributions for the networks generated with *m*=5 links per step and different preferential attachment levels (α), gradually increased from 0 (a), to 2 (k) with increment 0.2 at each plot. The last plot (l) corresponds to random α applied at each step of the growing network, from a uniform distribution between 0 and 2. Both axes are logarithmically scaled. Power-law fits with degree exponents are presented for each plot, after least-squared (LS) fitting performed.

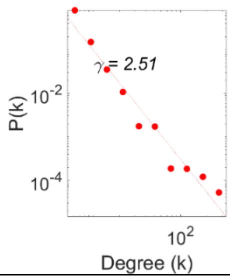
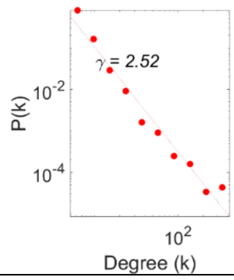
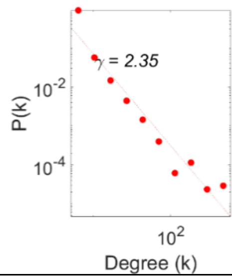
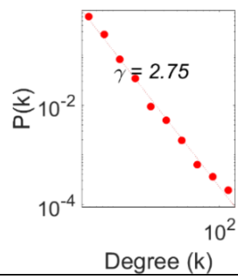
To further investigate the advantage of random α on clustering mentioned above, we applied four different patterns of random α generation procedures, as given in Table 3. As seen from the function plots, the first 3 functions are expected to generate α parameters with mean values greater than 1 (which will result in improved clustering). The last function is expected to produce symmetric α values around 1, but also promote α values rather closer to 0 and 2. All network generation procedures are performed with *m*=5 links per step, and 1000 steps of simulation.

Table 3. Functions applied for random α generation procedures. Uniform random numbers are generated within the interval given in the rightmost column, and the corresponding y values are used as random α at each step of a growing network.

Graph	Function	Applied Interval
	$y = 1 + e^{-x}$	[0,3]
	$y = 2 * sin(x)$	[0,π]
	$y = 1.25 + e^{-x}sin(x)$	[0,π]
	$y = \frac{2}{1 + e^{-x}}$	[-4,4]

Applying random level of preferential attachment at each step of growing network, we generated four network topologies. Generating uniform random numbers within the given intervals for each function, the generated *y* values are constituted for α at each step of edge generation. We present the results of network analysis for each setup in Table 4, together with the resulting degree distributions.

Table 4. Network parameters for the four different patterns of random preferential attachment. Corresponding degree distributions are also presented at the end of each column.

	$\boldsymbol{y = 1 + e^{-x}}$	$\boldsymbol{y = 2 * \sin(x)}$	$\boldsymbol{y = 1.25 + e^{-x}}$	$\boldsymbol{y = \frac{2}{1 + e^{-x}}}$
Avg. Clust. C.	0,1638	0,1856	0,3039	**0,12**
Avg. Path Length	2,5885	2,6250	2,3189	**2,8**
Degree Dist.				

As seen in Table 4, the first 3 functions promoting α values greater than 1, result in greater clustering and low separation, while the last function (a type of sigmoid function) generating symmetric α values around 1 yields less clustering but better consistency with power-law behavior. Although the first 3 distribution plots show convex tendencies, they seem to be acceptable to be labeled as scale-free. The most deviated one from pure power-law consistency is the third one, which dominantly produces greater α values from 1.25. The last model with symmetric α distribution around 1 deviates from the uniformly generated α model in Table 2, by promoting α values rather close to 0 and 2. As a result, both networks have similar average path length values, while power-law consistencies and degree coefficients also seem to be identical. Additionally, the latter one seems to be more successful in improving clustering property (0.077 vs. 0.12). This result indicates that, beyond generating random α parameters symmetric around 1, promoting the values close to 0 and 2 also has a significant impact on improving clustering.

Conclusion

We applied various levels of fixed preferential attachment levels, together with random number generation processes at each step of wiring in growing BA-like networks. We first outlined that clustering coefficient and modularity measures are inversely proportional with the preferential attachment level. Picking random levels of preferential attachment (α) at each wiring procedure emerges as a key factor for improving clustering, while it preserves average path length. Picking α from a uniform distribution between 0 and 2 results in approximately improving clustering twice, while this improvement is three times for a sigmoid function symmetric around 1. We conclude these results that fixed level of preferential attachment may emerge as an inhibiting factor on clustering, in growing networks.

References

Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics, 6*(3), 403-412.

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

Alstott, J., Klymko, C., Pyzza, P. B., & Radcliffe, M. (2016). Local rewiring algorithms to increase clustering and grow a small world. *arXiv preprint arXiv:1608.02883*.

Bansal, S., Khandelwal, S., & Meyers, L. A. (2009). Exploring biological network structure with clustered random networks. *BMC bioinformatics, 10*(1), 405.

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

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

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

Colomer-de-Simon, P., & Boguná, M. (2012). Clustering of random scale-free networks. *Physical Review E, 86*(2), 026120.

Dereich, S., & Mörters, P. (2009). Random networks with sublinear preferential attachment: degree evolutions. *Electronic Journal of Probability, 14*, 1222-1267.

Dereich, S., & Mörters, P. (2011). Random networks with concave preferential attachment rule. *Jahresbericht der Deutschen Mathematiker-Vereinigung, 113*(1), 21-40.

Herrero, C. P. (2015). Ising model in clustered scale-free networks. *Physical Review E, 91*(5), 052812.

Holme, P., & Kim, B. J. (2002). Growing scale-free networks with tunable clustering. *Physical Review E, 65*(2). doi: 10.1103/PhysRevE.65.026107

Jeong, H., Néda, Z., & Barabási, A.-L. (2003). Measuring preferential attachment in evolving networks. *EPL (Europhysics Letters), 61*(4), 567.

Johnson, S. L., Faraj, S., & Kudaravalli, S. (2014). Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment. *Mis Quarterly, 38*(3), 795-808.

Kim, J., & Diesner, J. (2017). Over-time measurement of triadic closure in coauthorship networks. *Social Network Analysis and Mining, 7*(1), 9. doi: 10.1007/s13278-017-0428-3

Manna, S. S., & Sen, P. (2002). Modulated scale-free network in Euclidean space. *Physical Review E, 66*(6), 066114.

Milojević, S. (2010). Modes of collaboration in modern science: Beyond power laws and preferential attachment. *Journal of the Association for Information Science and Technology, 61*(7), 1410-1423.

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

Poncela, J., Gómez-Gardenes, J., Floría, L. M., Sánchez, A., & Moreno, Y. (2008). Complex cooperative networks from evolutionary preferential attachment. *Plos One, 3*(6), e2449.

Türker, İ. (2018). Generating clustered scale-free networks using Poisson based localization of edges. *Physica A: Statistical Mechanics and its Applications, 497*, 72-85. doi: https://doi.org/10.1016/j.physa.2018.01.009

Xie, Z., Ouyang, Z., & Li, J. (2016). A geometric graph model for coauthorship networks. *Journal of Informetrics*, *10*(1), 299-311.

Xulvi-Brunet, R., & Sokolov, I. M. (2002). Evolving networks with disadvantaged long-range connections. *Physical Review E, 66*(2), 026118.

Author Information

Gokhan Kutluana	**Ilker Turker**
Bartın University	Karabük University
Health Services Vocational School	Faculty of Engineering, Computer Engineering Dept.
	Contact e-mail: *iturker@karabuk.edu.tr*