2014 American Control Conference (ACC)
June 4-6, 2014. Portland, Oregon, USA

Characterizing the Relationship Between Degree Distributions and Community Structures

Pablo Moriano
School of Informatics and Computing
Indiana University
Bloomington, IN, USA
moriano@ieee.org

Jorge Finke
Department of Electrical Engineering and Computer Science
Pontificia Universidad Javeriana
Santiago de Cali, Colombia
finke@ieee.org

***Abstract*— Extended power laws and inhomogeneous connections are structural patterns often found in empirical networks. Mechanisms based on the formation of triads are able to explain the power law behavior of the degree distribution of such networks. The proposed model introduces a two-step mechanism of attachment and triad formation that illustrates how preferential linkage plays an important role in shaping the inhomogeneity of connections and the division of the network into groups of nodes (i.e., the growth of community structures). In particular, we identify conditions under which the scaling exponent of the power law correlates to a widely-used modularity measure of non-overlapping communities. Our analytical results characterize the asymptotic behavior of both the scaling exponent and the modularity, as a function of the strength with which nodes with similar characteristics tend to link to each other.**

I. INTRODUCTION

Dynamic network models describe the evolution of structural properties of interconnected systems. Large networks arise by the gradual addition of new members to an existing set of nodes, often with a strong preference to attach to nodes with similar characteristics (e.g., due to homophily). The proximity between nodes and the neighbors of their neighbors further shapes how nodes link to each other over time. The premise that there are well-defined trends underlying the emergence of particular network structures lies at the heart of recent efforts to identify what principles explain observable patterns found in empirical data.

There are striking similarities in the structure among empirical networks that are otherwise quite different in their nature. Understanding the relationships between structural properties requires the development of models that characterize each property and its evolution (e.g., to explain the processes from which they may emerge). Highly clustered networks, for example, are a particular class of models where the networks share many close instead of distant connections. They include technological (e.g., the Internet [1], [2]), information (e.g., the World Wide Web [3]), social (e.g., collaborations between authors [4]), and biological systems (e.g., metabolic networks [5]).

All of these networks follow a single power law for a large part of their probability degree distributions [6]-[8]. Two well-known principles which lead to power law behavior are preferential attachment [9] and triad formation [10]. Resting solely on the assumption that a new node is more likely to connect to nodes with a higher degree, mechanisms of preferential attachment are not able to capture the clustering properties of large networks. Triad formation mechanisms, on the other hand, induce clustering by presuming that after a new node randomly attaches to the network, there exists a strong preference to link to other nodes in its neighborhood. The work in [11], [12] deduces analytical expressions for both the power law exponent and the global and local clustering coefficients of networks models that evolve based on triad formation. Moreover, it shows that under proper conditions, the nodes of the network may divide into groups (communities) with a higher density of edges within and a sparser density between them [13].

Most of the work on network communities has focused on (i) identifying and quantifying the strength of (overlapping and non-overlapping) communities [14]-[16]; and (ii) developing dynamic models that explain their formation [17]-[21]. The model in [17] illustrates the formation of communities in social networks based on the desire of individuals to differentiate themselves from the average (i.e., a seceder model). The work in [18] proposes a bipartite model that leads to growing networks with well-defined communities. In [19], community formation is studied through a model based on social distance (i.e., interaction is based on a decreasing function that quantifies the degree of closeness of an individual towards others). The models in [20] and [21] rest on refinements of preferential attachment that generate community structures based on inner- and inter-community measures.

While all the above models [11]-[21] illustrate a wide range of mechanisms that drive the formation of communities, analytical results that characterize the relationship of community structures to other structural properties and their evolution are lacking (a noteworthy exception is the work in [22], which illustrates that for a certain parameter range the degree distribution and the size of emerging communities both satisfy power laws). This paper introduces a two-step mechanism of attachment and triad formation that captures the interdependence between degree distributions and community structures. To our knowledge, the proposed model is novel in that it generates directed networks in which triad formation induces (i) extended power law behavior with

978-1-4799-3271-9/$31.00 ©2014 AACC

a scaling exponent $\alpha \geq 3$ (i.e., a degree distribution that follows a power law for nodes with a high degree); and (ii) inhomogeneity of connections (i.e., a distinct community structure with a modularity $-0.5 \leq Q \leq 0.5$). Our work extends the results in [10] by verifying analytically that the proposed mechanism of attachment and triad formation is sufficient for the emergence of highly clustered power law networks.

The remaining sections are organized as follows: Section 2 introduces a model that captures the connectivity dynamics of a growing network. Theorem 1 in Section 3 shows that the asymptotic behavior of the in-degree distribution follows a power law distribution $p_k \sim k^{-\alpha}$ above a certain threshold ε and an exponential distribution $p_k \sim e^{-\lambda k}$ otherwise. It presents analytical results for the values of α, λ, and ε. Theorem 2 yields an analytical expression for the value of Q for networks that divide into two communities. Section 4 illustrates a particular criterion for defining the likelihood with which nodes with similar characteristics may link to each other during triad formation. Section 5 presents Monte Carlo simulations that capture the effect of the combined mechanism on the scaling behavior and the modularity of the network. Finally, Section 6 draws some conclusions and future research directions.

II. A NETWORK FORMATION MODEL

Let $\mathcal{H}_t = \{1, \ldots, N_t\}$ be a finite set of interconnected nodes at time index t. The set $\mathcal{A}_t = \{(i,j) : i,j \in \mathcal{H}_t\}$ represents the relationships between nodes, where (i,j) indicates that there exists a directed edge between nodes i and j. Let $\mathcal{G}_t = (\mathcal{H}_t, \mathcal{A}_t)$ represent the network at time index t. Assume that nodes tend to attract or repel each other based on the assessment of a particular trait. In other words, we view the network $\mathcal{G}_t$ as being composed of two types of nodes denoted by $\delta \in \{1,2\}$ where δ_i specifies the type of node i. Two nodes are of the same type if they share common characteristics (e.g., beliefs, values, status). Let $q_i(t) = \{j \in \mathcal{H}_t : (j,i) \in \mathcal{A}_t\}$ represent all nodes that link to node i at time t (i.e., its incoming neighbors). For any node $i \in \mathcal{H}_t$, let $k_i(t) = |q_i(t)|$ represent the in-degree of node i.

A. Node attachment

Every time index t a new node is added to the network and it attaches to m different nodes. The type δ_j for the new node $j \notin \mathcal{H}_{t-1}$ takes the value of 1 with probability $\frac{1}{2}$. When node j attaches to the network, it connects to a node $j' \in \mathcal{H}_{t-1}$ of the same type ($\delta_j = \delta_{j'}$) with probability p_r (and with probability $1 - p_r$ to a node of a different type).

B. Triad formation

Conditions for the formation of triad junctions are similar to [10]. When node $j \notin \mathcal{H}_{t-1}$ attaches to some node $j' \in \mathcal{H}_{t-1}$, it may also establish an additional link to one of the outgoing neighbors of node j'. If $j \in q_{j'}(t)$ and $j' \in q_i(t)$ for some node i, node j links to node i with probability $x_i(t)$, forming a triad whenever the event occurs. The probability $x_i(t)$ that node j establishes an additional edge to node i is influenced by δ_j and δ_i. A multivariate random variable X_t^δ with a positive expected value $p_t^\delta = E[X_t^\delta] = f(\sigma_1, \cdots, \sigma_s) d\sigma_1 \cdots d\sigma_s$ captures the likelihood of establishing a link between nodes j and i, when $\sigma_1, \cdots, \sigma_s$ are independent factors that influence the process. Note that if the set of outgoing neighbors of node j' is a subset of the set of outgoing neighbors of node j then there is no possibility of forming triads. The process of triad formation repeats for every edge established by a newly added node (m times) before another node attaches to the network. Let $X^\delta = \{X_t^\delta\}$ be the random process associated to triad formation with stationary mean $p_\Delta > 0$. To ensure that the two-step mechanism is well-defined we require the following assumption.

Assumption 1: The initial network $\mathcal{G}_0$ is connected ($q_i(0) > 0$ for all $i \in \mathcal{H}_0$) and every node has at least m neighbors ($N_0 \geq m$). Moreover, if $p_\Delta = 1$ then each node in $\mathcal{H}_0$ must have at least one outgoing neighbor.

III. ANALYSIS

The relationship between the in-degree distribution and the modularity of the network rests on the following characterizations of the asymptotic behavior of the network (the proofs of Theorems 1 and 2 are presented in the Appendix).

Theorem 1 *(in-degree distribution): For all networks $\mathcal{G}_0$ that satisfy Assumption* 1*, the in-degree distribution p_k of $\mathcal{G}_t$ follows an extended power law as $t \to \infty$. The scaling exponent $\alpha = 1 + \frac{1}{\tau}$ where $\tau = \left(\frac{p_r p_\Delta}{1+p_\Delta} + \frac{(1-p_r)(1-p_\Delta)}{2-p_\Delta}\right)$ and the exponential exponent $\lambda = \frac{\alpha}{\alpha-1}$, with threshold $\varepsilon = (\alpha-1)\,m$.*

Theorem 1 implies that, as the network grows, the scaling exponent of the in-degree distribution depends solely on the preference to establish links between nodes of the same type during node attachment and triad formation. The resulting distribution follows a strict power law for nodes with a degree greater than $(\alpha - 1)\,m$.

Theorem 2 *(community modularity): Given a network $\mathcal{G}_0$ that satisfies Assumption* 1*, the modularity Q of $\mathcal{G}_t$ tends to a constant value that depends only on p_r and p_Δ as $t \to \infty$.*

Theorem 2 implies that the modularity measure of communities reaches a stationary value. As is the case for the degree distribution, the community modularity depends on the probability to connect to nodes of the same type.

IV. AN EXAMPLE

Similar to the work in [23], [24] we let the probability of establishing additional links due to triad formation be

$$x_i(t) = \begin{cases} p_\Delta - \frac{c}{uk_i}, & \text{if } \delta_j = \delta_i \\ (1-p_\Delta) - \frac{c}{uk_i}, & \text{if } \delta_j \neq \delta_i \end{cases} \quad (1)$$

where u captures the compatibility between nodes and is chosen from a uniformly random distribution with support on $[0, 1]$ (in other words the random variable X_t^δ takes values $x_i(t)$). The parameter c, $0 < c < u$, represents the cost of establishing additional links and $0 \leq p_\Delta \leq 1$. The expected value of X_t^δ at time t is given by

$$p_t^\delta = \begin{cases} \int_0^\infty \int_0^1 \left(p_\Delta - \frac{c}{uk_i(t)}\right) p_u p_k \; du \; dk_i, & \text{if } \delta_j = \delta_i \\ \int_0^\infty \int_0^1 \left((1 - p_\Delta) - \frac{c}{uk_i(t)}\right) p_u p_k \; du \; dk_i, & \text{if } \delta_j \neq \delta_i \end{cases} \tag{2}$$

where $p_u = \frac{1}{u}$ and p_k is the probability distribution of $k_i(t)$. If $\delta_j = \delta_i$ the process of triad formation has a stationary mean p_Δ (i.e., equation (2) converges to p_Δ). Otherwise, if $\delta_j \neq \delta_i$ it has stationary mean $1 - p_\Delta$.

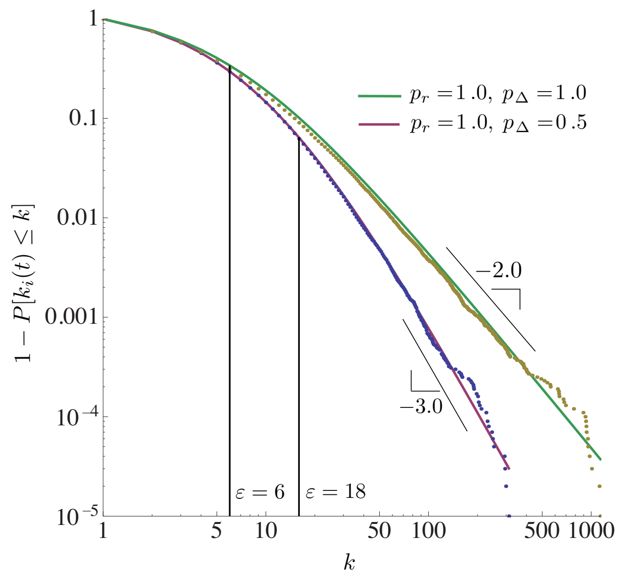
V. SIMULATIONS

To gain insight into the network formation dynamics, let $N_0 = 20$, $c = 0.1u$ and consider $x_i(t)$ as defined in (1). Figure 1(a) shows the in-degree distribution of the network at $t = 10^5$ for $p_r = 1.0$ and different values of p_Δ. For nodes with a low degree, the complementary cumulative degree distribution follows an exponential form with $\lambda = \frac{3}{2}$ (for $p_\Delta = 0.5$, $\alpha = 4$ and $\lambda = \frac{4}{3}$). In particular, the equation $\varepsilon = (\alpha - 1)m$ characterizes the transition from an exponential to a power law distribution.

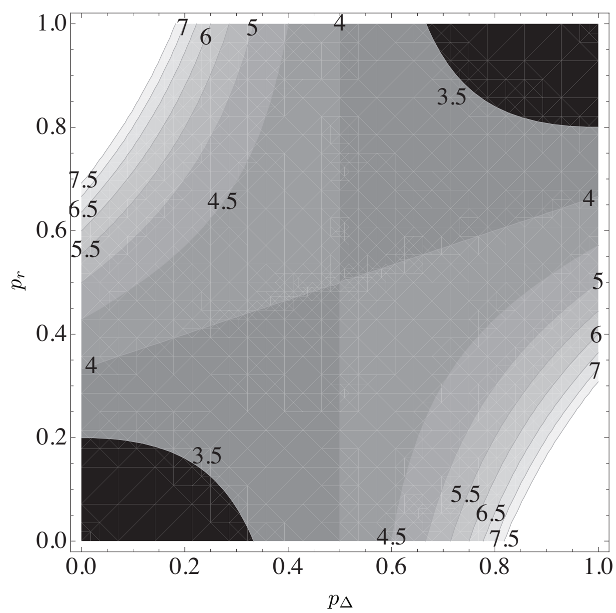
Figure 1(b) illustrates the dependence of the scaling exponent on p_r and p_Δ. Note that a very strong or a very weak tendency to link similar nodes for both node attachment and triad formation generates networks with a scaling exponent below 3.5. On the other hand, when one probability is large but the other small, the scaling exponent of the resulting network is above 4.5.

Figure 2(a) shows the evolution of the modularity of the network for different values of p_r and p_Δ. The dots represent simulation results; the solid lines are the theoretical prediction derived from Theorem 2. As the network grows the modularity tends to a stationary value that depends solely on p_r and p_Δ for any initial network. Figure 2(b) shows the value of the modularity Q for different values of p_r and p_Δ. Note that if $p_r = 1$ then $Q = 0.5$. Not surprisingly, p_Δ has no impact on the modularity during random attachment when nodes only connect to other nodes of the same type (see Case 1 in the proof of Theorem 2). Overall, fig. 2(b) suggests that p_r has a stronger impact on the modularity than p_Δ. However, note that for high values of p_r, increasing p_Δ can still have a noticeable effect. The highlighted region illustrates the parameter regime where the network has modularity $0.3 \leq Q \leq 0.5$ as often found in empirical data [16].

Next, fig. 3 shows the modularity of the network for different values of p_r when $p_\Delta = 1$. It confirms that increasing the likelihood to link similar nodes during random attachment strongly impacts the modularity of the resulting network even if during triad formation nodes always connect to nodes of the same type (see Case 2 in the proof of Theorem 2).



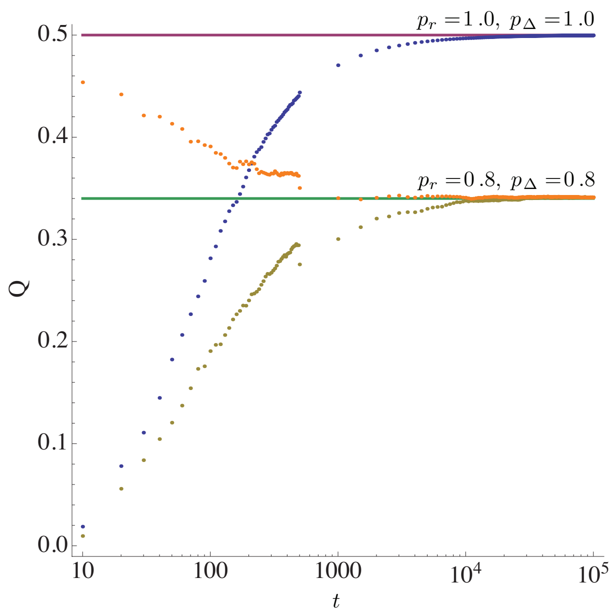
(a) Complementary cumulative in-degree distribution for $m = 3$.



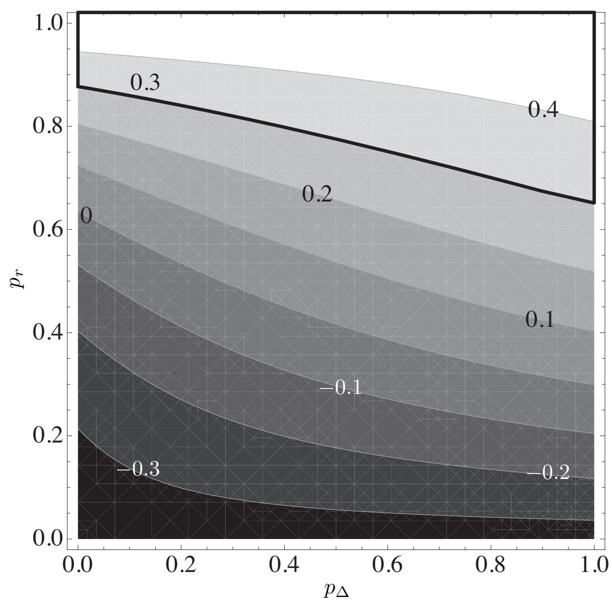
(b) Scaling exponent.

Fig. 1: (a) Complementary cumulative in-degree distribution p_k. The solid lines are the theoretical predictions based on (5); the dots represent distributions from simulations for different values of p_r and p_Δ; and (b) Scaling exponent for different values of p_r and p_Δ.

Finally, fig. 4 characterizes the relationship between modularity Q and the scaling exponent α. Note that for certain parameters there is a strong correlation, but it may be positive or negative depending on their values. Table I describes the range of p_r values used to generate the relationships in fig. 4 for a constant preferential linkage during triad formation. For negative correlations (i.e., for lines $l_4 - l_8$) the model produces similar results as the empirical measures in [25].



(a) Evolution of the modularity of the network.



(b) Resulting modularity of the network.

Fig. 2: (a) Modularity for different values of p_r and p_Δ. The solid lines are the theoretical prediction from Theorem 2; the dots represent the simulated behavior of the modularity over time; and (b) Contour plot of the modularity measure as a function of p_r and p_Δ.

VI. CONCLUSIONS

This paper introduces a dynamic network model that captures the relationship between degree distributions and community structures based on a simple two-step mechanism: (i) a step of attachment in which a newly added node links to nodes of the same type with probability p_r; and (ii) a step of triad formation in which the newly added node

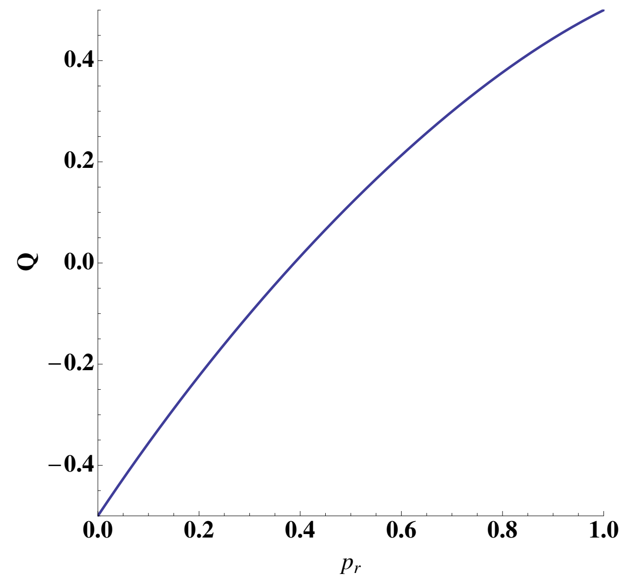


Fig. 3: Modularity of the network for $p_\Delta = 1$ and different values of p_r.

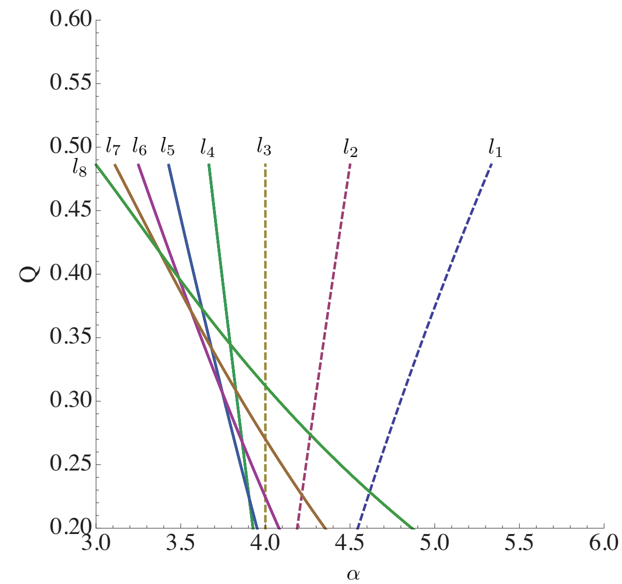


Fig. 4: Relationship between the modularity Q and the scaling exponent α for different values of p_r and p_Δ. Table I shows the parameter regime for the resulting lines.

TABLE I: Parameter regime for the correlations in fig. 4.

	Type of correlation	Range of p_r	p_Δ
l_1	Positive	[0.83,1]	0.3
l_2	Positive	[0.80,1]	0.4
l_3	–	[0.78,1]	0.5
l_4	Negative	[0.76,1]	0.6
l_5	Negative	[0.73,1]	0.7
l_6	Negative	[0.71,1]	0.8
l_7	Negative	[0.68,1]	0.9
l_8	Negative	[0.65,1]	1.0

may establish an additional link to one of the neighbors of the node it attaches to with probability p_Δ.

The proposed mechanism is of interest because the correlation between the degree distribution and the community structure resembles empirical network data. Characterizing the relationship between community structure and local clustering coefficient provides an important direction for future research.

VII. APPENDIX

Proof of Theorem 1. We assume that the in-degree of node i is a continuous variable $k_i \in \mathbb{R},\ k_i > 0$. Every time step t a newly added node $j \notin \mathcal{H}_{t-1}$ attaches to m different nodes in $\mathcal{H}_{t-1}$, selected according to a uniform distribution process over the N_0+t-1 existing nodes. The probability that node j attaches at time t to an existing node $i \in \mathcal{H}_{t-1}$ is

$$\frac{p_r m}{N_0+t-1} + \frac{(1-p_r)m}{N_0+t-1} = \frac{m}{N_0+t-1}$$

The triad formation step that follows random attachment adds to the rate of change of node i with in-degree $k_i(t-1)$ by

$$\begin{aligned} & \left(\frac{p_r m k_i(t)}{N_0+t-1}\right)\left(\frac{1}{m(1+p_\Delta)}\right)p_\Delta \\ + & \left(\frac{(1-p_r) m k_i(t)}{N_0+t-1}\right)\left(\frac{1}{m(1+(1-p_\Delta))}\right)(1-p_\Delta) \\ = & \left(\frac{k_i(t)}{N_0+t-1}\right)\left(\frac{p_r p_\Delta}{m(1+p_\Delta)} + \frac{(1-p_r)(1-p_\Delta)}{m(1+(1-p_\Delta))}\right) \end{aligned}$$

The terms $\frac{p_r m k_i(t)}{N_0+t-1}$ and $\frac{(1-p_r) m k_i(t)}{N_0+t-1}$ are the probabilities of selecting, during random attachment, an incoming neighbor of node i of the same type or different type respectively (*i.e.,* some node $j' \in q_i(t)$). Additionally, the terms $\frac{1}{m(1+p_\Delta)}$ and $\frac{1}{m(1+(1-p_\Delta))}$ are the probabilities that node j' is an incoming neighbor of node i of the same type or different type respectively (*i.e.,* $j' \in q_i(t)$). Finally, the probability p_Δ and $(1-p_\Delta)$ are the stationary mean of the random process of forming triads when nodes are of the same or different type. The multiplication of all 3 terms define the probability of forming a triplet with an edge that contributes to the in-degree of node i. The overall rate of change of $k_i(t)$ is

$$\begin{aligned} \frac{dk_i(t)}{dt} &= \frac{m}{N_0+t-1} \\ &+ \left(\frac{p_r p_\Delta}{1+p_\Delta} + \frac{(1-p_r)(1-p_\Delta)}{2-p_\Delta}\right)\frac{k_i(t)}{N_0+t-1} \end{aligned} \quad (3)$$

with boundary condition $k_i(t_i) = 0$. The solution to (3) is

$$k_i(t) = \left(\frac{m}{\tau}\right)\left(\frac{N_0+t-1}{N_0+t_i-1}\right)^{\tau} - \frac{m}{\tau} \quad (4)$$

with $\tau = \left(\frac{p_r p_\Delta}{1+p_\Delta} + \frac{(1-p_r)(1-p_\Delta)}{2-p_\Delta}\right)$. Using (4), the analytical expression for the cumulative distribution of the in-degree $P[k_i(t) \le k]$ of node i equals

$$\begin{aligned} & P\left[\left(\frac{m}{\tau}\right)\left(\frac{N_0+t-1}{N_0+t_i-1}\right)^{\tau} - \frac{m}{\tau} \le k\right] \\ = & P\left[t_i \ge \left(\frac{\frac{m}{\tau}}{k+\frac{m}{\tau}}\right)^{\frac{1}{\tau}}(N_0+t-1) - (N_0-1)\right] \end{aligned}$$

And as $t \to \infty$

$$P[k_i(t) \le k] = 1 - \left(\frac{\frac{m}{\tau}}{k+\frac{m}{\tau}}\right)^{\frac{1}{\tau}} \quad (5)$$

Finally,

$$p_k = \frac{dP[k_i(t) \le k]}{dk} = \frac{1}{\tau}\left(\frac{m}{\tau}\right)^{\frac{1}{\tau}}\left(k+\frac{m}{\tau}\right)^{-\left(1+\frac{1}{\tau}\right)} \quad (6)$$

Note that (6) exhibits the extended power law of the form

$$p_k \sim (k+\varepsilon)^{-\alpha}$$

where $\alpha = 1+\frac{1}{\tau}$ and $\varepsilon = (\alpha-1)\,m$. When $k \gg \varepsilon$, (6) is reduced to a single power law $p_k \sim k^{-\alpha}$. On the other hand, when $k \ll \varepsilon$ we have

$$\begin{aligned} \ln p_k \sim -\alpha\ln(k+\varepsilon) &= -\alpha\left[\ln\left(1+\frac{k}{\varepsilon}\right) + \ln\varepsilon\right] \\ &\sim -\alpha\left[\frac{k}{\varepsilon} + \ln\varepsilon\right] \end{aligned}$$

and obtain

$$p_k \sim \varepsilon^{-\alpha}\exp\left(-\alpha\frac{k}{\varepsilon}\right)$$

Thus, (6) is proportional to the exponential form $p_k \sim \exp(-\lambda k)$ with $\lambda = \frac{\alpha}{\alpha-1}$.

□

Proof of Theorem 2. To measure the modularity Q in the formation of communities we use the metric defined in [15]. Let $\mathbf{e}$ be a 2×2 matrix whose element e_{ij} represents the fraction of edges that connect the nodes in two non-overlapping communities i and j. The trace $\mathrm{Tr}\,\mathbf{e} = \sum_i e_{ii}$ gives the fraction of the links in the network that connect nodes in the same community. The row (or column) sums $a_i = \sum_j e_{ij}$ represents the fraction of links with at least one node in community i. It can be shown that if edges are build randomly (without considering communities) then $e_{ij} = a_i a_j$ [14]. The modularity Q measures the number of edges within each community minus the expected number of such edges if they fall randomly between the nodes and is defined as

$$Q = \sum_i (e_{ii} - a_i^2) = \mathrm{Tr}\,\mathbf{e} - \|\mathbf{e}^2\| \quad (7)$$

where $\|\mathbf{e}\|$ indicates the sum of the elements of the matrix $\mathbf{e}$. To calculate the fraction of edges that connect every type of nodes, we consider two possible scenarios for the dynamic evolution of (i) the number of links that connect nodes of the same type; and (ii) the number of links that connect nodes of different type. In scenario (i), there are an expected

$$m(p_r + p_r p_\Delta)t + e_{ii}^0 \quad (8)$$

number of links between nodes of the same type, where e_{ii}^0 is the amount of links between nodes of the same type in the initial network $\mathcal{G}_0$.

In scenario (ii), the number of links created between nodes of different type is

$$m\left((1-p_r) + (1-p_r)(1-p_\Delta)\right)t + e_{ij}^0 \quad (9)$$

where e_{ij}^0 is the amount of links between nodes of different type in the initial network $\mathcal{G}_0$. The total amount of links created in the network depends on the dynamic behavior of the in-degree of the nodes (equation. (4)). The sum over all nodes can be written in the integral form as

$$d(t) = \int_1^t \left\{ \left(\frac{m}{\tau}\right) \left(\frac{N_0 + t - 1}{N_0 + t_i - 1}\right)^\tau - \frac{m}{\tau} \right\} dt_i \qquad (10)$$

Finally we can express the matrix $\mathbf{e}$ as 0.5 times

$$\begin{bmatrix} \frac{m(p_r + p_r p_\triangle)t + e_{ii}^0}{d(t)+e} & \frac{m((1-p_r)+(1-p_r)(1-p_\triangle))t + e_{ij}^0}{d(t)+e} \\ \frac{m((1-p_r)+(1-p_r)(1-p_\triangle))t + e_{ij}^0}{d(t)+e} & \frac{m(p_r + p_r p_\triangle)t + e_{ii}^0}{d(t)+e} \end{bmatrix}$$

where $e = \sum_{i \in \mathcal{H}_0} q_i(0)$ is the initial number of edges in $\mathcal{G}_0$. The factor of 0.5 comes from the fact that the expected type of node of an incoming node is equal for both types. In general, as $t \to \infty$ the value of Q tends to a constant that depends on the link acquisition tendency of the two-step mechanism of attachment and triad formation, p_r and $p_\triangle$. Consider the following values for p_r and $p_\triangle$ which represent particular cases of interest.

Case 1 *(Full preference during random attachment):* When $p_r = 1$

$$Q = \lim_{t \to \infty} \sum_i (e_{ii} - a_i^2) = \lim_{t \to \infty} \operatorname{Tr} \mathbf{e} - \|\mathbf{e}^2\| = 0.5 \qquad (11)$$

Case 2 *(Full preference during triad formation):* When $p_\triangle = 1$

$$\begin{aligned} Q &= -0.125 p_r^4 + 0.25 p_r^3 \\ &- 0.625 p_r^2 + 1.5 p_r - 0.5 \end{aligned} \qquad (12)$$

Note that (12) correspond to the plot shown in fig. 3.

□

References

[1] G. Caldarelli, R. Marchetti, and L. Pietronero, "The fractal properties of Internet," *Europhys. Lett.*, vol. 52, no. 4, pp. 386–391, 2000.

[2] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, "Dynamical and correlation properties of the Internet," *Phys. Rev. Lett.*, vol. 87, no. 25, p. 258701, 2001.

[3] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the World-Wide Web," *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.

[4] M. E. J. Newman, "The structure of scientific collaboration networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 98, no. 2, pp. 404–409, 2001.

[5] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, "Hierarchical organization of modularity in metabolic networks," *Science*, vol. 297, no. 5586, p. 1551, 2002.

[6] M. Mitzenmacher, "A brief history of generative models for power law and lognormal distributions," *Internet Mathematics*, vol. 1, no. 2, pp. 226–251, 2004.

[7] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Rev.*, vol. 51, no. 4, pp. 661–703, 2009.

[8] P. Moriano and J. Finke, "Power-law weighted networks from local attachments," *Europhys. Lett.*, vol. 99, no. 1, p. 18002, 2012.

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

[10] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Phys. Rev. E*, vol. 65, no. 2, p. 026107, 2002.

[11] P. Moriano and J. Finke, "Structure of growing networks with no preferential attachment," in *Proceedings of the American Control Conference*, Washington, DC, June 2013.

[12] P. Moriano and J. Finke, "On the formation of structure in growing networks," *J. Stat. Mech. Theor. Exp.*, vol. 6, p. P06010, 2013.

[13] E. A. Leicht and M. E. J. Newman, "Community structure in directed networks," *Phys. Rev. Lett.*, vol. 100, no. 11, p. 118703, 2008.

[14] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 99, no. 12, pp. 7821–7826, 2002.

[15] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, p. 026113, 2004.

[16] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 23, pp. 8577–8582, 2006.

[17] A. Grönlund and P. Holme, "Networking the seceder model: Group formation in social and economic systems," *Phys. Rev. E*, vol. 70, no. 3, p. 036108, 2004.

[18] J. D. Noh, H.-C. Jeong, Y.-Y. Ahn, and H. Jeong, "Growing network model for community with group structure," *Phys. Rev. E*, vol. 71, no. 3, p. 036131, 2005.

[19] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, "Models of social networks based on social distance attachment," *Phys. Rev. E*, vol. 70, no. 5, p. 056122, 2004.

[20] C. Li and P. K. Maini, "An evolving network model with community structure," *J. Phys. A: Math. Gen.*, vol. 38, no. 45, pp. 9741–9749, 2005.

[21] C. Li and G. Chen, "Modelling of weighted evolving networks with community structures," *Physica A*, vol. 370, no. 2, pp. 869–876, 2006.

[22] M. H. Li, S. G. Guan, and C.-H. Lai, "Formation of modularity in a model of evolving networks," *Europhys. Lett.*, vol. 95, no. 5, p. 58004, 2011.

[23] M. O. Jackson and B. W. Rogers, "Meeting strangers and friends of friends: How random are social networks?," *Am. Econ. Rev.*, vol. 97, no. 3, pp. 890–915, 2007.

[24] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill Higher Education, 4th ed., 2002.

[25] Z. Lai, J. Su, W. Chen, and C. Wang, "Uncovering the properties of energy-weighted conformation space networks with a hydrophobic-hydrophilic model," *Int. J. Mol. Sci.*, vol. 10, no. 4, pp. 1808–1823, 2009.