Power-law weighted networks from local attachments

P. Moriano(a) and J. Finke(b)

Department of Electrical Engineering and Computer Science, Pontificia Universidad Javeriana
Santiago de Cali, Colombia

received 30 January 2012; accepted in final form 5 June 2012
published online 9 July 2012

PACS 89.75.Da – Systems obeying scaling laws
PACS 89.75.Fb – Structures and organization in complex systems

Abstract – This letter introduces a mechanism for constructing, through a process of distributed
decision-making, substrates for the study of collective dynamics on extended power-law weighted
networks with both a desired scaling exponent and a fixed clustering coefficient. The analytical
results show that the connectivity distribution converges to the scaling behavior often found in
social and engineering systems. To illustrate the approach of the proposed framework we generate
network substrates that resemble steady state properties of the empirical citation distributions of
i) publications indexed by the Institute for Scientific Information from 1981 to 1997; ii) patents
granted by the U.S. Patent and Trademark Office from 1975 to 1999; and iii) opinions written by
the Supreme Court and the cases they cite from 1754 to 2002.

Copyright © EPLA, 2012

Introduction. – Understanding structure lies at the
very heart of the study of complex networks. A network is
a collection of a large number of interconnected elements
(units or agents) whose interaction with each other and
with the surroundings leads to characteristic properties
that can only be attributed to the network as a whole [1].
Networks often develop distinct structural steady state
patterns. Studying these patterns, promises to enhance
our understanding of the dynamics underlying collective
human responses [2], corrupt behavior [3], and economic
development [4].

Random graph models fail to capture key features
of real-world networks (e.g., clustering coefficients
and degree correlations). Recent efforts to understand
network structure have focused on connectivity distribu-
tions underlying a number of social and engineering
systems which, rather than following the Poisson dis-
tribution of random networks (bounded by Chebyshev’s
inequality), have heavy tails [5]. Heavy-tailed distribu-
tions in empirical data suggest the existence of causal
mechanisms that shape the structure and function of
real-world networks [6]. In the era of “big data,” the
development of formal frameworks that quantify patterns
of interaction of networks has set the research agendas
across various disciplines (e.g., more recently across the
data driven computational social sciences).

Power laws, a particular type of heavy-tailed distribu-
tions, have received significant attention in recent years.
For a network with an extended power-law connectiv-
ity distribution, if the number of connections of a node
is much larger than \( x_0 \), the probability that the node
connects to \( x \) other nodes is proportional to \( x^{-\alpha} \) for some
positive constants \( \alpha \) and \( x_0 \) [7]. As a result, the tail of the
distribution has no exponential bound and the connectiv-
ity of the nodes of the network comprises different orders
of magnitude, with a few nodes being highly connected.

Key to modeling power-law networks is the characteriza-
tion of hubs (highly interconnected nodes). In the context
of the spread of disease, measuring patterns in regions
that are more vulnerable to infection (hubs) allows us to
respond more effectively to the potential spread of large-
scale epidemics [8]. The ability to understand and recre-
ate the structure of epidemic networks allows us to design
strategies that embrace how interconnected regions influ-
ence one another (as a result of the evolution of social
systems) in order to quantify and predict the dimensions
of disease.

To capture the relationships between the elements of a
network, e.g., duration, emotional intensity, or intimacy,
models define weights as an inherent property between
nodes [9]. Recent models of weighted networks have
focused on attachment strategies in which nodes are added
according to probability distributions on the existing
weights across the entire network. The network model
introduced in [10] captures the evolution of weights driven
by preferential strength attachment, a mechanism in which
newly added nodes are more likely to connect to nodes associated with larger weights. Lacking local competitive factors between nodes, the resulting networks exhibit power-law distributions where the hubs correspond to the nodes that have been part of the network the longest.

This letter introduces a wide class of attachment strategies which promote the formation of hubs based on both the length of time a node has been part of the network (i.e., node longevity) and its ability to compete for weights with surrounding neighbors (i.e., node fitness). Because the connectivity dynamics of the nodes depend on their attractiveness to compete for weights (as in [11]), older nodes are not necessarily more successful in acquiring weights. To our knowledge the proposed mechanism is novel in that it generates weighted directed networks with extended power-law strength distributions i) in a distributed fashion (decision-making strategies are based on local information; we do not assume any type of global information to generate the desired network structure); ii) for an arbitrary scaling exponent $\alpha > 2$ and a fixed clustering $c \in (0, 1)$ (as in [12]); and iii) for values greater than a particular threshold $\bar{s} > 0$ (for the case when only the tail of the distribution obeys a power law).

The remaining sections are organized as follows. First we introduce a model that captures the connectivity and growth dynamics of the gradual addition of nodes to an existing network component and proposes attachment strategies for local rearrangement of weights between pairs of nodes. We prove that for any connected network there exists a distribution of the total weight from neighboring nodes (node strength) that is asymptotically stable (i.e., the proposed strategies lead to a Nash equilibrium [13]). Moreover, as the network grows, consecutive achievements of this network state lead to weighted directed networks with extended power-law strength distributions and distinctive clustering coefficients (defined as the ratio of the total average weight of transitive triplets over the total weight of possible triplets). We present simulations that capture the effect of node fitness and illustrate the application of the proposed model to generate various citation networks. Finally, we draw some conclusions and future research directions.

A model of network topology and growth. – Consider a directed network that captures weighted relationships between a set of nodes. As the network grows, more nodes join the network, each possessing a small “budget” used to construct directed links to some existing nodes. When, node $i$ establishes a link to node $j$ (by passing some of its budget to node $j$) node $j$ has more budget to spend, which it may do by increasing its weighted connections to other nodes. Broadly speaking, every node wishes to spend its budget, but the more it spends the less willing it is to spend more. Nodes will locally rearrange their weights until every node reaches an equilibrium. At the equilibrium all nodes have associated gains that are equal and there are no further incentives to rearrange connections.

To formalize this idea let us introduce the following notation. Let $\mathcal{H}_1 = \{1, \ldots, N_1\}$ be a finite set of nodes at generation $k = 1$. Nodes represent elements (acting units) that establish connections to other nodes. We represent the relationship between nodes using a weighted matrix $W_{ij} = [w_{ij}]_{N_i \times N_j}$, where $w_{ij} \in \mathbb{R}_+$ quantifies the relationship between node $i$ and $j$. If $w_{ij} > 0$, then there exists some kind of action from $i$ to $j$ with weight $w_{ij}$. It may capture, for instance, the extent to which node $i$ influences node $j$. Let $G_k = (\mathcal{H}_k, W_k)$ represent the network at generation $k$ (because in general $w_{ij} \neq w_{ji}$, the network is modeled as a directed graph). For a fixed generation, let $p(i) = \{j: w_{ij} > 0\}$ represent all nodes which influence node $i$ (incoming neighbors). Similarly, let $q(i) = \{j: w_{ij} > 0\}$ represent all nodes influenced by node $i$ (outgoing neighbors). A gain function $g_i(s_i)$ is associated to each node $i \in \mathcal{H}_k$ and characterizes the marginal benefit that results from its current set of connections, where $s_i = \sum_{j \in p(i)} w_{ij}, s_i \in \mathbb{R}_+$. Note that $s_i$ is a scalar that represents the incoming strength of node $i$ (referred to as node strength hereafter). The following network assumptions are needed:

A1) Finite network strength: The total weight of the initial network $P_1 = \sum_{i=1}^{N_1} s_i, P_i \in \mathbb{R}_+$, is finite. In other words, the extent to which any node in the network can be influenced by other nodes is bounded.

A2) Connectedness: Every node is influenced to some extent by another node. At each generation $k$, $s_i \geq \epsilon > 0, \forall i \in \mathcal{H}_k$.

A3) Bounded marginal gains: The gain function $g_i(s_i)$ associated to node $i \in \mathcal{H}_k$ satisfies

$$-a_i \leq g_i(y_i) - g_i(z_i) \leq -b_i$$

for any $y_i, z_i \in \mathbb{R}_+, y_i \neq z_i$ and some constants $a_i \geq b_i > 0$. In other words, the marginal gain associated with each node decreases with increasing strength. Equation (1) eliminates the possibility that a very small difference in node strength may result in an unbounded change in gain. Note that if $g_i$ is differentiable and has a negative derivative it satisfies eq. (1).

Next, we use $t \geq 0$ to specify the time index of events. Let $t = \tau_k$ be the time instant when a new node is added to form the network $G_k$ (i.e., the start of generation $k$). Let $\tau^+_k$ be the instant right before the new node is added to $G_k$ (i.e., the start of generation $k + 1$). When $t = \tau_k + 1$, $G_k$ evolves into $G_{k+1}$. For generation $k$ let the set of states

$$S_k = \left\{ s \in \mathbb{R}_+^N : \sum_{i=1}^{N_k} s_i = P_k \right\}$$

be the simplex over which the connectivity dynamics evolve. Constraints on our model below will ensure that
for all nodes \( i \in \mathcal{H}_k \), \( s(t) \in \mathcal{S}_k \) for all \( \tau_k \leq t < \tau_{k+1} \). We assume that as \( t \to \tau_{k+1}^+ \), the time allowed for the events that drive the connectivity dynamics during generation \( k \) goes to infinity. Let \( s(t) = [s_1(t), \ldots, s_{N_k}(t)]^\top \in \mathcal{S}_k \) be the state vector for \( G_k \) at time \( t \) (i.e., the incoming strength distribution of the entire network).

**Connectivity dynamics.** We first focus on the dynamics of \( s(t) \) for \( \tau_k \leq t < \tau_{k+1} \) (i.e., within a fixed generation).

In particular, we want to define the singleton
\[
\mathcal{S}_k^* = \{ s \in \mathcal{S}_k : \text{for all } i, j \in \mathcal{H}_k, \, g_i(s_i) = g_j(s_j) \},
\]
(2) such that any strength distribution that belongs to this set represents a distribution where all nodes in \( \mathcal{H}_k \) have equal gain levels. To capture the connectivity dynamics that lead to \( \mathcal{S}_k^* \), let \( e^{\sigma(i)}_\mu \) represent the decision of node \( i \) to weaken its relation from some node \( j \) in \( p(i) \) while strengthening its relation to other nodes in \( \mathcal{H}_k \). Let the set \( \sigma(i) = (\sigma_1(i), \sigma_2(i), \ldots, \sigma_{|p(i)|}(i)) \) such that \( j < j' < \cdots < j'' \) and \( j, j', \ldots, j'' \in \mathcal{H}_k \) be composed of elements \( \sigma_\mu(i) \) that denote the weight to be added or created to the to link \( \omega_{ij} \) between node \( i \) and node \( j \in \mathcal{H}_k \). For convenience, we will denote this list by \( \sigma(i) = (\sigma_\mu(i), j \in \mathcal{H}_k) \). Similarly, let the list \( (\mu_j(i); j \in p(i)) \) be composed of elements \( \mu_j(i) \) that denote the weight to be subtracted from the link \( w_{ij} \) where node \( j \in p(i) \).

Let \( \{ e^{\sigma(i)}_\mu \} \) denote the set of all possible combinations of how node \( i \) can weaken or strengthen its relations to other nodes. Let the set of events be described by \( \mathcal{E}_1 \equiv \mathcal{P}((\{ e^{\sigma(i)}_\mu \}) - \{ \emptyset \} \) (\( \mathcal{P}(\cdot) \) denotes the power set). We call \( e_1(t) \), \( \tau_k \leq t < \tau_{k+1} \), events of type 1; they drive the connectivity dynamics within a network generation. Notice that each event \( e_1(t) \in \mathcal{E}_1 \) is defined as a set, with each element of \( e_1(t) \) representing the potential rearrangement of multiple weights between nodes, and multiple elements in \( e_1(t) \) representing the simultaneous rearrangements among multiple nodes.

An event \( e_1(t) \) may occur only if it belongs to the set defined by an enable function \( h_1; \mathcal{S}_k \rightarrow \mathcal{P}(\mathcal{E}_1) - \{ \emptyset \} \), specified for node \( i \in \mathcal{H}_k \) as follows:

- If \( g_i(s_i) \geq g_j(s_j) \) for all \( j \in q(i) \), then \( e^{\sigma(i)}_\mu \in e_1(t) \) such that \( \sigma(i) = (0, \ldots, 0) \) and \( \mu(i) = (0, \ldots, 0) \) is the only enabled event. Hence, node \( i \) does not modify its relationships to others nodes (i.e., the strength of node \( i \) does not change).

- If \( g_i(s_i) < g_j(s_j) \) for some \( j \in q(i) \), then the only \( e^{\sigma(i)}_\mu \in e_1(t) \) are ones with \( \sigma(i) = (\sigma_j(i); j \in \mathcal{H}_k) \) and \( \mu(i) = (\mu_j(i); j \in p(i)) \) such that
\[
\begin{align*}
C1) \quad & \sum_{j \in \mathcal{H}_k} \sigma_j(i) = \sum_{j \in p(i)} \mu_j(i), \\
C2) \quad & \sigma_j(i) \geq \frac{1}{a_i} \gamma (g_j(s_j) - g_i(s_i)), \\
C3) \quad & \sum_{j \in p(i)} \mu_j(i) \leq \frac{1}{b_i} (g_j(s_j) - g_i(s_i)) - \sigma_j(i),
\end{align*}
\]

for some \( j^\ast \in \{ j; g_j(s_j) \geq g_i(s_i), \text{ for } r \in q(i) \} \) and \( \gamma \). The parameter \( \gamma \in (0,1) \) regulates the speed at which weights are rearranged and affects the transitivity of the network (i.e., if a node \( j \) is connected to node \( j^\ast \) and node \( j^\ast \) to node \( j'' \), the probability that node \( j \) is also connected to node \( j'' \)). Low values of \( \gamma \) lead to slower convergence processes which increase the probability of forming transitive triples and lead to higher clustering coefficients.

Condition (C1) implies that a node can only establish or strengthen its relations to other nodes by weakening incoming weights (the sum of incoming weights must equal the sum of outgoing weights). It implies that \( G_k \) conserves total network strength, i.e., \( P_k = \sum_{i=1}^{N_k} s_i(t) \) is constant. To interpret (C2) and (C3) it is useful to remember that reducing (increasing) the strength of a node always increases (decreases, respectively) its gain. Both conditions constrain how nodes can modify their weights in terms of the gain of outgoing neighbors. Condition (C2) implies that if the gain of node \( i \) differs from any of its outgoing neighbors, then the relation to some neighbor with the highest gain must be strengthened by some amount. Condition (C3) implies that, when node \( i \) weakens incoming weights, node \( i \) cannot exceed the highest gain of at least one outgoing neighbor. Together they guarantee that the highest gain of the network is strictly monotonically decreasing over time (as we prove in Theorem 1).

Next, state transitions are defined by the operator \( f_1; \mathcal{S}_k \rightarrow \mathcal{S}_k \) where \( e_1(t) \in \mathcal{E}_1 \). For a fixed generation \( k \), if \( e_1(t) = h_1(s(t)) \), \( e^{\sigma(i)}_\mu \in e_1(t) \), then \( s(t+1) = f_1(s(t)) \), where
\[
\begin{align*}
s_i(t+1) &= s_i(t) + \sum_{\{j \in \mathcal{H}_k; e^{\sigma(i)}_\mu \in e_1(t)\}} \sigma_j(i) - \sum_{\{j \in p(i), e^{\sigma(i)}_\mu \in e_1(t)\}} \mu_j(i).
\end{align*}
\]
Equation (3) means that the strength at node \( i \) at time \( t+1 \) equals the strength of node \( i \) at time \( t \), plus the total weight added by the nodes that strengthened their relationship to node \( i \), minus the total weight reduced by nodes that weakened their relation to node \( i \) at time \( t \).

Let \( E_1 \) denote the set of all infinite sequence of events \( \mathcal{E}_1 \). Let \( E_{1}^\ast \) denote the sequence of events \( e_1(0), \ldots, e_1(t-1) \) and let the value of the function \( S(s(0), E_{1}^\ast, t) \) denote the state reached at time \( t \) from the initial state \( s(0) \) by the application of the sequence \( E_{1}^\ast \) of events of type 1. We assume that each event of type 1 occurs infinitely often on each event trajectory \( E_{1}^\ast \), \( \tau_k \leq t < \tau_{k+1} \). This assumption is met if nodes persistently try to rearrange weights. The enable function \( h_1 \) together with state transition operator \( f_1 \) define the evolution of the connectivity dynamics of the network.

**Growth dynamics.** We now turn our attention to the evolution of the network as it grows. To capture a
nodes’s advantage of longevity let \( k_i \) be the generation when node \( i \) is added to the network. We denote \( n_i = \frac{k_i}{k} \) as the fraction of generation indices node \( i \) was not part of the network component. Moreover, to capture a node’s competitive advantage in acquiring weights we associate to every node a fitness \( \beta_i \), where \( \beta_i \in (0,1) \). Let \( s_0 \geq 0 \) be a constant amount of strength such that \( s_i > s_0 \). Let the gain function (marginal utility) associated to node \( i \in \mathcal{H}_k \) during generation \( k \) be

\[
g_i(s_i) = \frac{1}{s_i - s_0} \left( \frac{1}{n_i} \right)^{\beta_i}.
\]

Higher values of \( \beta_i \) characterize nodes that are more attractive in the sense that they can carry more weight without greatly reducing their gain. Both high values of \( n_i \) (representing the fact that node \( i \) has been part of the growing network for only a few generations) and low values of \( \beta_i \) (representing the fact that the node has a low competitive advantage for acquiring weights) have a negative effect on the gain of node \( i \).

Let \( e^{\sigma(i)} \) represent the attachment of a new node \( i \) to the network at the beginning of generation \( k \) (when \( t = \tau_k \)). Let \( m = \sum_i \sigma_i(i) \) be the total (constant) weight of a newly added node. A node attaches to the network component by i) randomly distributing its weight \( \sigma(i) \) across some nodes and ii) establishing a non-empty set of incoming neighbors (i.e., some node must connect to it). We call the attachment of nodes to \( \mathcal{G}_k \), \( k = 1,2,\ldots \) events of type 2. Let \( \mathcal{E}_2 = \{e^{\sigma(i)}\} \) denote all possible combinations of how node \( i \) can attach to the network component. An event \( e_2(k) \in \mathcal{E}_2 \) may occur if it is defined by an enable function \( h_2: \mathcal{S}_k \rightarrow e^{\sigma(i)} \), specified for a newly added node as follows:

- Node \( i \) attaches to the network only if the associated gain function \( g_i(s_i) \) follows the general form of (4) with longevity and fitness parameters that satisfy
  \[
  \text{C4)} \quad n_i = 1,
  \]
  \[
  \text{C5)} \quad \beta_i = \beta \quad \forall i \in \mathcal{H}_k.
  \]

Condition C4) follows from letting \( k_i = k \) for the newly added node (at generation \( k \) node \( i \) has been part of network for one generation). Condition C5) specifies an equal fitness value for every node (as is the case for networks with linear growth under preferential attachment).

The transition \( e_2(k) \in \mathcal{E}_2 \) is defined by the operator \( f_2: \mathcal{S}_k \rightarrow \mathcal{S}_{k+1} \). If \( e_2(k) \in h_2(s(\tau_k)) \), then \( s(\tau_{k+1}) = f_2(s(\tau_k)) \) where \( s(\tau_{k+1}) = m \) only if node \( i \) is the newly added node. Let \( \mathcal{E}_2 \) denote the set of all infinite sequence of events \( \mathcal{E}_2 \). Let \( \mathcal{E}_2^k \) denote the sequence of events of type 2, \( e_2(1), \ldots, e_2(k) \). We assume that each event of type 2 occurs infinitely often on each event trajectory \( E_2^k \). The assumption is met if nodes constantly attach to the existing network component. The enable function \( h_2 \) together with the transition operator \( f_2 \) define the growth dynamics of the network.

**Analysis.** Next, we present stability properties of the invariant set \( \mathcal{S}_k^* \) and deduce the average gain level of the network \( \mathcal{G}_k \). We then prove that, for values greater than a threshold \( \hat{s} \), the strength distribution converges to a scaling behavior.

**Theorem 1:** Suppose A1)–A3) and C1)–C3) hold. Then \( \mathcal{S}_k^* \) is an invariant set and has region of asymptotic stability equal to \( \mathcal{S}_k \).

Theorem 1 guarantees that for any generation \( k \), initial network state \( s(0), \) and event sequence \( E_2^1, S(s(0), E_2^1, t) \rightarrow \mathcal{S}_k^* \) as \( t \rightarrow \tau_k+1 \) for generation \( k \). Broadly speaking, the conditions in Theorem 1 capture the dynamic coupling between different nodes that lead to a Nash equilibrium. By attaining the same gain level no node can increase its gain by changing its connections unilaterally without making the average gain of all other nodes worse off. When \( \mathcal{S}_k^* \) is reached the average gain of the network an instant before the start of generation \( k+1 \) is given by

\[
\mathcal{C}_k = \frac{1}{N_1 + k} \sum_{i \in \mathcal{H}_k} g_i(s_i(\tau_k^+)).
\]

As the network grows, the behavior of the average gain is characterized by the following lemma:

**Lemma 1:** Suppose A1)–A3) and C1)–C5) hold. Moreover, \( \forall k \) let \( s(\tau_k^+) \in \mathcal{S}_k^* \) then \( \mathcal{C}_k \rightarrow 1/(m + s_0)(1 - \beta) \) as \( k \rightarrow \infty \).

Lemma 1 implies that at the desired strength distribution \( \mathcal{S}_k^* \), the average gain tends to \( \mathcal{C}_k \rightarrow 1/(m + s_0)(1 - \beta) \) as \( k \rightarrow \infty \).

The following theorem implies that as the network grows, it develops an extended power-law structure driven by the marginal benefit of the allocation of weights across nodes and quantifies the value \( \hat{s} \) above which the scaling behavior emerges.

**Theorem 2:** Suppose A1)–A3) and C1)–C5) hold. Moreover, \( \forall k \) let \( s(\tau_k^+) \in \mathcal{S}_k^* \). Then the strength distribution \( P(s_i - s_0 > \omega) \) of the network \( \mathcal{G}_k(H_k, W_k) \) follows an extended power law with scaling exponent \( \alpha = 1/\beta + 1 \) as \( k \rightarrow \infty \). The scaling behavior holds for values greater than \( \hat{s} = (m + s_0)/(1 - \beta) \).

Note that if \( s_0 = 0 \) the model yields power law rather than extended power-law distributions (as in preferential attachment with linear growth) for values greater than \( m(1 - \beta) \) [10].

Extended power-law distributions emerge as a result of both the interaction between local mechanisms that lead to Nash equilibria and the continuum attachment of new nodes to the network. In particular, when the network is at a Nash and a new node is added, it introduces a perturbation to the existing set of strategies. Conditions C1)–C3) force the network to return to a state which again represents a Nash, with subsequent achievements of Nash equilibria shaping the structure of the network.

---

1. The proofs of all theorems can be found in the supplement to this letter in the e-print in ref. [14].
Simulations. – To gain insight into the connectivity dynamics let $\beta \equiv 1$, $m = 1$, $s_0 = 0$, $N_1 = 2$, and consider a network after $k = 1000$ generations. Figure 1 shows the value of the clustering coefficient $c = \frac{\sum_{i,j} w_{i,j}}{\sum_i w_i}$ (i.e., the ratio of the total average weight of transitive triplets over the total weight of possible triplets) as a function of the size. Note that for any $\gamma \in (0, 1)$ the clustering properties remain constant as the network grows. Figure 2 shows the effect of varying node fitness, where $\beta_i$ is chosen from a uniform distribution with support $(0, 1)$. Figure 2(a) shows the evolution of the node’s strength for different values of $\beta$. Note that $s_i(\tau_{k+1}^\gamma)$ follows a power law for all values of $\beta_i \sim U(0, 1)$. Because of their competitive advantage, there are some nodes with more strength $s_i$ which have been part of the network for only a few generations. It is possible for a node to join the network at a more recent generation and become more attractive than other nodes that have been part of the network for longer. In particular, fig. 2(a) shows that the node added at generation $k = 105$ with $\beta_{105} = 0.9$ overcomes older nodes with $\beta_{35} = 0.6$ and $\beta_5 = 0.3$. In fig. 2(b), the cumulative strength distribution for the entire network suggests a power law with a logarithmic corrective term similar to the theoretical prediction in [11] where $p_c(\omega) \sim \frac{1}{\log(\omega)} \omega^{-(1+C^*)}$ with $C^* = 1.255$. Finally, fig. 3 shows empirical data on the citation distribution of articles indexed by the Institute for Scientific Information (ISI); patents granted by the U.S. Patents and Trade Office; and opinions written by the U.S. Supreme Court and the cases they cite. Figure 3(a) illustrates the case for scientific papers published in 1981 and cited between 1981 and 1997 [15]. The authors of [6] estimated both the scaling exponent $\alpha^* = 3.16$ and the threshold $\hat{s}^* = 160 \pm 35$ at which the scaling behavior emerges. Figure 3(b) represents citations on the main subnetwork of U.S. patents granted between 1963 and 1999 and references made to these patents between 1975 and 1999 [16]. Figure 3(c) shows the majority opinions written by the U.S. Supreme Court and the cases they cite from 1754 to 2002 [17]. All three citation networks follow extended power-law distributions (for the last two examples we estimate the values of $\alpha^*, \hat{s}^*$, and $e^*$ from empirical data).

Finally, we compare the distributions from empirical data with the distributions predicted by the proposed and previous models [6,7]. We measure the greatest discrepancy between the empirical and the expected distribution (D-statistic), as well as the sum of squares of the deviations between the two distributions. Table 1 summarizes the model parameters and the results. Note that the performance of other (perhaps simpler) models degrades when the entire range is considered.

Discussion. – The proposed model generates extended power-law distributions from consecutive achievements of

| Table 1: Model parameters for the three citation networks. |
|-------------------------|-------------------|-------------------|
|                        | Papers            | Patents           | Court cases |
| Empirical network      |                   |                   |             |
| $\alpha^*$             |                   |                   |             |
| $\hat{s}^*$            |                   |                   |             |
| $e^*$                  |                   |                   |             |
| Generated network      |                   |                   |             |
| $\beta$                |                   |                   |             |
| $\gamma$               |                   |                   |             |
| $\alpha$               |                   |                   |             |
| $\hat{s}$              |                   |                   |             |
| $e$                    |                   |                   |             |
| Model parameters       |                   |                   |             |
| $m$                    |                   |                   |             |
| $s_0$                  |                   |                   |             |
| $C^*$                  |                   |                   |             |
| D-statistic            |                   |                   |             |
| Sum of squares         |                   |                   |             |
| Range                  |                   |                   |             |

(a) For the paper citation network we use the data and the distribution predicted by the model introduced in [6]. For the U.S. patent citation network we use the data presented in [16] and the distribution predicted by the model introduced in [7] with $\alpha = 4.9$ and $\hat{s} = 9.6$. For the U.S. Supreme Court citation network we use the data presented in [17] and the distribution predicted by the model introduced in [7] with $\alpha = 4.1$ and $\hat{s} = 35$, respectively.

Empirical paper network:
- $\alpha^* = 3.16$, $\hat{s}^* = 160 \pm 35$, $e^* = 0.31$

Empirical patent network:
- $\alpha^* = 3.25$, $\hat{s}^* = 250 \pm 50$, $e^* = 0.32$

Empirical court cases:
- $\alpha^* = 3.33$, $\hat{s}^* = 300 \pm 60$, $e^* = 0.33$

Power-law weighted networks from local attachments
stable strength distributions $S_k^*$. Although it does not pretend to empirically validate real-world mechanisms behind citation networks, the model may be of interest in the following context. First, it can be shown that the state $S_k^*$ is a Nash, which implies that when a network reaches the equilibrium there is not any node that can gain by unilaterally rearranging weights to neighboring nodes (there are no incentives to change or establish new relationships). By focusing on the dynamics that drive the network to $S_k^*$ we capture the coupling between different nodes, characterizing how relationships between any pair of nodes affect other nodes in the network. Second, the proposed strategies allow us to control the connectivity dynamics of nodes based on local attachment strategies (C1–C5)), allowing us to generate network substrates through distributed decision-making. Finally, the ability to control the rate at which attachment strategies lead to the scaling behavior allows us to obtain non-negligible clustering coefficients for large networks.

We focused on two types of network incentives: i) Longevity rewards nodes that have been part of the network for a long time (they have the ability to acquire more weight compared to recently added ones); ii) Fitness rewards nodes that are highly competent (they are more suitable to compete and maintain weights). Modeling nodes with varying fitness allows “latecomers” to overcome nodes that have been in the network for longer generations.

Following similar ideas as in Theorems 1 and 2, the proposed framework can be extended to generate exponential strength distributions. In particular, if we consider the gain function of the general from $g_i(s_i) = \alpha_i \ln \left( \frac{1}{n_i} + \kappa \right)$, where $\kappa > 0$, the proposed strategies lead to weighted networks with $P[s_i > \omega] \sim e^{-\omega}$. A mathematical framework that allows us to generate various strength distributions for different domain intervals provides an important direction for future research.

REFERENCES

[1] Amaral L. A. N. and Ottino J. M., Eur. Phys. J. B, 38 (2004) 147.
[2] Cho A., Science, 325 (2009) 406.
[3] Bohannon J. S., Science, 325 (2009) 409.
[4] Farmer J. D. and Foley D., Nature, 460 (2009) 685.
[5] Finke J., Quijano N. and Passino K. M., EPL, 82 (2008) 28004.
[6] Clauset A., Shalizi C. R. and Newman M. E. J., SIAM Rev., 51 (2009) 661.
[7] Valverde S., Solé R. V., Bedau M. A. and Packard N., Phys. Rev. E, 76 (2007) 056118.
[8] Christakis N. A. and Fowler J. H., PLoS ONE, 5 (2010) e12948.
[9] Barrat A., Barthelemy M., Pastor-Satorras R. and Vespignani A., Proc. Natl. Acad. Sci. U.S.A., 101 (2004) 3747.
[10] Barrat A., Barthelemy M. and Vespignani A., Phys. Rev. Lett., 92 (2004) 228701.
[11] Bianconi G. and Barabási A.-L., Europhys. Lett., 54 (2001) 436.
[12] Holme P. and Kim B. J., Phys. Rev. E, 65 (2002) 026107.
[13] Jackson M. O., Group Formation in Economics: Networks, Clubs and Coalitions (Cambridge University Press) 2005, Chapt. 1, pp. 11–57.
[14] Moriano P. and Finke J., arXiv:1110.0751v5 [physics.soc-ph], http://arxiv.org/abs/1110.0751.
[15] Redner S., Eur. Phys. J. B, 4 (1998) 131.
[16] Hall G. H., Jaffe A. B. and Trajtenberg M., The NBER patent citation data file: Lessons, insights and methodological tools, nBER Working Papers 8498, National Bureau of Economic Research (October 2001).
[17] Fowler J. H. and Jeon S., Soc. Netw., 30 (2008) 16.