

G OPEN ACCESS

Citation: Pandey PK, Singh M (2019) MSP-N: Multiple selection procedure with '*N*' possible growth mechanisms. PLoS ONE 14(12): e0224383. https://doi.org/10.1371/journal. pone.0224383

Editor: Federico Botta, University of Warwick, UNITED KINGDOM

Received: May 10, 2019

Accepted: October 11, 2019

Published: December 12, 2019

Copyright: © 2019 Pandey, Singh. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The US Powergrid dataset is available at (http://konect.uni-koblenz.de/ networks/opsahl-powergrid). The Arxiv HEP-TH is avialable at (https://snap.stanford.edu/data/cit-HepTh.html).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

RESEARCH ARTICLE

MSP-N: Multiple selection procedure with '*N*' possible growth mechanisms

Pradumn Kumar Pandey¹*, Mayank Singh²

1 Computer Science and Engineering, Indian Institute of Technology Roorkee, Roorkee, Uttrakhand, India, 2 Computer Science and Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat, India

* pandeypradumn16@gmail.com

Abstract

Network modeling is a challenging task due to non-trivial evolution dynamics. We introduce multiple-selection-procedure with 'N' possible growth mechanisms (*MSP-N*). In *MSP-N*, an incoming node chooses a single option among *N* available options to link to pre-existing nodes. Some of the potential options, in case of social networks, can be standard preferential or random attachment and node aging or fitness. In this paper, we discuss a specific case, *MSP-2*, and shows its efficacy in reconstructing several non-trivial characteristic properties of social networks, including networks with power-law degree distribution, power-law with an exponential decay (exponential cut-off), and exponential degree distributions. We evaluate the proposed evolution mechanism over two real-world networks and observe that the generated networks highly resembles the degree distribution of the real-world networks. Besides, several other network properties such as high clustering and triangle count, low spectral radius, and community structure, of the generated networks are significantly closer to the real-world networks.

Introduction

We witness a variety of complex social systems and non-trivial interactions among actors of the network [1]. In real-world networks, actors are represented as nodes, and interactions among actors are represented as edges. The definition of nodes is contextual, for example, in World Wide Web (WWW) network [2], web pages are considered as the nodes while in protein-protein interaction network [3, 4], proteins act as nodes. This diversity results in the non-trivial distribution of fundamental properties including degree distribution [5, 6], clustering coefficient, triangle distribution [7], small-world property [8–14], low average path length, assortativity [15–17], and community structure [18]. The *degree* of a node in a network is the number of connections it has w other nodes, and the *degree distribution* is the probability distribution of these degrees over the whole network. The degree distribution varies from power-law (a.k.a., *scale-free*) [1, 19, 20] to exponential. Similarly, a *clustering coefficient* is a measure of the extent to which nodes in a graph tend to group together [21]. Clustering coefficient varies from meager value to very high-value [1]. Assortativity represents a tendency of nodes to

connect to other nodes that posses similar properties, for example, the tendency of actors to connect others having similar degree. In literature, for different networks, assortativity varies from negative (disassortative mixture) to positive values (assortative mixture) [15-17]. A network is considered to have *community structure* if the nodes of the network can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally. Modularity [22] is often used in optimization methods for detecting community structure in networks.

In past, several classical growth mechanisms [8, 23–33] have been proposed to explain the network properties. Some of these interesting hypotheses include scaling behavior of the degree distribution, node aging, cost of link formation, randomness and preferential attachment in link formation. Interestingly, in 1955, Herbert Simon [34] confirmed the existence of *'rich get richer'* phenomena leading to power-law tail in degree distribution of several real-world networks. Later, Derek De Solla Price [35] proposed a similar idea in the context of bibliographic networks. Albert and Barabási [8] (hereafter *BA model*) proposed a degree-based preferential attachment is that the probability (p_{ij}) of connecting a newborn node *i* to an older (pre-existing) node *j* is an affine function of the degree of node *j* given by

$$p_{ij} = \frac{k_j + x_j}{\sum_j k_j + x_j}.$$
(1)

where x_j is a constant and k_j is the degree of node j. If N denotes the total number of nodes, then $l \in [1, N]$. The BA model proposed a constant value (= 3) for the power-law exponent with the clustering coefficient ($\mathcal{O}(1/n)$) vanishing as the network grows. However, majority of the real-world networks possess a wide range of exponent values and non-zeros clustering coefficients. Some of the interesting real-world networks that follow power-law distribution are phones call graphs [36], Internet [37], Web [8, 38, 39], click-stream data [40], who-trustswhom social network [41], etc.

Similarly, random attachment mechanisms emphasize on the uniform-random attachment of nodes and edges. Empirically, they generate networks with lower clustering coefficient along with Binomial or Poisson degree distribution [1]. In 1998, Watts and Strogatz [14] proposed a variant based on the random rewiring of a regular network. This mechanism generates networks with higher clustering along with a bell-shaped degree distribution. Additionally, several other network models have been proposed in the literature that capture different statistical and spectral properties of real-world networks [8, 42–48]. We claim that these mechanisms do not perform well in isolation with each other. Thus, we encompass above mechanisms in a more generic network growth model based on *Multiple Selection Procedure* (MSP).

In real-world networks, a node can interact with other nodes in more than one possible linking mechanism. A simple analogy is people traveling from one place to another using different modes of travel. Each person (node) may use a different mode of traveling, depending on his financial conditions, comfort level, age, popularity, and delay estimate. Similarly, MSP-N assumes the availability of N-possible linking mechanisms. An incoming node chooses one out of N mechanisms.

Fig 1 shows the graphical representation of MSP-N. Specifically, we derive and investigate MSP-2 that encompasses preferential attachment [8], cost of linking [29, 30], local dynamics [31], and aging [32].

We show that MSP-2 leads to non-trivial characteristic features of a larger class of social networks, including networks with power-law degree distribution, power-law with an



Fig 1. Graphical representation of Multiple-Selection-Procedure (MSP-N). Each incoming node can choose one out of N mechanisms for link formation.

https://doi.org/10.1371/journal.pone.0224383.g001

exponential decay (exponential cut-off), and exponential degree distributions. The generated class of scale-free networks (power-law degree distribution) shows power-law exponent $(1 + 1/\beta)$ in the range $(2, \infty)$ and phase transition in average connectivity, derived as $(2\beta - 1)$, at $\beta = 1/2$. We also present bounds on the average connectivity of the networks under different settings of model parameters. We demonstrate the high similarity between generated networks against real-world networks. The simulated networks show high clustering, slow growth of hubs (high degree nodes), edge densification, and community structure—presenting a good resemblance with real-world networks. We also show that the edge densification restricts the growth of the diameter in a random network.

Materials and methods

Real networks

We leverage two real-world network datasets to evaluate the generative ability of our proposed model by fitting the parameter values corresponding to each dataset. Intuitively, the aim is to find whether proposed models can mimic real-world networks efficiently. The two datasets are:

1. High Energy Physics-theory citation network (**ca- HepTh**) [49]: It is a collaboration network of scientists working in High Energy Physics-theory field. The network consists of authors as nodes and co-authorship relation between authors as an edge. It contains 8, 638 nodes and 24, 806 edges.

2. Power-grid Network (**PGN**) [14]: It is an electricity transport network in which a node represents either a generator, a transformer or a substation and an edge corresponds to a power supply line between two nodes. It contains 4, 941 nodes and 6, 594 edges.

Multiple selection procedure

We propose a network growth mechanical model based on the idea of multiple selection procedures (with N = 2). In particular, the model focuses on two plausible selection procedures closely resembling the observed processes in real-world networks. We combine these real-world processes in an MSP framework leading to a more realistic and generalized evolution process of networks. An incoming node in the system connects to the existing nodes in the network based on any of the following two selection procedures:

- 1. Preferential attachment with aging.
- 2. Random attachment with local growth.

Preferential attachment with aging. The first procedure combines degree-based preferential attachment with self aging. At each time-step (t), a new node enters the system. An already existing node (i) with degree $k_i(t)$ attracts the new node j due to its preferential ability (π) as

$$\pi_i(t) = \frac{k_i(t)}{t} \tag{2}$$

In contrast to the above preferential mechanism, the self-aging mechanism restricts the ability of an existing node to attract incoming nodes. For example, in a paper citation network, older papers receive fewer citations than similar newer papers due to field obsolescence [50, 51]. A young node possesses more ability to attract incoming nodes than an older node [52]. Similar growth behavior was observed by Dorogovtsev et al. [24]. They found that the aging is proportional $\tau^{-\alpha}$, where τ is the age of the node and α is the aging exponent. They also claimed that the network shows scaling behavior only in the region $\alpha > 1$. Even though aging function can exist in several possible mathematical forms, we propose a novel variant parameterized by the current degree of the node. The intuition lies in the hypothesis that the willingness of a node to accept new connections decreases as the current number of neighbors increases. The self-aging function $f_i(t)$ is a non-increasing function of time (t) defined below

$$f_i(t) = \frac{1}{1 + b_i k_i(t)},$$
(3)

where, b_i is a positive constant that controls the rate of the aging of node *i*. As evident from Eq (3), $f_i(t + 1) \le f_i(t) \forall t$ and $f_i(t) < 1$, $\forall i$ and $t \in [1 \infty)$. The self-aging restricts the growth in the number of connections of a node as time advances. More specifically, self-aging restricts the growth of hubs in a network. Next, we combine preferential attachment (described in Eq (2)) with self-aging expression (described in Eq (3)) in single formulation $F_i(t)$ given by

$$F_i(t) = \pi_i(t) * f_i(t) = \frac{k_i(t)}{(1+b_ik_i(t))t} = \frac{k_i}{(1+bk_i)t}.$$
(4)

The rightmost derivation in Eq (4) simplifies the nomenclature by replacing $k_i(t)$ with k_i and b_i with b (assuming that all nodes have similar aging rate). Next, we describe the second procedure.

Random attachment with local growth. The second procedure accounts random attachment with local growth. It facilitates attachment in two scenarios; direct (DRA) and indirect random attachment (IDRA). In DRA, the incoming node (*i*) attaches to an already existing node (*i*), randomly. Thus each existing node at time-step (*t*) attracts the incoming node with equal probability ($=\frac{1}{t}$). In IDRA, an initial attachment to a neighbor of node *i* favors link formation with *i* (similar to the link-copying mechanism proposed by Kumar et al. [53]). However, due to limited link formation capacity (of node *i*), the IDRA probability decreases as the degree (of node *i*) increases. We also term IDRA as "random attachment mechanism with local growth". The above two scenarios are combined in a single formulation $\phi_i(t)$ given by

$$\phi_i(t) = \text{DRA} + \text{IDRA},\tag{5}$$

$$\phi_i(t) = \frac{1}{t} + \frac{g_i(t)}{t} \left(1 - \frac{1}{t} \right),$$
(6)

For larger value of t, $\left(1 - \frac{1}{t}\right) \rightarrow 1$

$$\phi_i(t) = \frac{1}{t} + \frac{g_i(t)}{t},\tag{7}$$

where 1/t denotes DRA probability, $g_i(t)$ is a non-negative real-valued function which associates cost of linking with the local growth. Even though local growth dynamics helps in increasing the concentration of triangles in the network, the cost of link formation restricts the growth of the degree of nodes.

A plausible cost function. Consider a node *i* placed under IDRA process. Node *i* has k_i neighbours (degree). Node *j* joins the network formation process at time (t + 1). Node *j* connects to a neighbour of node *i* with probability 1/t. Later, *j* attempts to connect with *i*. The probability of a connection between node *i* and *j* depends on the linking cost. We argue that this indirect linking cost is inversely proportional to the current degree $\left(\propto \frac{1}{k_i(t)}\right)$. The intuition lies in the capacity constraints in link formation [54]. In general, researchers have shown that the formation of a link can inhibit the formation of another one, typically due to time, space, or capacity constraints [55]. The process is explained in Fig 2. Node *i* has k_i chances to receives



Random attachment

Local growth having cost of linking

Fig 2. Random attachment with local growth having cost of linking. Sub-figure (a) is showing the random attachment of node *j* to a neighbour of node *i* by probability 1/t (first step). In sub-figure (b), after connecting to a neighbour of node *i* in the first step, node *j* has linking cost inversely proportional to the degree of node *i* under local growth.

new links due to the local growth process of its neighbors. This results in

$$g_i(t) = \alpha \sum_{x=1}^{k_i} \frac{1}{k_i} = \alpha \tag{8}$$

where α (>0) is a proportionality constant for a network. Different networks might exhibit different values of α .

Consider a new node *j* randomly connects with an older node *r*, then attempts to form a link with first neighbors of node *r*. Let node *i* is one of the first neighbors of node *r* then cost of linking of nodes *i* and *j* depends on $1/k_i$ which tells that higher degree node has higher cost or low probability of being connected to a new node under local growth scheme. At the same time in the same process of link formation, for an older node *i*, $1/k_i$ is fixed for all the possibilities when a new node joins one of its neighbors and tries to connect with node *i* with probability $1/k_i$ and collectively results in a constant value α . In another way, we can understand the cost of linking to an older node. Constant α tells that an older node of higher degree has higher chances to be connected with a new joining node as compared to an older node of low degree but in both the cases gain is same which signifies the high linking cost of higher degree node.

MSP-2: The proposed network growth model. We combine the above two selection procedures in an MSP framework (hereafter *MSP-2*). The combined effect is reformed in terms of a differential equation using standard mean-field approximation [56] given by

$$\frac{\partial k_i}{\partial t} = \beta_j F_i(t) + (1 - \beta_j)\phi_i(t).$$
(9)

where $\beta_j \in (0, 1)$. Here $F_i(t)$ and $\phi_i(t)$ corresponds to expected gain in degree of node *i* due to preferential attachment with aging and random attachment with local growth, respectively. To keep the notation simple, we assume that all incoming nodes have same β , thus $\beta_j = \beta \forall j$.

The generative algorithm

- 1. Consider an initial connected network of m_0 nodes.
- 2. A new node *j* joins the network.
- 3. Toss coin with head probability $\beta_i = \beta$,
- If *head*: node *j* chooses preferential attachment mechanism with self-aging. The probability of connecting a node *i* in the pre-existing network is given by $\frac{k_i}{(1+bk_i)t}$.
- If *tail*: node *j* randomly selects a node *i* and gets connected to it. After that node *j* connects with the first neighbors of the node *i* according to the probability $\alpha/k_l \leq 1$, where *l* is a neighboring node of the node *i*.
- 4. Repeat steps 2and3 until the network has desired number of nodes.

Results

Next, we analyze the degree distribution, average connectivity, and edge densification. We also explore various structural and spectral measures such as triangle count, modularity structure and spectral radius, by theoretical calculations, and numerical validations.

The degree distribution

We, first of all, derive the generic degree distribution by combining Eqs (4), (7), (8) and (9), that results in the following formulation

$$\frac{\partial k_i}{\partial t} = \beta \frac{1}{1+bk_i} \frac{k_i}{t} + (1-\beta) \frac{1+\alpha}{t}.$$
(10)

rearranging above equation results in

$$\left(\frac{k_i + 1/b}{c_1(k_i + c_2)}\right)\frac{\partial k_i}{\partial t} = \frac{1}{t},\tag{11}$$

where $c_1 = \beta/b + (1 - \beta)(1 + \alpha)$ and $c_2 = \frac{(1 - \beta)(1 + \alpha)}{\beta + b(1 - \beta)(1 + \alpha)}$.

$$\frac{1}{c_1} \left(\frac{k_i + c_2 + 1/b - c_2}{k_i + c_2} \right) \frac{\partial k_i}{\partial t} = \frac{1}{t},\tag{12}$$

$$\begin{split} &\frac{1}{c_1}\left(1+\frac{1/b-c_2}{k_i+c_2}\right)\frac{\partial k_i}{\partial t}=\frac{1}{t},\\ &\frac{1}{c_1}\left(1+\frac{\gamma}{k_i+c_2}\right)\frac{\partial k_i}{\partial t}=\frac{1}{t}, \end{split}$$

where

$$\gamma = 1/b - c_2$$

After solving the PDE given above

$$\frac{1}{c_1} \left(k_i - k_i^0 + \gamma \ln \frac{k_i + c_2}{k_i^0 + c_2} \right) = \ln \frac{t}{t_i}$$
(13)

where k_i^0 is the initial degree of node *i* which appears at time t_i .

$$k_{i} - k_{i}^{0} + \gamma \ln \frac{k_{i} + c_{2}}{k_{i}^{0} + c_{2}} = c_{1} \ln \frac{t}{t_{i}}$$
(14)

$$\ln e^{k_i - k_i^0} + \gamma \ln \frac{k_i + c_2}{k_i^0 + c_2} = c_1 \ln \frac{t}{t_i}$$
$$\ln \frac{e^{k_i}}{e^{k_i^0}} + \ln \left(\frac{k_i + c_2}{k_i^0 + c_2}\right)^{\gamma} = \ln \left(\frac{t}{t_i}\right)^{c_1}$$
$$\ln \frac{e^{k_i}(k_i + c_2)^{\gamma}}{e^{k_i^0}(k_i^0 + c_2)^{\gamma}} = \ln \left(\frac{t}{t_i}\right)^{c_1}$$
$$e^{k_i}(k_i + c_2)^{\gamma} \frac{1}{e^{k_i^0}(k_i^0 + c_2)^{\gamma}} = \left(\frac{t}{t_i}\right)^{c_1}.$$

By the law of large numbers, at large t, $e^{k_i^0}(k_i^0 + c_2)^{\gamma} \rightarrow \langle e^{k_i^0}(k_i^0 + c_2)^{\gamma} \rangle = \eta^{c_1}$ (average of expected initial degree of nodes), where η is a positive constant.

$$e^{k_i}(k_i+c_2)^{\gamma}rac{1}{\eta^{c_1}}=\left(rac{t}{t_i}
ight)^{c_1}.$$

 $k < k_i$

Consider,

$$\begin{split} e^k (k+c_2)^{\gamma} \frac{1}{\eta^{c_1}} &< e^{k_i} (k_i+c_2)^{\gamma} \frac{1}{\eta^{c_1}}, \\ e^k (k+c_2)^{\gamma} \frac{1}{\eta^{c_1}} &< \left(\frac{t}{t_i}\right)^{c_1}, \\ t_i &< \eta e^{-k/c_1} (k+c_2)^{-\gamma/c_1} t. \end{split}$$

As the network is growing uniformally, the probability of selecting a node t_i at time t, $P(t_i) = 1/(t + m_0)$, where m_0 is the size (number of nodes) of initial seed network [56], and

$$P(k_i > k) = P(t_i < \eta e^{-k/c_1} (k + c_2)^{-\gamma/c_1} t)$$

= $\frac{t}{m_0 + t} \eta e^{-k/c_1} (k + c_2)^{-\gamma/c_1}$

As $t \to \infty$,

$$P(k_i > k) = \eta e^{-k/c_1} (k + c_2)^{-\gamma/c_1},$$
(15)

The above equation represents a class of networks that demonstrate power-law degree distribution with exponential cut-off. Specifically, the two classes can be derived as follows:

• Exponential degree distribution: Assume that if $\beta \to 0$, then $c_2 \to \frac{1}{b}$ and $\gamma = 1/b - c_2 \to 0$. The cumulative degree distribution expression Eq.(15) reduces to

$$P(k_i > k) \to \eta e^{-k/(1+\alpha)},\tag{16}$$

Eq (16) refers to the networks of exponential degree distribution with $k_j^0 \rightarrow (1 + \alpha)$ and average degree $\bar{k} \rightarrow 2(1 + \alpha)$.

• Scale-free degree distribution: Assume that if b = 0, then from Eq (10)

$$\frac{\partial k_i}{\partial t} = \beta \left(\frac{k_i}{t}\right) + (1 - \beta) \left(\frac{1 + \alpha}{t}\right),\tag{17}$$

and

$$P(k_i > k) \propto (k + (1 - \beta)(1 + \alpha)/\beta)^{-1/\beta}.$$
 (18)

Eq (18) refers to the class of scale-free networks that follow power-law degree distribution

with $k_i^0 = \beta \bar{k}_t + (1 - \beta)(1 + \alpha)$ and

$$\bar{k}_{t+1} = \frac{t\bar{k}_t + 2k_j^0}{t+1} = \frac{t\bar{k}_t + 2\beta\bar{k}_t + 2(1-\beta)(1+\alpha)}{t+1},$$

$$\bar{k}_{t+1} = (t+1)^{2\beta-1}(1+\theta) - \theta, \tag{19}$$

where \bar{k}_t is the average connectivity of the network at time t, $\theta = \frac{2(1-\beta)(1+\alpha)}{(2\beta-1)}$ and $\beta \in (0, 1)$. This represents a densification power law (DPL) [57], exhibiting phase transition at $\beta = 1/2$. For $\beta > 1/2$, the average degree increases as the network evolves over the period of time and for $\beta \leq 1/2$ average degree of evolving network approaches the limiting value of $2(1 - \beta)(1 + \alpha)/(1 - 2\beta)$ as $t \to \infty$ (see Fig 3).

The average connectivity

Next, we investigate the dynamics of the average connectivity of MSP-2. At each time step, a node arrives and attaches itself to pre-existing nodes in the network. Consider node *j* being introduced in the network formation process at the time step (t + 1). From Eq.(10), the initial



Fig 3. Average degree, $\bar{k}(t)$, is plotted for different size of networks simulated under MSP-2 by setting $\beta = 0.4$ (plot in dots) and $\beta = 0.6$ (plot in circles), and $\alpha = 1$ in both the cases. For $\beta = 0.4$ average degree reaches to steady state (constant) and for $\beta = 0.6$ (> 0.5) average degree increases continuously.

degree of node j is given by

$$k_j^0 = \sum_{i \in N_t} \left(\beta \left(\frac{1}{1 + bk_i} \frac{k_i}{t} \right) + (1 - \beta) \left(\frac{1 + \alpha}{t} \right) \right). \tag{20}$$

Average connectivity of the network at time t + 1 will be

$$\begin{split} \bar{k}_{t+1} &= \frac{t\bar{k}_t + 2k_j^0}{t+1}, \\ \bar{k}_{t+1} &= \frac{t\bar{k}_t + 2(1-\beta)(1+\alpha) + 2\beta/b - \frac{2\beta}{b}\sum_i \frac{1}{(bk_i+1)t}}{t+1}, \\ \bar{k}_{t+1} &\geq \frac{t\bar{k}_t + 2(1-\beta)(1+\alpha) + 2\beta/b - \frac{2\beta}{b}\sum_i \frac{1}{(1+b)t}}{t+1}, \\ \bar{k}_{t+1} - \bar{k}_t &\geq \frac{-\bar{k}_t + 2\frac{\beta}{b} + 2(1-\beta)(1+\alpha) - \frac{2\beta/b}{1+b}}{t+1}, \\ \frac{\partial\bar{k}_{t+1}}{\partial t} &\geq \frac{-\bar{k}_t + 2\frac{\beta}{b} + 2(1-\beta)(1+\alpha) - \frac{2\beta/b}{1+b}}{t+1}. \end{split}$$
(21)

The solution of Eq(21) is given by,

 \bar{k}

$$\bar{k}_{t+1} \ge (t+1)^{-1}(1-\bar{k}_L) + \bar{k}_L,$$

where $\bar{k}_L = \frac{2\beta}{1+b} + 2(1-\beta)(1+\alpha)$. As $t \to \infty \Rightarrow \bar{k}_\infty \to \bar{k}_L$. Similarly,

$$k_{t+1} \le (t+1)^{-1}(1-k_U) + k_U$$

where $\bar{k}_U = 2\beta/b + 2(1-\beta)(1+\alpha)$. As $t \to \infty \Rightarrow \bar{k}_\infty \to \bar{k}_U$. \bar{k}_L and \bar{k}_U are lower and upper bounds of the average connectivity of the network model defined by Eq.(10).

Edge densification

Edge densification prevents the growth of the diameter of networks generated under the proposed model. Consider two networks with same number of nodes corresponding to β_1 and β_2 ($\geq \beta_1$) (discussed in Eq (19)), MSP-2 generate networks with different average connectivity and different diameter growth rate. Let D_1 and D_2 be the diameters of the networks generated under Eq (19) by setting $\beta = \beta_1$ and β_2 , respectively, such that $\beta \geq 1/2$ and b = 0. The proposed modelling scheme results in networks with $D \propto (1 - \beta)$ (proof, similar to [33], is omitted due to space constraints). Again consider

$$\begin{split} \beta_1 &\leq \beta_2 \Rightarrow k(\beta_1) \leq k(\beta_2), \\ \beta_1 &\leq \beta_2 \Rightarrow 1-\beta_1 \geq 1-\beta_2 \Rightarrow D_1 \geq D_2. \end{split}$$

So, if β has different values and keeping the rest of the parameters constant in Eq. (19), then

$$k(eta_1) \leq k(eta_2) \Rightarrow eta_1 \leq eta_2 \Rightarrow D_1 \geq D_2.$$

Modularity or community structure

Real-world networks inherit community structure. Community is a group of nodes possessing more number of links than expected in random networks [18]. In the context of social structure, a community is a group of similar people having a significantly high number of

connections within the community and lesser connections to the outside world. To measure the quality of community structure inside a network, modularity index Q [18] is defined as

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta_{c_i c_j}$$

where *m* is the number of edges in the network, k_i and k_j are the degrees of the nodes *i* and *j* which belong to communities c_i and c_j respectively, and δ is the Kronecker delta function.

Next, we conduct the theoretical analysis of the possibility of community structure in MSP-2. We argue that similar results hold for higher-order MSP variations (N > 2). Consider the connection probability p_{ij} between two nodes *i* and *j* which appear at time t_i and $t_j(>t_i)$, respectively

$$p_{ij} = \beta \frac{k_i}{1+bk_i} \frac{1}{t} + (1-\beta) \frac{1+\alpha}{t}.$$

As we know that at each time a new appears, so $t = t_j$ and we get

$$p_{ij} = \beta \frac{k_i}{1+bk_i} \frac{1}{t_j} + (1-\beta) \frac{1+\alpha}{t_j}.$$

Here, $(1 + bk_i)$ is applied to restrict the growth of the degree of node *i*. Hence, without loss of generality, we consider a simpler case, where b = 0. We claim that similar results true for other cases of the proposed model. We consider

$$p_{ij} = \beta \frac{k_i(t_j)}{t_j} + (1 - \beta) \frac{1 + \alpha}{t_j}.$$
(22)

and after solving PDE Eq (17), we have

$$k_{i}(t_{j}) = c_{i}^{0} \left(\frac{t_{j}}{t_{i}}\right)^{\beta} - \frac{(1-\beta)(1+\alpha)}{\beta},$$

$$c_{i}^{0} = \left(k_{i}^{0} + \frac{(1-\beta)(1+\alpha)}{\beta}\right).$$
(23)

Let *A* be the adjacency matrix associated with a network *G*, where $A_{ij} = 1$, if nodes *i* and *j* are connected otherwise 0. Corresponding *null-model* is defined in the following way: if k_i and k_j are the degrees of nodes *i* and *j*, respectively, and *m* is the number of edges in the network *G* then $k_ik_j/2m$ is the expected value of A_{ij} under null-model and community structure is measured by the non-zero heterogeneous values of $A_{ij} - \frac{k_ik_j}{2m}$. Here, we have only expected values of A_{ij} of a network obtained under MSP-2 that is $E[A_{ij}] = p_{ij}$. Using Eqs (22) and (23), we have

$$E[A_{ij}] = \frac{\beta c_i^0}{t_i^\beta t_j^{1-\beta}}$$
(24)

and

$$\frac{k_{i}k_{j}}{2m} = \begin{cases} \frac{c_{i}^{0}c_{j}^{0}}{(1+\theta)t_{j}^{\beta}t_{i}^{\beta}} & \text{if } \beta \ge 1/2\\ \\ \frac{c_{i}^{0}c_{j}^{0}}{\theta t_{j}^{\beta}t_{i}^{\beta}t^{1-2\beta}} & \text{if } \beta < 1/2 \end{cases}$$
(25)



Fig 4. Modularity index, Q, is plotted for different size of networks generated by the proposed model with different values of β . Horizontal-axis represents the number of nodes in the networks and the vertical-axis represents modularity index of the networks.

https://doi.org/10.1371/journal.pone.0224383.g004

where c_i^0 and c_j^0 are initial conditions. As we know that β controls the contribution of two micro-level network growth processes in the resulting evolution of the network, so we analyze the effect of the parameter β over the strength of community structure of the obtained network under the proposed model. We consider the condition when β approaches to 1 (main contribution made by preferential attachment scheme), and have

$$p_{ij} = E[A_{ij}] = \frac{\beta c_i^0}{t_i^\beta t_j^{1-\beta}}$$
(26)

and as $\beta \to 1 \Rightarrow c_j^0 \to k_j^0 \to \beta \bar{k}_{t_j} \to t_j^{2\beta-1}$, $\theta \to 0$ that results

$$\frac{k_i k_j}{2m} \sim \frac{\beta c_i^0}{t_i^\beta t_j^{1-\beta}} \tag{27}$$

From Eqs (26) and (27), when β approaches to 1, the distribution of links (p_{ij}) in the network simulated under the proposed model approaches to null model $\binom{k_i k_j}{2m}$ and strength of community structure (*Q*) gets reduced, see Fig 4.

Again we consider another condition when β approaches to 0, and have

$$p_{ij} = E[A_{ij}] = \frac{1+\alpha}{t_j} \tag{28}$$

and

$$\frac{k_i k_j}{2m} \sim \frac{1+\alpha}{t} \log \frac{t}{t_i} \log \frac{t}{t_i} \tag{29}$$

From Eqs (28) and (29), when β approaches to 0, the distribution of links (p_{ij}) are different from null model $(\frac{k_i k_j}{2m})$ and strong community structure appears, see Fig 4.

Modularity index Q is defined over matrix B, where

$$B_{ij} = A_{ij} - \frac{k_i k_j}{2m} \sim E[A_{ij}] - \frac{k_i k_j}{2m} = p_{ij} - \frac{k_i k_j}{2m}.$$

We discuss the strength of modularity in terms of positive eigenvalues of matrix *B*. We have already shown two extreme ends when $\beta = 1$ and $\beta = 0$. As β approaches to 1 $B_{ij} = p_{ij} - \frac{k_i k_j}{2m}$ approaches to 0 and maximum eigenvalue approaches to 0 and weak community structure arises. In another direction, when β approaches to 0 we have heterogeneous values of $B_{ij} = p_{ij} - \frac{k_i k_j}{2m}$, and eigenvalues spread over a wide range (it is true due to constant volume of matrix *B*) and gets higher maximum eigenvalue of *B*, close to 1 (strong community structure). While finding communities in a network, we decompose the network into small sub-networks that is equivalent to decomposition of matrix *B* in such a way that diagonal of *B* has blocks of positive entries and off-diagonal blocks have maximum negative entries, and it can happen only when *B* has heterogeneous structure which appears when β has lower value. We can conclude that local dynamics is responsible for community structure.

The distribution of expected links in the proposed model is different from the corresponding distribution of $k_i k_i/2m$ values. If the probability of the existence of a link between nodes i and j is different from the value $k_i k_i / 2m$, which is the probability of the existence on a link between nodes *i* and *j* under null model, then the structure of the network obtained under the model would be different as compared to the expected null structure (under null model) for given degree sequence. This leads to the existence of community structure in the network (higher value of maximum eigenvalue of B explained in the previous paragraph). As the β reduces, the heterogeneity of $E[A_{ij}] - \frac{k_i k_j}{2m}$ increases and network generated under MSP-2 shows strong community structure. Fig 4 demonstrates the modularity index of the networks generated by the proposed model. It clearly shows the existence of the community structure. The contribution of local dynamics, $(1 - \beta)$, in the network evolution, affects the community structure of the resulting network positively. A smaller value of β does not indicate that the network is denser as we have the condition that if $(2\beta - 1)$ is greater than 1/2 then network would show densification. However, as β decreases $(1 - \beta)$ increases but $(2\beta - 1)$ decreases. Theoretical analysis and numerical simulation show similar results. As β decreases, the value of modularity index (Q) increases (see Fig 4). It is needless to mention that this phenomenon is common in various social networks [58]. We leverage Louvain algorithm [59] to detect community structures in networks.

Number of triangles

In a network, a triangle is a cycle of three nodes. The high concentration of triangles in a network is a fundamental property of many real networks. In real-world networks, social phenomena such as *"friends of a friend are friends"* beautifully explain the high concentration of triangles. Several social networks have a high density of triangles, for example, ego-Facebook network, ego-Gplus network, and ego-Twitter [7].

- Node *j*, first, connects to one of neighbours of node *i* with probability 1/*t*. Later, *j* connects to *i* with probability α/k_i . Thus, node *i* has k_i chances to form a triangle, each with probability $\frac{\alpha}{k_i}$.
- Similarly, node *j* can first link to node *i* with probability 1/t and then connects with the neighbours of *i* with probability α/k_b where node *l* is the neighbour of node *i* other than *j*.

Next, we compute expected number of triangles generated by above two scenarios at time *t*. At each time step, network generates atleast *Tr* triangles, given by

$$Tr(t) = \sum_{i=1}^{t} \frac{(1-\beta)\alpha}{t} + \sum_{i=1}^{t} \frac{(1-\beta)\alpha}{t} \sum_{l=1}^{t} A_{il} \frac{1}{k_l} = 2(1-\beta)\alpha.$$

The expected number of total triangles in the network is given by

$$\Delta_t \geq 2(1-\beta)\alpha t.$$

The number of triangles in a network produced by the proposed model is lower bounded by a linear function which has slope $2(1 - \beta)\alpha$, and β controls the density of triangles in the model generated networks.

Simulation experiments validate the above theoretical bounds. We generate networks using MSP-2 by setting $\beta = 0.1$, $\alpha = 1$ and b = 1. The theoretical and numerical results are plotted in Fig 5. The theoretical result provides a good estimate of numerical values for triangle count in the networks generated under MSP-2.

Spectral radius

The largest eigenvalue of an adjacency matrix associated with a network is known as the spectral radius (SR) of the network. Empirically, the reciprocal of SR quantifies the threshold of viral propagation in the network [60]. The networks with smaller spectral radius have larger robustness against the spread of viruses [60]. We derive bounds on SR for the networks produced under MSP-2. The bounds on SR can be leveraged to attain bounds on diffusion threshold. Let $\lambda_1(A)$ is the largest eigenvalue of the adjacency matrix *A* associated with a network produced by MSP-2. From [61], we know that

$$\varepsilon_{\max} \ge \lambda_1(A) \ge \bar{k},$$
(30)

where k_{max} is the maximum degree and \bar{k} is the average degree of the network.

By Eq (13), we get the expected value of maximum degree in a network of size *n* nodes obtained under MSP-2 by setting $k_i^0 = 1$, $t_i = 1$ and t = n that is

$$k_{\max} + \gamma \ln (k_{\max} + c_2) = 1 + \gamma \ln (1 + c_2) + c_1 \ln n.$$

After simplifications and approximations, we get

$$1 + c_1 \ln n \ge k_{\max} \ge \frac{c_1}{1 + \gamma} \ln n. \tag{31}$$



Fig 5. Number of triangles (Δ) is plotted for different size of networks generated by the proposed model by setting $\beta = 0.1$, b = 1, and $\alpha = 1.0$. Theoretically calculated lower bound $(2(1 - \beta)\alpha)$ is also plotted. Horizontal-axis represents the number of nodes in the networks and the vertical-axis represents number of triangles in the networks.

https://doi.org/10.1371/journal.pone.0224383.g005

By Eqs (30) and (31) and lower bound on average connectivity \bar{k}_{L} (Eq (21))

$$1 + c_1 \ln n \ge \lambda_1(A) \ge \frac{2\beta}{b+1} + 2(1-\beta)(1+\alpha).$$

SR of a growing network under MSP-2 has a growth rate lower than the logarithm of the size of the network. Using the bounds on SR, we can select the model parameters to generate a network of the desired property.

Reconstructing real networks

As discussed earlier, the majority of the real-world networks follow a combination of multiple degree distributions and not just power-law distribution [5, 6]. The neuronal network of the worm *C. elegans* and the power-grid network of Southern California shows exponential decay in cumulative connectivity of the nodes [5] while networks of scientific collaborators exhibit power-law with exponential cut-off [6].

Parameter tuning. Assume that we produce a model network corresponding to a real network which has Δ_r triangles, k_{max} maximum degree and *n* number of nodes. We know that, if β and α are model parameters then the network obtained under the proposed model would have at-least $\alpha(1 - \beta)n$ triangles. We consider the relation

$$\frac{\Delta_r}{n} = \alpha (1 - \beta), \tag{32}$$

to ensure that the reconstructed network has at-least Δ_r triangles. We have another relation

between maximum degree and the size of the network.

$$e^{k_{\max}}(k_{\max}+c_2)^{\gamma}\frac{1}{e(1+c_2)^{\gamma}}=n^{c_1}$$
(33)

First we discretize the interval [0, 1], and let say *S* be the set of those discrete points. Select $\beta \in S$ and using Eq (32) compute the value of α . Again, using the obtained values of α and β , we compute the value of *b* by solving Eq (33), numerically. Now, we simulate model for the computed values of (β , α , *b*) and select the model network which has minimum value of $|P(k_i \ge k) - P_r(k_i \ge k)|$ for the set *S*, where $P(k_i \ge k)$ and $P_r(k_i \ge k)$ are cumulative degree distributions of model network and given real network, respectively.

Here, we reconstruct two real networks described in Material and Methods section; one is a collaboration network (ca - HepTh), which is an example of power-law degree distribution with exponential cut-off [49] and second is a Power-Grid-Network (PGN) which has exponential degree distribution. Cumulative degree distribution of both the networks (in black squares) and corresponding model networks (in blue stars) are plotted in Fig 6a and 6b, respectively. The model parameters are tuned as $\beta = 0.15$, $\alpha = 3.8$ and b = 0.05 to generate a network corresponding to '*ca* – *HepTh*' network. Similarly, $\beta = 0.5$, $\alpha = 0.26$ and b = 0.1 are the parameter values to generate model network corresponding to PGN network. The overlapping of the plots in Fig 6a and 6b clearly shows that the MSP-2 generative mechanism is significantly capable enough to capture the degree distribution of different classes of real-world networks. In Fig 6a and 6b, the degree distribution of model networks are corresponding to a single snapshot. Apart from degree distribution, we compare other statistical properties such as triangles' count Δ , spectral radius SR, clustering coefficient CC, and modularity index Q of real networks and corresponding model networks. Results are tabulated in Table 1. It is observed that the proposed model produces networks which have statistical properties close to real data, for example, low-spectral radius, large triangles' count, clustering, and modularity index. It is more generalized as compared to the other previous growth models.



Fig 6. (a) Cumulative degree distribution of collaboration network (in black) and Model network (in blue). The parameter values are $\beta = 0.15$, $\alpha = 3.8$ and b = 0.05. (b) Cumulative degree distribution of power grid network (in black) and Model network (in blue). The parameter values are $\beta = 0.5$, $\alpha = 0.26$ and b = 0.1.

	ca-HepTh (MSP-2)	PGN (MSP-2)
SR	31 (20.3)	7.5 (5.9)
Δ	27869 (38251)	651 (646)
CC	0.48 (0.39)	0.08 (0.11)
Q	0.72 (0.66)	0.93 (0.75)

Table 1. Structural and spectral properties of real-world networks and corresponding synthetic networks obtained under our proposed model. Values in brackets represent proposed model properties.

https://doi.org/10.1371/journal.pone.0224383.t001

Discussion

The previous section demonstrates how well the MSP-2 mechanism can imitate real-world structural properties. Also, the proposed model is useful in generating several types of networks under different conditions depending on the settings of the parameters in Eq (10). Intuitively, we can generate the following three classes of networks conditioned on the selection procedure and parameter initialization.

- *Empty network*: If all incoming nodes consider only the first branch of the MSP-2 ($\beta = 1$) along with high aging factor ($b \rightarrow \infty$), then an empty network with a single edge will be generated. The generated network is closer to an extreme of the network structures (see Fig 7a).
- *Tree network*: If all incoming nodes consider only the second branch of the MSP-2 ($\beta = 0$) and the cost of linking under local growth is very high, then a tree network is generated (see Fig 7b).
- *Complete network*: Again, consider that the incoming nodes consider only the second branch of the MSP-2 ($\beta = 0$). If the cost of the local connection is zero, then it will generate a complete network (see Fig 7c). It is another extreme of the network topology of the possible network structures.



Fig 7. $\beta = 1$ and $b \rightarrow \infty$ (left most network). $\beta = 0$ and $g(t) \rightarrow 0$, very high cost of link formation (network in the middle). $\beta = 1$ and b = 0, or $\beta = 0$ and cost of link formation is zero (right most network).

If all incoming nodes only consider preferential attachment without aging ($\beta = 1$ and b = 0), then the generated network would be almost complete. However, node aging affects its growth during network evolution. $b \neq 0$ bounds the degree of the node and average degree of the network.

All the network structures lie between the two extremes, the empty network, and the complete network. Most of the real-world networks lie between the tree and complete network topology. The simulated results described in the previous section show the potential of MSP-2 to reconstruct real-networks from different classes structurally.

Conclusion

In this work, we propose a novel random network growth model (MSP-2) based on the observed phenomena of link formation. MSP-2 combines concepts of preferential attachment, random selection, aging, and the cost of link formation in a single process of network evolution. The combined dynamics leads to non-trivial properties exhibited by real-world networks. The proposed model successfully generates the networks corresponding to the real-world networks, like collaboration network and Power-Grid-Network. We find that the degree distribution of the real-world networks is significantly closer to the corresponding modeled networks. The properties of the networks generated under the proposed model are similar to the real-world networks. We also show that densification in a random network restricts the growth of diameter of that network. In the future, more generalized cost, and aging functions can be implemented using complex MSP-N (N>2) to incorporate various other physical phenomena.

Acknowledgments

The authors would like to thank Dr. Bibhas Adhikari, assistant professor in the Mathematics department of Indian Institute of Technology Kharagpur, for his valuable suggestions and discussions to improve the manuscript.

Author Contributions

Conceptualization: Mayank Singh.

Formal analysis: Pradumn Kumar Pandey.

Investigation: Mayank Singh.

Methodology: Pradumn Kumar Pandey, Mayank Singh.

Project administration: Pradumn Kumar Pandey.

Supervision: Pradumn Kumar Pandey.

Validation: Pradumn Kumar Pandey.

Visualization: Mayank Singh.

Writing - original draft: Pradumn Kumar Pandey.

Writing - review & editing: Mayank Singh.

References

- 1. Newman M. Networks: an introduction. Oxford University Press; 2010.
- 2. Barabási AL, Albert R, Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web. Physica A: statistical mechanics and its applications. 2000; 281(1):69–77.

- Asur S, Ucar D, Parthasarathy S. An ensemble framework for clustering protein–protein interaction networks. Bioinformatics. 2007; 23(13):i29–i40. https://doi.org/10.1093/bioinformatics/btm212 PMID: 17646309
- Wu J, Vallenius T, Ovaska K, Westermarck J, Mäkelä TP, Hautaniemi S. Integrated network analysis platform for protein-protein interactions. Nature methods. 2009; 6(1):75–77. https://doi.org/10.1038/ nmeth.1282 PMID: 19079255
- Amaral LAN, Scala A, Barthelemy M, Stanley HE. Classes of small-world networks. Proceedings of the national academy of sciences. 2000; 97(21):11149–11152. https://doi.org/10.1073/pnas.200327197
- Newman ME. The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences. 2001; 98(2):404–409. https://doi.org/10.1073/pnas.98.2.404
- Leskovec J, Mcauley JJ. Learning to discover social circles in ego networks. In: Advances in neural information processing systems; 2012. p. 539–547.
- Barabási AL, Albert R. Emergence of scaling in random networks. Science. 1999; 286(5439):509–512. https://doi.org/10.1126/science.286.5439.509 PMID: 10521342
- Bollobás B, Riordan O. The diameter of a scale-free random graph. Combinatorica. 2004; 24(1):5–34. https://doi.org/10.1007/s00493-004-0002-2
- Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, et al. Graph structure in the web. Computer networks. 2000; 33(1):309–320. https://doi.org/10.1016/S1389-1286(00)00083-9
- Chung F, Lu L. The average distances in random graphs with given expected degrees. Proceedings of the National Academy of Sciences. 2002; 99(25):15879–15882. https://doi.org/10.1073/pnas. 252631999
- 12. Kleinberg J. Small-world phenomena and the dynamics of information. Advances in neural information processing systems. 2002; 1:431–438.
- 13. Milgram S. The small world problem. Psychology today. 1967; 2(1):60-67.
- Watts DJ, Strogatz SH. Collective dynamics of 'small-world'networks. Nature. 1998; 393(6684):440– 442. https://doi.org/10.1038/30918 PMID: 9623998
- Martinez ND. Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. Ecological Monographs. 1991; p. 367–392. https://doi.org/10.2307/2937047
- Barabási AL, Albert R. Emergence of scaling in random networks. Science. 1999; 286(5439):509–512. https://doi.org/10.1126/science.286.5439.509 PMID: 10521342
- 17. Huxham M, Beaney S, Raffaelli D. Do parasites reduce the chances of triangulation in a real food web? Oikos. 1996; p. 284–300. https://doi.org/10.2307/3546201
- Newman ME, Girvan M. Finding and evaluating community structure in networks. Physical review E. 2004; 69(2):026113. https://doi.org/10.1103/PhysRevE.69.026113
- Barabási AL, et al. Scale-free networks: a decade and beyond. Science. 2009; 325(5939):412. https://doi.org/10.1126/science.1173299 PMID: 19628854
- Eguiluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV. Scale-free brain functional networks. Physical review letters. 2005; 94(1):018102. <u>https://doi.org/10.1103/PhysRevLett.94.018102</u> PMID: 15698136
- Luce RD, Perry AD. A method of matrix analysis of group structure. Psychometrika. 1949; 14(2):95– 116. https://doi.org/10.1007/bf02289146 PMID: 18152948
- 22. Li W, Schuurmans D. Modular community detection in networks. In: Twenty-Second International Joint Conference on Artificial Intelligence; 2011.
- Pandey PK, Adhikari B. A parametric model approach for structural reconstruction of scale-free networks. IEEE Transactions on Knowledge and Data Engineering. 2017; 29(10):2072–2085. <u>https://doi.org/10.1109/TKDE.2017.2725264</u>
- Dorogovtsev SN, Mendes JFF. Evolution of networks with aging of sites. Physical Review E. 2000; 62(2):1842. https://doi.org/10.1103/PhysRevE.62.1842
- Dorogovtsev S, Mendes J, Samukhin A. Size-dependent degree distribution of a scale-free growing network. Physical Review E. 2001; 63(6):062101. https://doi.org/10.1103/PhysRevE.63.062101
- Chakrabarti D, Faloutsos C. Graph mining: Laws, generators, and algorithms. ACM computing surveys (CSUR). 2006; 38(1):2. https://doi.org/10.1145/1132952.1132954
- Toivonen R, Onnela JP, Saramäki J, Hyvönen J, Kaski K. A model for social networks. Physica A: Statistical Mechanics and its Applications. 2006; 371(2):851–860. https://doi.org/10.1016/j.physa.2006.03. 050
- Kossinets G, Watts DJ. Empirical analysis of an evolving social network. Science. 2006; 311(5757):88– 90. https://doi.org/10.1126/science.1116869 PMID: 16400149

- Jackson MO. A survey of network formation models: stability and efficiency. Group Formation in Economics: Networks, Clubs, and Coalitions. 2005; p. 11–49.
- Slikker M, van den Nouweland A. Network formation models with costs for establishing links. Review of Economic Design. 2000; 5(3):333–362. https://doi.org/10.1007/PL00013692
- Bornholdt S, Rohlf T. Topological evolution of dynamical networks: Global criticality from local dynamics. Physical Review Letters. 2000; 84(26):6114. https://doi.org/10.1103/PhysRevLett.84.6114 PMID: 10991137
- Barabâsi AL, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T. Evolution of the social network of scientific collaborations. Physica A: Statistical mechanics and its applications. 2002; 311(3):590– 614.
- Pandey PK, Adhikari B. Context dependent preferential attachment model for complex networks. Physica A: Statistical Mechanics and its Applications. 2015; 436:499–508. https://doi.org/10.1016/j.physa. 2015.04.038
- Simon HA. On a class of skew distribution functions. Biometrika. 1955; 42(3/4):425–440. https://doi.org/ 10.2307/2333389
- Price DdS. A general theory of bibliometric and other cumulative advantage processes. Journal of the American society for Information science. 1976; 27(5):292–306. <u>https://doi.org/10.1002/asi.</u> 4630270505
- Abello J, Buchsbaum AL, Westbrook JR. A functional approach to external graph algorithms. In: Algorithms—ESA'98. Springer; 1998. p. 332–343.
- Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology. In: ACM SIGCOMM computer communication review. vol. 29. ACM; 1999. p. 251–262.
- Huberman BA, Adamic LA. Internet: growth dynamics of the world-wide web. Nature. 1999; 401 (6749):131–131. https://doi.org/10.1038/43604
- Kumar R, Raghavan P, Rajagopalan S, Tomkins A. Trawling the Web for emerging cyber-communities. Computer networks. 1999; 31(11):1481–1493. https://doi.org/10.1016/S1389-1286(99)00040-7
- Bi Z, Faloutsos C, Korn F. The DGX distribution for mining massive, skewed data. In: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining. ACM; 2001. p. 17–26.
- 41. Chakrabarti D, Zhan Y, Faloutsos C. R-MAT: A Recursive Model for Graph Mining. In: SDM. vol. 4. SIAM; 2004. p. 442–446.
- Christensen K, Donangelo R, Koiller B, Sneppen K. Evolution of random networks. Physical Review Letters. 1998; 81(11):2380. https://doi.org/10.1103/PhysRevLett.81.2380
- Krapivsky PL, Redner S, Leyvraz F. Connectivity of growing random networks. Physical review letters. 2000; 85(21):4629. https://doi.org/10.1103/PhysRevLett.85.4629 PMID: 11082613
- Ohtsuki H, Hauert C, Lieberman E, Nowak MA. A simple rule for the evolution of cooperation on graphs and social networks. Nature. 2006; 441(7092):502–505. https://doi.org/10.1038/nature04605 PMID: 16724065
- 45. Krapivsky PL, Redner S. Organization of growing random networks. Physical Review E. 2001; 63(6):066123. https://doi.org/10.1103/PhysRevE.63.066123
- Dorogovtsev SN, Goltsev AV, Mendes JFF. Pseudofractal scale-free web. Physical Review E. 2002; 65(6):066122. https://doi.org/10.1103/PhysRevE.65.066122
- Dorogovtsev SN, Goltsev AV, Mendes JFF. Ising model on networks with an arbitrary distribution of connections. Physical Review E. 2002; 66(1):016104. https://doi.org/10.1103/PhysRevE.66.016104
- Dorogovtsev SN, Mendes JFF. Effect of the accelerating growth of communications networks on their structure. Physical Review E. 2001; 63(2):025101.
- Leskovec J, Krevl A. SNAP Datasets: Stanford Large Network Dataset Collection; 2014. <u>http://snap.stanford.edu/data</u>.
- Singh M, Sarkar R, Goyal P, Mukherjee A, Chakrabarti S. Relay-linking models for prominence and obsolescence in evolving networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM; 2017. p. 1077–1086.
- Singh M, Sarkar R, Goyal P, Mukherjee A, Chakrabarti S. Sic Transit Gloria Manuscriptum: Two Views of the Aggregate Fate of Ancient Papers. arXiv preprint arXiv:151108310. 2015;.
- Zhu H, Wang X, Zhu JY. Effect of aging on network structure. Phys Rev E. 2003; 68:056121. <u>https://doi.org/10.1103/PhysRevE.68.056121</u>
- 53. Kumar R, Raghavan P, Rajagopalan S, Sivakumar D, Tomkins A, Upfal E. Random graph models for the web graph. In: FOCS; 2000. p. 57–65.

- König MD, Tessone CJ, Zenou Y. From assortative to dissortative networks: the role of capacity constraints. Advances in Complex Systems. 2010; 13(04):483–499. https://doi.org/10.1142/ S0219525910002700
- 55. Battiston P. Constrained Network Formation. Italian Economic Journal. 2016; 2(3):347–362. https://doi. org/10.1007/s40797-016-0040-0
- Barabási AL, Albert R, Jeong H. Mean-field theory for scale-free random networks. Physica A: Statistical Mechanics and its Applications. 1999; 272(1):173–187.
- Leskovec J, Kleinberg J, Faloutsos C. Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. ACM; 2005. p. 177–187.
- Girvan M, Newman ME. Community structure in social and biological networks. Proceedings of the national academy of sciences. 2002; 99(12):7821–7826. https://doi.org/10.1073/pnas.122653799
- Blondel V, Guillaume J, Lambiotte R, Lefebvre E. Fast Unfolding of Community Hierarchies in large network. J Stat Mech P. 2008; 1008.
- 60. Jamakovic A. Characterization of complex networks: application to robustness analysis. 2008;.
- 61. Chung FR, Graham FC. Spectral graph theory. 92. American Mathematical Soc.; 1997.