IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Emergent network modularity

This content has been downloaded from IOPscience. Please scroll down to see the full text.

J. Stat. Mech. (2017) 073405

(http://iopscience.iop.org/1742-5468/2017/7/073405)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 168.122.66.195
This content was downloaded on 25/07/2017 at 22:59

Please note that terms and conditions apply.

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Interdisciplinary statistical mechanics

Emergent network modularity

P L Krapivsky[1] **and S Redner**[2]

[1] Department of Physics, Boston University, Boston, MA 02215, United States of America
[2] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, United States of America
E-mail: redner@santafe.edu

Received 7 June 2017
Accepted for publication 16 June 2017
Published 25 July 2017

Online at stacks.iop.org/JSTAT/2017/073405
https://doi.org/10.1088/1742-5468/aa7a3f

Abstract. We introduce a network growth model based on complete redirection: a new node randomly selects an existing target node, but attaches to a random neighbor of this target. For undirected networks, this simple growth rule generates unusual, highly modular networks. Individual network realizations typically contain multiple macrohubs—nodes whose degree scales linearly with the number of nodes N. The size of the network 'nucleus'—the set of nodes of degree greater than one—grows sublinearly with N and thus constitutes a vanishingly small fraction of the network. The network therefore consists almost entirely of leaves (nodes of degree one) as $N \to \infty$.

Keywords: network dynamics, random graphs, networks

J. Stat. Mech. (2017) 073405

© 2017 IOP Publishing Ltd and SISSA Medialab srl
1742-5468/17/073405+22$33.00

Contents

1. Introduction **2**

2. Perfect and near-perfect star graphs **4**

3. Maximal degrees **7**

4. Macrohubs **8**

4.1. Two large hubs: $m, n \gg 1$ . . . 8

4.2. One large and one small hub: m finite and $n \gg 1$ . . . 9

4.2.1. $m = 1$ . . . 9

4.2.2. $m = 2$ . . . 10

4.3. $m = O(1)$ . . . 11

4.4. More than two hubs . . . 11

5. Network nucleus **13**

5.1. Sublinear growth . . . 13

5.2. Anomalously small nucleus . . . 14

6. Multiple linking **15**

7. Discussion **18**

Acknowledgments **18**

Appendix A. Stars with two or more defects **18**

Appendix B. Multiple hubs **20**

References **21**

1. Introduction

Redirection is a fundamental network growth mechanism to determine how a new node attaches to a growing network. For *directed* networks, with a prescribed direction for each link, redirection is implemented as follows (figure 1(a)):

(i) A new node chooses a provisional target node uniformly at random.

(ii) With probability $0 \leqslant 1 - r \leqslant 1$, the new node attaches to this target.

(iii) With probability r, the new node attaches to the ancestor of the target.

By its very construction, an initial tree network always remains a tree.

Without the redirection step (iii), the above growth rules define the random recursive tree (RRT) [1–3], for which the average number of nodes of degree k is given by $N_k = N/2^k$, where N is the total number of nodes. Redirection represents a minimalist

J. Stat. Mech. (2017) 073405

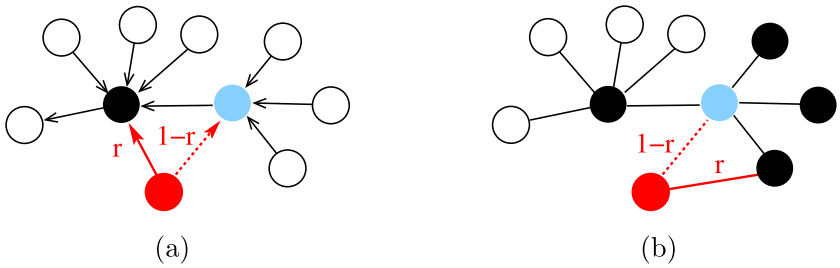


Figure 1. Redirection for (a) directed and (b) undirected networks. (a) The new node attaches by redirection to the unique ancestor (black) of the provisional target (light blue). (b) With the same target, the new node attaches to one of the black nodes.

extension of the RRT; this idea was suggested in [4] and was made more concrete and developed mathematically in [5]. The latter work showed that redirection is equivalent to *shifted linear preferential attachment*, in which the rate of attaching to a pre-existing network node of degree k is proportional to $k + \lambda$, with $\lambda = \frac{1}{r} - 2$. That is, redirection transforms a purely local growth mechanism into a global mechanism. Redirection is also extremely efficient algorithmically; to build a network of N nodes requires a computation time that scales linearly with N, with a prefactor of the order of one.

Because of its useful qualities, redirection has been extended in many ways: connecting to (i) more distant ancestors [6]; (ii) an arbitrary ancestor [7]; (iii) all earlier ancestors [8]; (iv) multiple ancestors [9]; and (v) the ancestor with a probability that depends on the degrees of the provisional target and the ancestor [10, 11]. Each of these scenarios has revealed intriguing features that highlight the richness of the redirection mechanism.

Although this growth mechanism has been applied to directed networks, *undirected* graphs are more pertinent for many applications. In social networks, for example, directionality plays a limited role because friendship is inherently a two-way relationship [12]. This observation motivates us to extend redirection to *undirected* networks. Such an isotropic network again grows according to the rules enumerated above, but now redirection can occur to *any* of the neighbors of the provisional target (figure 1(b)). We define this process as *isotropic redirection* (IR). While the behavior of this IR model for general redirection probability $0 < r < 1$ is interesting in its own right [13, 14], here we focus on the parameter-free case of $r = 1$, where the new node always attaches to a random neighbor of the provisional target.

The consequences of this IR growth rule are surprisingly profound, as highly modular networks emerge (figure 2). Typical network realizations contain a number of well-resolved modules, each with a central macrohub whose degree is a finite fraction of the total number of nodes N. These modules visually resemble a variety of multiplex, or multilayer, networks [15–18]. Typical networks also consist almost entirely of leaves (nodes of degree 1) as $N \to \infty$; that is, the number of leaves satisfies $N_1/N \to 1$ as $N \to \infty$. Nodes with degrees $k > 1$ constitute what we term the 'nucleus' of the network. This nucleus comprises an infinitesimal fraction of the network, as the number of nucleus nodes $\mathcal{N} = \sum_{k\geqslant 2} N_k$ grows as N^{μ}, with $\mu \approx 0.566$.

The number of nodes of degree k grows in a similar manner: $N_k \sim N^{\mu}$ for any $k \geqslant 2$, with an algebraic tail

J. Stat. Mech. (2017) 073405

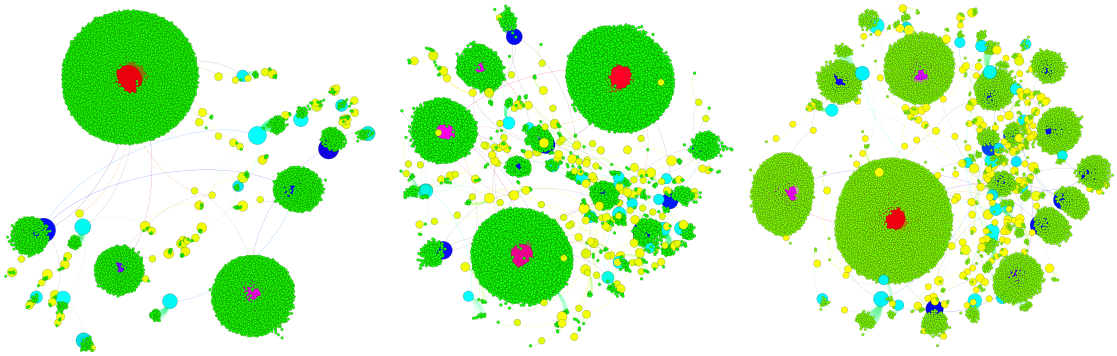


Figure 2. Examples of networks of 10^4 nodes grown by redirection. Green: nodes of degree $k = 1$; yellow, $2 \leqslant k \leqslant 10$; cyan, $11 \leqslant k \leqslant 99$; blue $100 \leqslant k \leqslant 500$; violet → red, $k > 501$. The node radius also indicates its degree.

$$N_k \sim \frac{N^{\mu}}{k^{1+\mu}} \tag{1}$$

when $k \gg 1$. Thus the degree distribution grows sublinearly with network size ($\mu < 1$) so that the degree exponent $1 + \mu$ is generally less than 2. These features strongly contrast with known sparse networks, where the nucleus and the degree distribution grows linearly with network size (see, e.g. [19]) and has the algebraic tail

$$N_k \sim \frac{N}{k^{\nu}}.$$

Here, the degree exponent satisfies $\nu > 2$ and depends on the model (see [19]), while in the IR model (as well as in the models of [11]) the growth exponent μ fixes the degree exponent to be $1 + \mu$, which ensures that it is less than 2. Finally, we emphasize that for directed networks, redirection with $r = 1$ generates a star graph; hence bidirectional links are needed to generate non-trivial networks.

In section 2, we begin by showing that star-like structures are surprisingly common in IR networks. We investigate the probability distribution for maximal degrees in section 3. We then show that multiple macrohubs arise with an anomalously large probability in section 4. In section 5, we discuss basic features of the nucleus of the network. Finally in section 6 we study the intriguing features that arise when a new node attaches to multiple neighbors of the provisional target.

2. Perfect and near-perfect star graphs

Unless otherwise stated, we assume that the initial network is a dimer: •—•. For $N = 3$, there is a single unique graph. For $N = 4$, a star occurs with probability $\frac{2}{3}$ by the new node selecting either of the leaves of the 3-node graph, after which redirection leads to attachment to the central node. Conversely, a linear chain is created with probability $\frac{1}{3}$. All IR network realizations of up to 6 nodes and their weights are shown in figure 3.

While it is impractical to extend this enumeration to large N, we can compute the probabilities to generate the special configurations of a perfect star and near-perfect stars

J. Stat. Mech. (2017) 073405

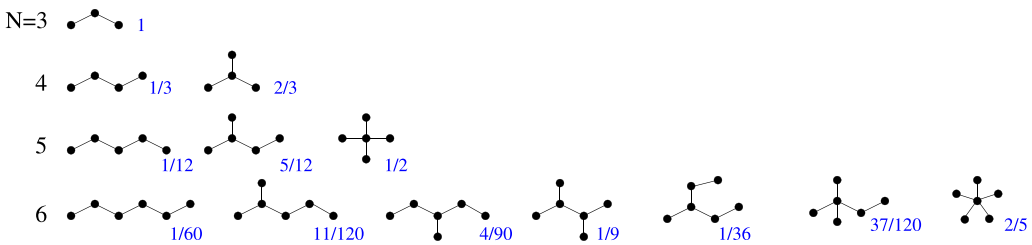


Figure 3. Enumeration of all network configurations up to $N = 6$ nodes.

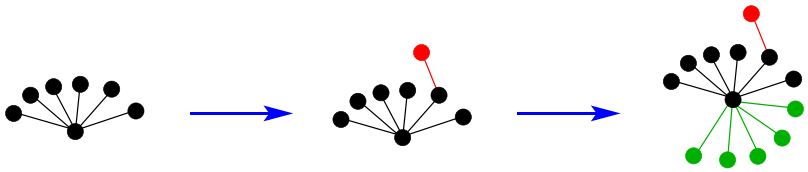


Figure 4. A single-defect star. A perfect star (black) is built to an intermediate stage, then an error occurs (red). All subsequent attachments (green) are to the hub.

of N nodes. Let S_N be the probability to create a perfect star of N nodes. To build this star, a new node has to provisionally select one of the periphery nodes (which occurs with probability $(k-1)/k$ for a star of k nodes), after which redirection shifts the attachment to the center of the star, thereby creating a perfect star of $k+1$ nodes. By this reasoning

$$S_N = \frac{2}{3} \times \frac{3}{4} \times \frac{4}{5} \times \cdots \times \frac{N-2}{N-1} = \frac{2}{N-1}. \tag{2}$$

The slower than exponential decay with N of the star probability provides a first clue that typical network realizations should be star like, as seen in figure 2. To make this surmise stronger, we compute the probability to create a star graph with a single defect. Such a network arises by first building a perfect star of k nodes, then making an 'error' in which the new node attaches to the periphery of the star, and finally building the rest of the star (figure 4).

From equation (2), the probability to build a perfect star of k nodes is $\frac{2}{k-1}$. A defect now occurs with probability $\frac{1}{k}$ because the new node must attach to the center of the star to create this defect. Finally, the probability that all remaining attachments occur to the hub is

$$\frac{k-3/2}{k+1} \times \frac{k-1/2}{k+2} \times \cdots \times \frac{N-5/2}{N} = \frac{\Gamma(k+1)}{\Gamma(k-3/2)} \frac{\Gamma(N-3/2)}{\Gamma(N+1)}. \tag{3}$$

To understand each factor in the product, note that in a network of n nodes with a single defect, there is one hub, one nucleus node (of degree 2), one leaf that attaches to the nucleus node, and $n-3$ leaves that attach to the hub (figure 4). To continue the star, the new node must either select one of the $n-3$ leaves attached to the hub and then redirect to the hub, or select the nucleus node and then redirect to the hub, which occurs with probability $\frac{1}{n} \times \frac{1}{2}$. The probability the new node is redirected to the hub therefore is $(n-3+\frac{1}{2})/n$.

https://doi.org/10.1088/1742-5468/aa7a3f

J. Stat. Mech. (2017) 073405

The probability to create a star of N nodes with a single defect after k nodes has been added, which we define as $S_{N,k}$, is

$$S_{N,k} = \frac{2}{k(k-1)} \frac{\Gamma(k+1)}{\Gamma(k-3/2)} \frac{\Gamma(N-3/2)}{\Gamma(N+1)} = 2\,\frac{\Gamma(k-1)}{\Gamma(k-3/2)} \frac{\Gamma(N-3/2)}{\Gamma(N+1)}. \tag{4}$$

Therefore the probability $S_N^{(1)}$ to build a star of N nodes with a single defect at any stage is given by

$$S_N^{(1)} = \sum_{k=3}^{N-2} S_{N,k} + \frac{2}{(N-1)(N-2)}, \tag{5}$$

where the last term is the probability to create the defect after building a perfect star of $N-1$ nodes. Using (4) we compute the sum in (5) and obtain [20]

$$S_N^{(1)} = \frac{4}{3N} - \frac{2}{(N-1)(N-2)} + \frac{9}{N(N-1)(N-2)} - \frac{4}{3\sqrt{\pi}} \frac{\Gamma(N-3/2)}{\Gamma(N+1)}. \tag{6}$$

Since dominant contribution to the sum comes from the terms with $k \gg 1$, the leading behavior can be extracting by using the asymptotic,

$$S_{N,k} = 2\,\frac{\Gamma(k-1)}{\Gamma(k-3/2)} \frac{\Gamma(N-3/2)}{\Gamma(N+1)} \simeq 2\,k^{1/2}\,N^{-5/2},$$

for $k \gg 1$, and replacing summation in (5) by integration:

$$S_N^{(1)} \simeq \int^N S_{N,k}\,\mathrm{d}k \simeq 2N^{-5/2} \int^N k^{1/2}\,\mathrm{d}k \simeq \frac{4}{3N}. \tag{7}$$

We will use this procedure to show that multiple-defect stars arise with roughly the same frequency as single-defect stars (appendix A).

Comparing (2) and (7) we see that a perfect star is 50% more common than a single-defect star. Naively, one would expect to find a prescribed structure with a probability that is inversely proportional to the total number of networks. The latter grows factorially with N; e.g. the number of labeled trees equals N^{N-2} [21–23]. By a computation similar to (2), the probability to build a linear graph in the IR model equals $2/(N-1)!$ and thus agrees with naive expectations. In contrast, perfect and slightly defective stars occur much more frequently than naively expected.

More importantly, the above reasoning shows that star-like subgraphs will be common in typical network realizations. Consider such a structure, in which the degree of the hub is n. As will be shown in the next section, n ranges from aN to bN, where $0 < a, b < 1$. Thus the probability that there is a star-like subgraph with the degree of the hub in this range is of the order of

$$\int_{aN}^{bN} \frac{\mathrm{d}n}{n} = \ln\frac{b}{a}. \tag{8}$$

That is, with a non-zero (and scale independent) probability, there will be a star-like structure whose degree is of the order of N, as observed in figure 2.

J. Stat. Mech. (2017) 073405

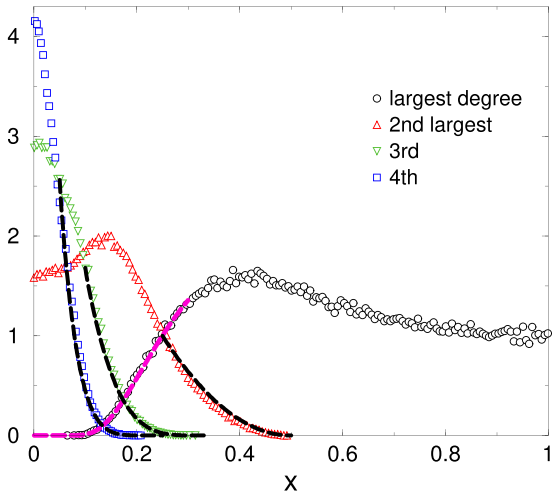


Figure 5. The distributions $\mathcal{M}^{(m)}(x)$ versus normalized degree $x = k/N$ for $m \leqslant 4$ for 10^6 realizations of networks with $N = 10^5$. The initially normalized distributions for the mth largest degree are divided by m, so that these rescaled distributions fit on the same plot. The dashed line that follows the data for $\mathcal{M}^{(1)}$ is the empirical fit $-\frac{1}{2x}\ln(1/(4x))$, while the dashed lines that follow that data for $\mathcal{M}^{(2)}$, $\mathcal{M}^{(3)}$, and $\mathcal{M}^{(4)}$ are the predictions from equation (31), with the amplitudes 16, 600, and 40 000 for the 2nd, 3rd, and 4th largest degree.

3. Maximal degrees

Because macrohubs—nodes whose degree is a finite fraction of N—are at the center of star-like graphs, they should occur at the same frequency as stars. This fact leads us to investigate the statistical properties of the maximal network degree, $k_{\max}$, and also the mth largest degree k_m, in IR networks. The value of $k_{\max}$ in the ensemble of all networks of N nodes is a random quantity that ranges between 2 and $N-1$. The smallest value $k_{\max} = 2$ arises for a linear graph, while the largest value $k_{\max} = N-1$ arises for a star. As figure 5 shows, $k_{\max}$, and indeed k_m for any finite m, scales linearly with N in our IR model. This scaling contrasts with sparse networks, where $k_{\max}$ typically scales sublinearly with N. Because there is a non-zero probability that the maximal degree is close to N, macrohubs will be common in isotropic networks that grow by redirection; a similar feature arises in enhanced redirection [11].

Let $M_N(k)$ denote the probability that the maximal degree in a network of N nodes equals k. This largest degree is distributed over a wide range, but is typically larger than $0.4N$ (figure 5). It is convenient to write this distribution in the scaling form

$$M_N(k) \to \frac{1}{N}\,\mathcal{M}^{(1)}(x), \qquad x = \frac{k}{N} \tag{9}$$

as $k, N \to \infty$, with finite rescaled degree x. The prefactor N^{-1} imposes the normalization $\int_0^1 \mathrm{d}x\,\mathcal{M}^{(1)}(x) = 1$. Because the distribution $\mathcal{M}^{(1)}(x)$ does not sharpen as N increases, moments of this distribution do not self-average[3]; moreover, the distribution is singular as $x \to 0$. Similar singularities arise in a variety of non-self-averaging

[3] For a review of non-self-averaging phenomena see, e.g. Derrida [24].

J. Stat. Mech. (2017) 073405

processes, such as random maps, random walks, spin glasses, and fragmentation processes [24–29]. In these systems, it was found that $\mathcal{M}^{(1)}(x)$ has an essential singularity of the form $\exp[-x^{-1}\ln(1/x)]$ as $x \to 0$. The same singularity apparently occurs here. Indeed, matching the exact result $M_N(2) = 2/(N-1)!$ for the minimal possible degree with the scaled form $\frac{1}{N}\,\mathcal{M}^{(1)}\left(\frac{2}{N}\right)$ gives

$$\ln \mathcal{M}^{(1)} \sim -\frac{2}{x}\left(\ln\frac{2}{x} - 1\right), \tag{10}$$

which qualitatively captures the small-x behavior of $\mathcal{M}^{(1)}(x)$. However, we must be cautious in making this connection because the scaled form is formally applicable when the rescaled degree is finite, while we used $x = \frac{2}{N} \to 0$ in connecting the data to the scaling form. In fact, we find a good visual fit using $\ln \mathcal{M}^{(1)}(x) \sim -\frac{1}{2x}\ln\left(\frac{1}{4x}\right)$.

More generally, the distributions $\mathcal{M}^{(m)}(x)$ for the mth largest degree have support on $\left[0, \frac{1}{m}\right]$ and exhibit power-law singularities as $x \to \frac{1}{m}$ from below (figure 5). We will derive this singular behavior in the next section from the limiting behavior of the probability to find macrohubs of specific topologies.

4. Macrohubs

As illustrated in figure 2, typical IR network realizations contain multiple macrohubs. To appreciate why such configurations are common, let us examine the likelihood that there are exactly two connected hubs, while all remaining nodes are leaves (see also [11]). Suppose that one hub is connected to m leaves and the other to n leaves, leading to what we define as the (m, n) graph (figure 6). The hub degrees are $m+1$ and $n+1$, respectively, and the total number of nodes is $m+n+2$. We term these nodes as hubs even if one of their degrees happens to be small.

Let $H_{m,n}$ be the probability to build an (m, n) graph. Because, this graph arises from $(m-1, n)$ and $(m, n-1)$ graphs, we can express $H_{m,n}$ through $H_{m-1,n}$ and $H_{m,n-1}$:

$$H_{m,n} = \frac{m - 1 + (n+1)^{-1}}{m+n+1}\,H_{m-1,n} + \frac{n - 1 + (m+1)^{-1}}{m+n+1}\,H_{m,n-1} \tag{11}$$

For example, to build an (m, n) graph from an $(m-1, n)$, the new node can either select one of the $m-1$ leaves on the left of an $(m-1, n)$ graph and redirect to the hub on the left. The probability for this event is $(m-1)/(m+n+1)$. Additionally, the new node can select the right hub and redirect to the left hub. This event occurs with probability $1/\big[(n+1)(m+n+1)\big]$.

We now solve (11) for several relevant cases that help reveal the star-like nature of typical IR network realizations.

4.1. Two large hubs: $m, n \gg 1$

When the hub degrees are both large, we treat m and n as continuous variables and expand $H_{m-1,n}$ and $H_{m,n-1}$ in the Taylor series

J. Stat. Mech. (2017) 073405

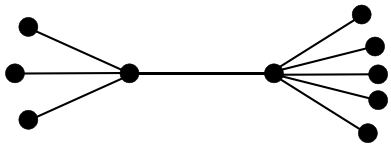


Figure 6. The $(m, n) = (3, 5)$ graph.

$$H_{m-1,n} = H - \frac{\partial H}{\partial m}, \quad H_{m,n-1} = H - \frac{\partial H}{\partial n},$$

where $H = H_{m,n}$. Using this in (11) gives

$$m\,\frac{\partial H}{\partial m} + n\,\frac{\partial H}{\partial n} = -3H.$$

The solution that satisfies the necessary symmetry requirement $H_{m,n} = H_{n,m}$ is

$$H_{m,n} = \frac{C_2}{(mn)^{3/2}}, \tag{12}$$

where the amplitude C_2 is not computable within the continuum approximation.

From (12), the probability for the $\left(\frac{N}{2}, \frac{N}{2}\right)$ graph scales as N^{-3}. This also gives the tail behavior of distribution of the second-largest degree, because the probability for the $\left(\frac{N}{2}, \frac{N}{2}\right)$ graph coincides with the probability that the second-largest degree equals $N/2$. Similar to the distribution of the largest degree (equation (9)), we anticipate that the second-largest degree distribution has the scaling behavior

$$\frac{1}{N}\,\mathcal{M}^{(2)}(x), \qquad \text{with} \qquad x = \frac{k}{N} \leqslant \frac{1}{2}. \tag{13}$$

This form is compatible with the above N^{-3} probability for the $\left(\frac{N}{2}, \frac{N}{2}\right)$ graph if

$$\mathcal{M}^{(2)}(x) \sim \left(\frac{1}{2} - x\right)^2 \quad \text{when} \quad x \to \frac{1}{2}. \tag{14}$$

Visually, this asymptotic behavior quantitatively agrees with simulation results when an appropriate amplitude is chosen (figure 5).

4.2. One large and one small hub: *m* finite and $n \gg 1$

4.2.1. $m = 1$ Suppose that the degree of the smaller hub $m = 1$. This $(1, n)$ graph is also just the single-defect star shown in figure 4. The number of nodes in such graph is $N = n + 3$ and the largest degree is $k_{\max} = n + 1 = N - 2$. To compute $H_{1,n}$, we must write the analog of the recursion (11) that applies for $m = 1$. This recursion is

$$H_{1,n} = \frac{2}{(n+1)(n+2)} + \frac{n - \frac{1}{2}}{n+2}\,H_{1,n-1}. \tag{15}$$

The first term on the right-hand side is the contribution that arises by creating the $(1, n)$ graph from a perfect star with $n + 2$ nodes. To solve (15), note that the homogeneous

J. Stat. Mech. (2017) 073405

version of (15) admits the solution $\Gamma(n+\frac{1}{2})/\Gamma(n+3)$. We use this solution as an integrating factor

$$H_{1,n} = \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+3)}\, A_n, \tag{16}$$

and substitute this form into (15) to give

$$A_n = A_{n-1} + 2\,\frac{\Gamma(n+1)}{\Gamma(n+\frac{1}{2})}.$$

Because $H_{1,2} = \frac{5}{12}$ (see figure 3), the initial condition for this recursion is $A_2 = 40/(3\sqrt{\pi})$. Thus [20]

$$A_n = 2\sum_{j=3}^{n} \frac{\Gamma(j+1)}{\Gamma(j+\frac{1}{2})} + \frac{40}{3\sqrt{\pi}} = \frac{2}{3}\left[\frac{2\Gamma(n+3) - 3\Gamma(n+2)}{\Gamma(n+\frac{3}{2})} + \frac{4}{\sqrt{\pi}}\right], \tag{17}$$

which leads to

$$H_{1,n} = \frac{2\Gamma(n+\frac{1}{2})}{3\Gamma(n+3)}\left[\frac{2\Gamma(n+3) - 3\Gamma(n+2)}{\Gamma(n+\frac{3}{2})} + \frac{4}{\sqrt{\pi}}\right] \tag{18}$$

for $n \geqslant 2$. The asymptotic behavior of (18) is

$$H_{1,n} \simeq \frac{U_1}{n}, \qquad U_1 = \frac{4}{3}, \tag{19}$$

which coincides with the asymptotic probability for a single-defect star given in equation (7). As another useful consequence of (18), notice that a node of degree $N-2$ appears only in the $(1, N-3)$ graph. Therefore $H_{1,N-3} = M_N(N-2)$, so that (18) also gives the probability that the largest degree in a graph of N nodes equals $N-2$.

4.2.2. $m = 2$ In analogy with (15), the recurrence for $H_{2,n}$ is

$$H_{2,n} = \frac{n+2}{(n+1)(n+3)}\, H_{1,n} + \frac{n-\frac{2}{3}}{n+3}\, H_{2,n-1}. \tag{20}$$

Using the homogeneous solution, one can again define the integrating factor $H_{2,n} = \left[\Gamma(n+\frac{1}{3})/\Gamma(n+4)\right] B_n$, which reduces (20) to

$$B_n = B_{n-1} + \frac{n+2}{n+1}\,\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+\frac{1}{3})}\, A_n,$$

with A_n given by (17). However, instead of deriving the exact solution to this equation, it is easier to extract the asymptotics by taking the continuum limit of equation (20) to give

$$\left(n\,\frac{\mathrm{d}}{\mathrm{d}n} + 3 + \frac{2}{3}\right) H_{2,n} \simeq \frac{4}{3n},$$

J. Stat. Mech. (2017) 073405

from which

$$H_{2,n} \simeq \frac{U_2}{n}, \qquad U_2 = \frac{1}{2}. \tag{21}$$

Note that this result also gives the probability for the 2-defect star shown in figure A1(b) in appendix A.

4.3. $m = O(1)$

More generally, for $m = O(1)$ and $n \gg 1$, we apply the continuum approach to the large variable in (11) and obtain

$$\left(n \frac{\partial}{\partial n} + m + 2 - \frac{1}{m+1} \right) H_{m,n} = H_{m-1,n} \tag{22}$$

Based on (19) and (21) we again expect that

$$H_{m,n} = \frac{U_m}{n} \tag{23}$$

for $m = O(1)$ and $n \gg 1$. Substituting (23) into (22), the amplitudes satisfy the recursion

$$U_m = \left(m + 1 - \frac{1}{m+1} \right)^{-1} U_{m-1},$$

from which

$$U_m = 2 \prod_{j=1}^{m} \left(j + 1 - \frac{1}{j+1} \right)^{-1} = 4 \frac{m+1}{(m+2)!}. \tag{24}$$

As a postscript, note that $D_2 \equiv \sum_{m \geqslant 1} U_m = 2$, so that the probability for all (m, n) graphs with $m = O(1)$ and $n \gg 1$ is dominated by the first two terms, for which $U_1 + U_2 = \frac{11}{6}$.

4.4. More than two hubs

To understand the general behavior, consider first the case of three macrohubs. Let ℓ, m, n denote the number of leaves connected to hubs of degrees $\ell + 1$, $m + 2$, and $n + 1$. Note that the 'central' hub with n leaves is special because it is linked to both other hubs (figure 7). The total number of nodes is $\ell + m + n + 3$. Denote by $H_{\ell,m,n}$ the probability to build such an (ℓ, m, n) graph. This graph can arise from $(\ell - 1, m, n)$, $(\ell, m - 1, n)$ and $(\ell, m, n - 1)$ graphs, which occur with probabilities $H_{\ell-1,m,n}$, $H_{\ell,m-1,n}$ and $H_{\ell,m,n-1}$.

These probabilities satisfy the recursion

$$H_{\ell,m,n} = \frac{\ell - 1 + (m+2)^{-1}}{\ell + m + n + 2} H_{\ell-1,m,n} + \frac{m - 1 + (\ell+1)^{-1} + (n+1)^{-1}}{\ell + m + n + 2} H_{\ell,m-1,n} + \frac{n - 1 + (m+2)^{-1}}{\ell + m + n + 2} H_{\ell,m,n-1}. \tag{25}$$

J. Stat. Mech. (2017) 073405

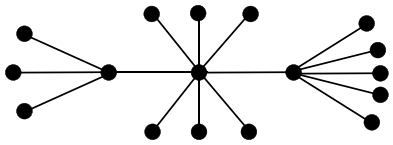


Figure 7. The $(\ell, m, n) = (3, 6, 5)$ graph.

Using continuum approach and assuming that all three hubs have large degrees, i.e. $\ell, m, n \gg 1$, we recast (25) into

$$\ell \frac{\partial H}{\partial \ell} + m \frac{\partial H}{\partial m} + n \frac{\partial H}{\partial n} + 5H = 0,$$

from which

$$H_{\ell,m,n} = \frac{C_3}{(\ell m n)^{5/3}}. \tag{26}$$

Suppose that all three hubs are macroscopic, i.e. their degrees are linear in the number of nodes: $(\ell, m, n) = N(a, b, c)$, with $a, b, c > 0$ and $a + b + c = 1$. The probability $H_N(a, b, c)$ for such a three-hub network is

$$H_N(a, b, c) = \frac{C_3 (abc)^{-5/3}}{N^5} \tag{27}$$

Thus the probability for the $\left(\frac{N}{3}, \frac{N}{3}, \frac{N}{3}\right)$ graph scales as N^{-5}, which coincides with the probability that the third-largest degree has the maximal possible size $\frac{N}{3}$. This third-largest degree also has the scaling behavior $N^{-1}\mathcal{M}^{(3)}(x)$, which is compatible with the N^{-5} extremal behavior for the probability of three macrohubs when

$$\mathcal{M}^{(3)}(x) \sim \left(\frac{1}{3} - x\right)^4 \qquad x \to \frac{1}{3}. \tag{28}$$

Generally, the probability for a graph with h macrohubs depends on the nature of the links between these hubs when $h \geqslant 3$. Nevertheless, if all the macrohub degrees m_j are large, the hub probability $H_{\mathbf{m}}$, with $\mathbf{m} = (m_1, \ldots, m_h)$, now satisfies

$$\sum_{j=1}^{h} m_j \frac{\partial H}{\partial m_j} + (2h - 1)H = 0$$

from which

$$H_{\mathbf{m}} = C_h \prod_{j=1}^{h} m_j^{1/h-2} \tag{29}$$

When all hubs are macroscopic, that is, $m_j = N a_j$, with $0 < a_j < 1$ and $\sum_{j=1}^{h} a_j = 1$, the network is realized with probability

$$H_N(\mathbf{a}) = \frac{C_h}{N^{2h-1}} \prod_{j=1}^{h} a_j^{1/h-2}, \tag{30}$$

J. Stat. Mech. (2017) 073405

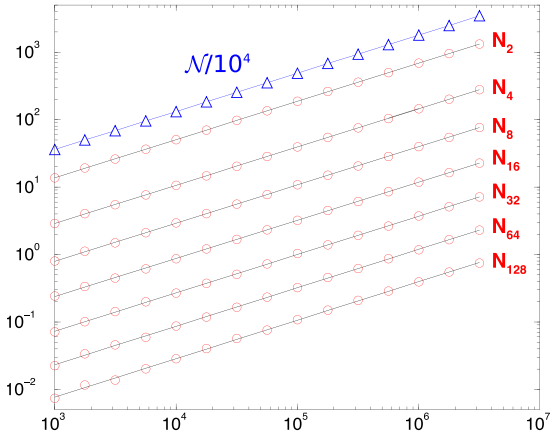


Figure 8. Dependence of $\mathcal{N}$ and N_k versus N for various k values.

where $\mathbf{a} = (a_1, \ldots, a_h)$. From (30), the scaled distribution of the hth largest-degree macrohub has the extremal behavior

$$\mathcal{M}^{(h)}(x) \sim \left(\frac{1}{h} - x\right)^{2h-2} \tag{31}$$

close to the maximal possible value $x \to \frac{1}{h}$. The simulation data shown in figure 5 is consistent with these singular behaviors for $h = 2$, 3, and 4.

5. Network nucleus

5.1. Sublinear growth

One of most enigmatic features of IR networks is that they consist almost entirely of leaves, namely, nodes of degree one (figure 2). Nodes of degree greater than one constitute what we term the *nucleus* of the network. Surprisingly, both the average number of nucleus nodes, $\mathcal{N} = \sum_{k\geqslant 2} N_k$, and indeed the average number of nodes N_k of any fixed degree $k \geqslant 2$, grow sublinearly with N (figure 8):

$$\mathcal{N} \sim N^{\mu}, \qquad N_k \sim N^{\mu}, \tag{32}$$

with exponent $\mu \approx 0.566$. The data are quite linear on the double logarithmic scale of the figure and a linear fit gives a correlation coefficient of 0.999 9956. In removing successive data points and performing the same regression analysis, there is no systematic change to the slope, which ranges between 0.5652 and 0.5673. Our quoted exponent value of $\mu = 0.566$ therefore seems accurate to within 0.001. Because of this sublinear growth, the nucleus represents a vanishing fraction of the entire network as $N \to \infty$, as is visually evident in figure 2. This behavior stands in stark contrast to that of sparse networks, where the nucleus represents a finite fraction of the whole.

Starting with the sublinear scaling (32), it is possible to generally show [11] that there is a power-law decay for the normalized degree distribution, $c_k = N_k/\mathcal{N}$

J. Stat. Mech. (2017) 073405

$$c_k \sim k^{-(1+\mu)} \qquad k \gg 1, \tag{33}$$

with the degree distribution exponent $1+\mu$ less than 2. Such an exponent value can be shown to be mathematically inconsistent [11] unless the size of the nucleus grows sublinearly with N.

We now demonstrate that the nucleus of IR networks must grow sublinearly with N. Let $N_{k,\ell}$ be the number of nodes of degree k that are connected to ℓ leaves, and let $c_{k,\ell} = N_{k,\ell}/\mathcal{N}$ be the density of such nodes. We make the mild assumptions that c_k and $c_{k,\ell}$ are both independent of N for $N \to \infty$. Then the number of nucleus nodes grows according to

$$\frac{\mathrm{d}\mathcal{N}}{\mathrm{d}N} = \frac{\mathcal{N}}{N} \sum_{k\geqslant 2} \sum_{\ell<k} c_{k,\ell}\, \frac{\ell}{k}. \tag{34}$$

That is, the size of the nucleus increases by 1 only when a new node initially selects a nucleus node and then redirects to a leaf. A nucleus node of degree k that is attached to ℓ leaves is selected with probability $N_{k,\ell}/N = \mathcal{N}\, c_{k,\ell}/N$, and redirection to a leaf occurs with probability $\frac{\ell}{k}$. Nucleus nodes that are not attached to leaves are characterized by $\ell = 0$; attaching to these nodes therefore does not affect the nucleus size.

If the densities $c_{k,\ell}$ are independent of N as $N \to \infty$, then (34) implies that

$$\mu = \sum_{k\geqslant 2} \sum_{\ell<k} c_{k,\ell}\, \frac{\ell}{k}. \tag{35}$$

We now obtain the strict upper bound $\mu < 1$ by replacing ℓ by its largest possible value, which is $k-1$, in the sum in (35). We also exploit the two obvious sum rules, $\sum_{\ell<k} c_{k,\ell} = c_k$ and $\sum_{k\geqslant 2} c_k = 1$ to give

$$\sum_{k\geqslant 2} \sum_{\ell<k} c_{k,\ell}\, \frac{\ell}{k} \leqslant \sum_{k\geqslant 2} c_k\, \frac{k-1}{k} = 1 - \sum_{k\geqslant 2} k^{-1} c_k < 1.$$

We conclude that $\mu < 1$, which gives the fundamental result that the nucleus grows sublinearly with N.

5.2. Anomalously small nucleus

We previously showed there is an anomalously large probability, of order $1/N$, to generate star-like graphs that necessarily have a small nucleus. Therefore we anticipate that the probability to generate a graph with a nucleus whose size is a finite number will also be proportional to $1/N$. We therefore focus on the probability that the nucleus has a finite size h:

$$P_h(N) \equiv \mathrm{Prob}\big[\big|\mathcal{N}\big| = h\big]. \tag{36}$$

Equivalently, this is the same as the probability that there are h macrohubs in the system. We already know the probability for the nucleus to consist of a single macrohub,

$$P_1(N) \equiv S_N = \frac{2}{N-1},$$

because this is the same as the probability to create a star graph.

J. Stat. Mech. (2017) 073405

Consider the probability that there two hubs. Let us first suppose that the degrees of both hubs are large; without loss of generality, we set $m \leqslant n$. Summing over all partitions of the two hub degrees, a graph with N nodes has exactly 2 hubs with probability

$$P_2(N) = \sum_{m=1}^{\lfloor N/2 \rfloor - 1} H_{m,N-2-m}. \tag{37}$$

When both hubs are macroscopic, we previously derived in equation (12) that $H_{m,n} \simeq N^{-3}$. Moreover, the number of such contributions to the above sum is of the order of N. Hence the overall contribution to this sum from macroscopic hubs scales as N^{-2}.

Thus the dominant contribution to the sum comes from terms with small m. We therefore use the result from equation (23) that $H_{m,n} = U_m n^{-1}$ for $m \sim O(1)$ and n large and substitute into (37) to obtain

$$P_2(N) \simeq \frac{\mathrm{D}_2}{N} = \frac{2}{N} \qquad N \to \infty. \tag{38}$$

with D_2 defined immediately after equation (24). Therefore the probabilities to generate either one hub—a star network—or two hubs are asymptotically of the same order.

Based on these results, we anticipate that

$$P_h(N) \simeq \frac{\mathrm{D}_h}{N} \tag{39}$$

for arbitrary h. In analogy with the discussion of two hubs, to justify (39) it is necessary to determine $H_{\mathbf{m}}$ near the 'corner' values of $\mathbf{m}$, where all hub degrees, apart from one, are small; the contribution from the cases where all hubs have macroscopic degrees is negligible. As a first step, we analyze two illustrative cases with three hubs in which one of them is macroscopic in appendix B. Namely, we show that the probabilities for the graphs $H_{1,m,1}$ and $H_{\ell,0,1}$ (see equations (B.3) and (B.6)) are indeed proportional to N^{-1}.

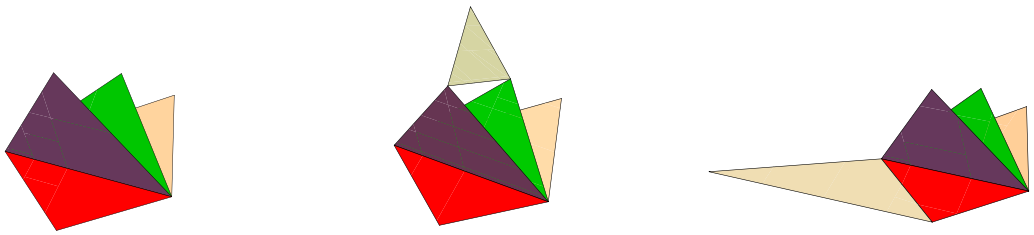
6. Multiple linking

We now extend IR networks in which a new node makes more than one link to the network. For sparse networks, this modification affects only the amplitude of the degree distribution. For example, for linear preferential attachment in which the new node makes m independent links to the network, the number of nodes of degree k asymptotically scales as (see e.g. [5, 19])

$$N_k \simeq \frac{2m(m+1)N}{k^3}.$$

That is, the exponent of the degree distribution does not depend on m, the out degree of the new node. In isotropic redirection, however, the degree m of the newly introduced node materially affects the degree distribution.

J. Stat. Mech. (2017) 073405



(a) (b) (c)

Figure 9. (a) Illustration of a book with $M = 4$ pages. (b) A defect in which the new node links to two lowest-degree nodes after a book of $M = 6$ pages is created. (c) A defect in which the new node links to a node of the lowest degree and a node of the highest degree.

There are two natural ways to implement multiple linking: (a) the new node selects a provisional target at random and attaches to m randomly selected but distinct neighbors of this single target, or (b) the new node selects m provisional random targets and attaches to a random neighbor of each of these targets. Let us first consider rule (a) with $m = 2$; here, it is convenient to take the initial condition as a triangle. We may visualize the growth process as an effective triangle being created each time a new node and two new links are introduced (figure 9). However, we emphasize that we are generating a random graph, and not a set of random triangular surfaces. (Random triangular surfaces and, more generally, random simplicial complexes are studied, e.g. in [31–36]).

More generally, we study networks where each new node makes m links to the network. Here, it is convenient to take the initial condition as a complete graph of $m + 1$ nodes. We define the nucleus of the network as the set of nodes with degree $k > m$. Once again the average number of nucleus nodes $\mathcal{N}$, as well as the average number of nodes of any fixed degree N_k, with $k > m$, both scale in the same way and sublinearly with N: $\mathcal{N} \sim N^{\mu}$ and $N_k \sim N^{\mu}$. The salient feature is that exponent μ is non-universal with respect to m (and also to the two linking rules, (a) and (b)). For the first few cases, the exponent values are:

$$\mu = \begin{cases} 0.74 & 1 \text{ target, } m = 2 \text{ links to its neighbors} \\ 0.83 & 1 \text{ target, } m = 3 \text{ links to its neighbors} \\ 0.88 & 1 \text{ target, } m = 4 \text{ links to its neighbors} \\ 0.83 & \text{link to neighbors of } m = 2 \text{ targets} \\ 0.93 & \text{link to neighbors of } m = 3 \text{ targets} \\ 0.97 & \text{link to neighbors of } m = 4 \text{ targets} \end{cases} \tag{40}$$

As m increases the scaling of $\mathcal{N}$ and N_k gradually approaches linearity in N. We note that our values for the degree distribution exponent $\nu = 1 + \mu$ are close to those reported in [37] for various Wikipedia pages where multiple linking is significant.

In analogy with star graphs when the new node make a single link, we also examine the corresponding extremal graphs when the new node makes $m > 1$ links. For $m = 2$ and rule (a), the analog of a star graph is a *book* (figure 9(a)). A book with N nodes has $N - 2$ triangular pages that all share a common link that acts as the binding

J. Stat. Mech. (2017) 073405

between the two highest degree nodes. (In [30] this graph was called an open book.) The probability B_N to build a book of N nodes is obtained by iterating the recursion $B_{N+1} = \frac{N-2}{N} B_N$. Starting with $B_4 = 1$, we obtain

$$B_N = \frac{6}{(N-1)(N-2)} \qquad N \geqslant 4. \tag{41}$$

Consider now books with one defect, as illustrated in figure 9(b). To compute the probability to create a single-defect book, one first generates a book of M nodes, and then make an error by selecting one of the highest degree nodes as the target and thereby link to two nodes of the lowest degree (figure 9(b)). All subsequent growth steps continue to build the book without any additional errors. This configuration occurs with probability

$$B(M,N) = B_M \, \frac{2}{M} \, \frac{M-3}{M-1} \prod_{k=M+1}^{N-1} \frac{k-\frac{13}{3}}{k}. \tag{42}$$

Here B_M accounts for generating a book with M nodes, the factor $\frac{2}{M}$ accounts for then 'erroneously' choosing one of the two highest degree nodes, and the factor $\frac{M-3}{M-1} = \binom{M-2}{2} / \binom{M-1}{2}$ accounts for the linking to the lowest degree nodes. The probability for the remaining attachments to occur without any errors is $(k-\frac{13}{3})/k$ when the total number of nodes equals k. Writing the product in terms of gamma functions gives

$$B(M,N) = 12 \, \frac{(M-3)\,\Gamma(M-2)}{(M-1)\,\Gamma(M-\frac{10}{3})} \, \frac{\Gamma(N-\frac{13}{3})}{\Gamma(N)} \simeq 12\, M^{4/3}\, N^{-13/3} \qquad M \to \infty. \tag{43}$$

Thus the probability to create a network with a single defect of the type illustrated in figure 9(b) is given by

$$B'_N = \sum_{4 \leqslant M \leqslant N-1} B(M,N) \simeq \frac{36}{7} \frac{1}{N^2} \qquad N \to \infty. \tag{44}$$

Analogously, we can compute the probability to create a book with a single error of the type shown in figure 9(c). The probability to create this defect after the network contains M nodes is

$$\begin{aligned} A(M,N) &= B_M \, \frac{2}{M} \, \frac{2}{M-1} \prod_{k=M+1}^{N-1} \frac{k-\frac{11}{3}}{k} = 24 \, \frac{\Gamma(M-2)}{(M-1)\,\Gamma(M-\frac{8}{3})} \, \frac{\Gamma(N-\frac{11}{3})}{\Gamma(N)}, \\ &\simeq 24\, M^{-1/3}\, N^{-11/3} \qquad M \to \infty. \end{aligned} \tag{45}$$

The total probability $A'_N = \sum_{4 \leqslant M \leqslant N-1} A(M,N)$ to have one defect of this type therefore scales as

$$A'_N \simeq \frac{36}{N^3}. \tag{46}$$

Thus defects of the type in figure 9(b) are more common than those shown in 9(c).

https://doi.org/10.1088/1742-5468/aa7a3f

J. Stat. Mech. (2017) 073405

7. Discussion

We introduced a parameter-free network growth mechanism—isotropic redirection (IR)—that represents a minimalist extension of the classic random recursive tree (RRT). In the RRT, new nodes connect to an existing network one by one. Each new node selects and connects to a target node in the existing network that is chosen uniformly at random. In our IR model, each new node again selects a random target node but then connects to one of its neighbors.

In spite of the homogeneity of this simple growth rule, highly modular networks emerge that contain multiple macrohubs—nodes whose degrees are macroscopic (figure 2). Visually, these networks share many features with multiplex networks; the latter are comprised of well-resolved individual networks that are weakly interconnected. It is remarkable that the modular configurations characteristic of multiplex networks arise essentially for free in our IR model.

A striking feature of network realizations in our IR model is that star-like structures are quite common. Naively, one might have anticipated that the probability for the occurrence of these extremal configurations would be exponentially small in N. By probabilistic reasoning, we showed that the likelihood of stars and near-perfect stars are both proportional to N^{-1}. This anomalous probability results from the inherent amplification of the redirection process as the degree of a central node starts to 'run away' from the typical network degree. Therefore star-like structures whose central degree is of the order of N occur with a non-zero probability.

An outstanding theoretical challenge is to determine the exponent μ that characterizes the sublinear scaling of the average size $\mathcal{N}$ of the 'nucleus' of the network. Numerically, we found $\mathcal{N} \sim N^{\mu}$, with $\mu \approx 0.566$. Thus the nucleus represents a vanishingly small fraction of the entire network. This behavior again starkly contrasts with sparse networks, where the size of the nucleus scales linearly with N. If the new node makes $m > 1$ connections to the network, then many of the anomalous features observed for the case $m = 1$ still arise, but the scaling of nucleus size on N appears to approach linearity as m increases.

Acknowledgments

We thank D ben-Avraham for discussions and collaboration on the IR model for the general case of $r < 1$. We are grateful to G S Redner for his assistance in translating our FORTRAN codes into C++ to take advantage of dynamic memory allocation, a helpful feature for growing undirected networks. We also acknowledge support from grants DMR-1608211 and DMR-1623243 from the National Science Foundation, and by the John Templeton Foundation (SR).

Appendix A. Stars with two or more defects

We extend the approach given in section 2 to compute the probability for a star with two or more defects. We will show that multiple-defect stars are more common than single-defect stars, thus providing evidence that typical network configurations

J. Stat. Mech. (2017) 073405

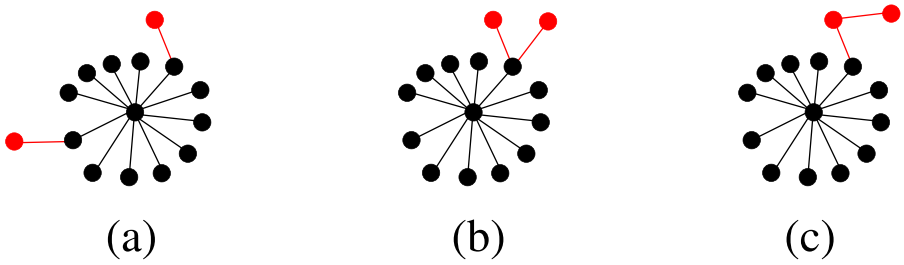


Figure A1. The three types of stars with 2 defects.

contain many star-like subgraphs. To create a star with two defects, the following must occur:

(i) First build a perfect star of k nodes.

(ii) Make an error when the next node is introduced.

(iii) Then continue attaching to the hub until a single-defect star of ℓ nodes is made.

(iv) Make a second error when the next node is introduced.

(v) Then continue attaching to the hub until a two-defect star of N nodes is made.

There are three distinct types of 2-defect stars (figure A1), and we now calculate the probability to create each of them. For each type, this probability generically has the asymptotic form

$$S_{N,k,\ell} = \frac{2}{(k-1)}\, d_1(k) \prod_{k+1}^{\ell} \left(1 - \frac{a_1}{n}\right)\, d_2(\ell) \prod_{\ell+1}^{N} \left(1 - \frac{a_2}{n}\right). \tag{A.1}$$

The first factor is the probability to create a perfect star of k nodes. The second factor, $d_1(k) = \frac{1}{k}$, is the probability to create the first defect. The next factor is the probability to add new nodes to the network without creating another defect; this probability was written in equation (3) with $a_1 = \frac{5}{2}$. The factor $d_2(\ell)$ is the probability to create the second defect when the network contains ℓ nodes. The last product gives the probability to build the network to its final state without any additional defects. Both $d_2(\ell)$ and a_2 depend on the topology of the 2-defect star that is created.

Before specifying $d_2(\ell)$ and a_2, we first determine the asymptotic behavior of the products in (A.1). While we can write them in terms of gamma functions, the following shortcut suffices for the asymptotic behavior. Generically, we write the products in equation (A.1) as

$$\prod_{n=k}^{\ell} \left(1 - \frac{a_1}{n}\right) = \exp\left[\sum_{n=k}^{\ell} \ln\left(1 - \frac{a_1}{n}\right)\right] \simeq \exp\left[-\int_k^{\ell} \frac{a_1}{n}\, \mathrm{d}n\right] \simeq \left(\frac{k}{\ell}\right)^{a_1}. \tag{A.2}$$

We now determine $d_2(\ell)$ and a_2 for the distinct 2-defect stars in figures A1(a), (b), and (c). For case (a), there is one hub, one core node, one leaf attached to the nucleus node, and $\ell - 3$ leaves attached to the hub. By enumerating all relevant states, the probability to create a second defect is

J. Stat. Mech. (2017) 073405

$$d_2(\ell) = \frac{1}{\ell}\,\frac{\ell-3}{\ell-2} \simeq \frac{1}{\ell}.$$

Once a second defect is created, a network of n nodes consists of a hub, two nucleus nodes, two leaves attached to nucleus nodes, and $n-5$ leaves attached to the hub. If the new node selects one of these $n-5$ leaves, then attachment to the hub occurs. If the new node selects one of the two nucleus nodes, then with probability 1/2, redirection to the hub occurs. Thus the probability that a new node attaches to the hub is $\big[n-5+2\times(1/2)\big]/n = 1-\frac{4}{n}$, so that

$$a_2 = 4.$$

By enumerations in the same spirit, the results for cases (b) and (c) are

$$d_2(\ell) = \frac{1}{\ell}\Big(1+\frac{1}{\ell-2}\Big) \simeq \frac{1}{\ell}, \qquad a_2 = \frac{11}{3},$$
$$d_2(\ell) = \frac{1}{2\ell}, a_2 = \frac{7}{2}.$$

Substituting these in equation (A.1), integrating over the possible values of k and ℓ, the probability to create an N node 2-defect star of type (a) is given by

$$\begin{aligned} S_N^{(2)} \equiv \sum S_{N,k,\ell} &\simeq \int_1^N \mathrm{d}k \int_k^N \mathrm{d}\ell\, d_2(\ell)\, S_{N,k,\ell}, \\ &\simeq \frac{2}{k^2}\int_1^N \mathrm{d}k \int_k^N \left(\frac{k}{\ell}\right)^{a_1} \frac{1}{\ell}\left(\frac{\ell}{N}\right)^{a_2}, \\ &\simeq \frac{2}{N}\,\frac{1}{(a_1-1)(a_2-1)} = \frac{4}{9N}. \end{aligned} \tag{A.3}$$

Note that this type of defective star can also be viewed as the graph of type $(1,m,1)$, whose probability is determined independently in the next section. The result (A.3) thus coincides with (B.3), the probability to create a $(1,m,1)$ graph. By similar calculations, the probability to create 2-defect stars of types (b) and (c) are $\frac{1}{2N}$ and $\frac{4}{15N}$, respectively. The former reproduces (21), while the latter reproduces (B.6), which will also derived in the next section.

Appendix B. Multiple hubs

To help understand the behavior of the probability to generate a small number of macrohubs, we compute the probability for specific networks with three hubs when two hub degrees are small. The enumerative procedure is straightforward, albeit a bit tedious. As a first example, consider $H_{1,m,1}$. In this case the analog of the recursion (25), for the case where the first and third arguments are small, is

$$H_{1,m,1} = \frac{m+1}{(m+2)(m+4)}\,H_{1,m+1} + \frac{m}{m+4}\,H_{1,m-1,1}, \tag{B.1}$$

J. Stat. Mech. (2017) 073405

with $H_{1,m+1}$ determined by (18). While this recurrence is soluble, it suffices to use the continuum approach to determine the asymptotic behavior. Using (19), we recast (B.1) into

$$\left(m\,\frac{\mathrm{d}}{\mathrm{d}m}+4\right)H_{1,m,1}\simeq\frac{4}{3m}, \tag{B.2}$$

from which

$$H_{1,m,1}\simeq\frac{4}{9m}. \tag{B.3}$$

This result coincides with (A.3), as it must.

Another illustrative example is when the degree of the central hub is the smallest; the simplest such example is the $(\ell,0,1)$ graph. Here the central hub is not linked to any leaf, but has degree 2. For this limiting case, the recursion for the number of such graphs is

$$H_{\ell,0,1}=\frac{1}{2(\ell+3)}\,H_{1,\ell}+\frac{\ell-\frac{1}{2}}{\ell+3}\,H_{\ell-1,0,1} \tag{B.4}$$

To obtain the asymptotic behavior, we use (19) and again take the continuum limit to recast (B.4) into

$$\left(\ell\,\frac{\mathrm{d}}{\mathrm{d}\ell}+\frac{7}{2}\right)H_{\ell,0,1}\simeq\frac{2}{3\ell}, \tag{B.5}$$

from which

$$H_{\ell,0,1}\simeq\frac{4}{15\ell}. \tag{B.6}$$

We have thus identified three-hub configurations whose occurrence probabilities in the ensemble of networks of N nodes is proportional to N^{-1}. By extending this reasoning, we anticipate that the probability to find N-node networks of h hubs in the full ensemble will also be proportional to N^{-1}.

References

[1] Na H S and Rapoport A 1970 *Math. Biosci.* **6** 313
[2] Moon J W 1974 *London Mathematical Society Lecture Notes Series 13* (Cambridge: Cambridge University Press) pp 125–32
[3] Meir A and Moon J W 1978 *Can. J. Math.* **30** 997
[4] Kleinberg J, Kumar R, Raghavan P, Rajagopalan S and Tomkins A 1999 *Proc. of the Int. Conf. on Combinatorics and Computing* (*Lecture Notes in Computer Science* vol 1627) (Berlin: Springer) pp 1–18
[5] Krapivsky P L and Redner S 2001 *Phys. Rev.* E **63** 066123
[6] Vázquez A 2003 *Phys. Rev.* E **67** 056104
[7] Ben-Naim E and Krapivsky P L 2010 *J. Stat. Mech.* P06004
[8] Krapivsky P L and Redner S 2005 *Phys. Rev.* E **71** 036118
[9] Rozenfeld H D and ben-Avraham D 2004 *Phys. Rev.* E **70** 056107
[10] Gabel A and Redner S 2013 *J. Stat. Mech.* P02043
[11] Gabel A, Krapivsky P L and Redner S 2013 *Phys. Rev.* E **88** 050802
Gabel A, Krapivsky P L and Redner S 2014 *J. Stat. Mech.* P04009
[12] Wasserman S and Faust K 1994 *Social Network Analysis* (Cambridge: Cambridge University Press)

J. Stat. Mech. (2017) 073405

[13] Frasco G F, Sun J, Rozenfeld H D and ben-Avraham D 2014 *Phys. Rev.* X **4** 011008
[14] ben-Avraham D, Krapivsky P L and Redner S unpublished
[15] Buldyrev S V, Parshani R, Paul G, Stanley H E and Havlin S 2010 *Nature* **464** 1025
[16] Boccaletti S, Bianconi G, Criado R, Del Genio C I, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 *Phys. Rep.* **544** 1
[17] Kivelä M, Arenas A, Barthelemy M, Gleeson J P, Moreno Y and Porter M A 2014 *J. Complex Netw.* **2** 203
[18] D'Agostino G and Scala A 2014 *Networks of Networks: the Last Frontier of Complexity* (Berlin: Springer)
[19] Newman M E J 2010 *Networks: an Introduction* (Oxford: Oxford University Press)
[20] Graham R L, Knuth D E and Patashnik O 1989 *Concrete Mathematics: a Foundation for Computer Science* (Reading, MA: Addison-Wesley)
[21] Cayley A 1889 *Q. J. Pure Appl. Math.* **23** 376
[22] Otter R 1948 *Ann. Math.* **49** 583
[23] Aigner M and Ziegler G M 1998 *Proofs from the Book* (Berlin: Springer)
[24] Derrida B 1997 *Physica* D **107** 186
[25] Derrida B and Flyvbjerg H 1987 *J. Phys. A: Math. Gen.* **20** 5273
Derrida B and Flyvbjerg H 1987 *Journale de Physique* **48** 971
Derrida B and Bessis D 1988 *J. Phys. A: Math. Gen.* **21** L509
[26] Higgs P G 1995 *Phys. Rev.* E **51** 95
[27] Frachebourg L, Ispolatov I and Krapivsky P L 1995 *Phys. Rev.* E **52** R5727
[28] Derrida B and Jung-Muller B 1999 *J. Stat. Phys.* **94** 277
[29] Krapivsky P L, Grosse I and Ben-Naim E 2000 *Phys. Rev.* E **61** R993
Krapivsky P L, Ben-Naim E and Grosse I 2004 *J. Phys. A: Math. Gen.* **37** 2863
[30] Krapivsky P L and Krioukov D 2008 *Phys. Rev.* E **78** 026114
[31] Pippenger N and Schleich K 2006 *Random Struct. Algorithms* **28** 247
Fleming K and Pippenger N 2010 *Random Struct. Algorithms* **37** 465
[32] Linial N and Meshulam R 2006 *Combinatorica* **26** 475
[33] Babson E, Hoffman C and Kahle M 2011 *J. Am. Math. Soc.* **24** 1
[34] Zuev K, Eisenberg O and Krioukov D 2015 *J. Phys. A: Math. Theor.* **48** 465002
[35] Bianconi G, Darst R K, Iacovacci J and Fortunato S 2014 *Phys. Rev.* E **90** 042806
Bianconi G and Rahmede R 2016 *Phys. Rev.* E **93** 032315
[36] Linial N and Peled Y 2016 *Ann. Math.* **184** 745
[37] Muchnik L, Pei S, Parra L C, Reis S D S, Andrade J S Jr, Havlin S and Makse H A 2013 *Sci. Rep.* **3** 1783

J. Stat. Mech. (2017) 073405