Probability in the Engineering and Informational Sciences (2024), **38:**1 100–114
doi:10.1017/S0269964822000523

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Random multi-hooking networks

Kiran R. Bhutani[1], Ravi Kalpathy[1]* and Hosam Mahmoud[2]

[1]Department of Mathematics, The Catholic University of America, Washington, DC 20064, USA
[2]Department of Statistics, The George Washington University, Washington, DC 20052, USA.
*Corresponding author. E-mail: kalpathy@cua.edu

Keywords: Degree profile, Hooking networks, Random graph, Recurrence, Asymptotic analysis

AMS Subject Classifications: Primary: 05C82, 90B15, Secondary: 60C05, 05C12

Abstract
We introduce a broad class of multi-hooking networks, wherein multiple copies of a seed are hooked at each step at random locations, and the number of copies follows a predetermined building sequence of numbers. We analyze the degree profile in random multi-hooking networks by tracking two kinds of node degrees—the local average degree of a specific node over time and the global overall average degree in the graph. The former experiences phases and the latter is invariant with respect to the type of building sequence and is somewhat similar to the average degree in the initial seed. We also discuss the expected number of nodes of the smallest degree. Additionally, we study distances in the network through the lens of the average total path length, the average depth of a node, the eccentricity of a node, and the diameter of the graph.

1. Introduction

Trees have long been in the focus of research on random graphs. The classic types, such as those that appear in data structures [3,10,14,15] and digital processing [6,12,20], grow incrementally, one node at a time. In more recent times, authors considered more complex types of random graphs grown by adjoining entire graphs to a growing network [1,2,5,7,11,13,16,21]. We consider a growing network model in which the number of components attached at a stage follows a predetermined building sequence of numbers.

Societies and social networks grow and change over time in multiple random ways, which include growth patterns that add "components" at each step. Networks grown by adding components reflect these dynamics better than networks evolving on single node additions. One can embed a graph in a predetermined growth structure leading to multiple scenarios of growing networks.

In this paper, we develop a model where networks grow by hooking multiple copies of the seed at multiple nodes of the growing network chosen in a random fashion and study the theoretical and statistical properties of the networks so generated.

2. The building sequence

We assume that a network grows by attaching a number of components at each step to the existing structure, which starts with $\tau_0 \geq 2$ vertices. In the next subsection, we give a formal definition. Here, we only say a word on the number of components added at each step. After n steps of growth, the number of components attached to obtain the next network is k_n, a predetermined sequence of nonnegative numbers.

2.1. Regularity conditions

Let $\tau_0 \geq 2$. This represents the number of nodes in a building block (a seed). We grow the network by adding a number of copies of the seed at places called latches. At each latch, a designated vertex in the seed (called the hook) is fused with the latch. A formal definition of this process is given in the sequel.

We shall consider adding $k_n \geq 1$ copies of the seed to construct the $(n+1)$st network, under the following *regularity conditions*:

(R1) $k_n \leq \tau_0 + (\tau_0 - 1)\sum_{i=0}^{n-1} k_i$.
(R2) $\lim_{n\to\infty} k_n/\sum_{i=0}^{n} k_i = a \in [0, 1]$.
(R3) $\lim_{n\to\infty} \tau_0/k_n = b \geq 0$.

A sequence of nonnegative integers $\{k_n\}_{n=0}^{\infty}$ satisfying (R1)–(R3) is called a building sequence. Condition (R1) is to guarantee the feasibility of choosing latches. At no point in time does the process require more (distinct) latches than the number of nodes existing in the network. Conditions (R2)–(R3) facilitate the existence of limits for properties of interest and expedite finding their values. Note that $a = 0$ and $b = 0$ are both allowed. For instance, for a constant sequence $k_n = k \in \mathbb{N}$, we have $a = 0$, and $b = \tau_0/k > 0$, whereas when $k_n = n + 1$, we have both $a = 0$ and $b = 0$.

Regularity conditions (R1)–(R3) are not too restrictive and the class covered by the investigation remains very broad. The examples that come up in practice satisfy these regularity conditions. For example, at one extreme, the building sequence $k_n = 1$ builds networks of linear growth, including trees. At the other extreme, the case of equality in Condition (R1) builds a deterministic network where the entire vertex set is chosen at each step (a take-all model); such extremal case grows the network exponentially fast.

3. The multi-hooking network

A network grows as follows. We start with a connected *seed* graph G_0 with vertex set of size τ_0 and edge set of size η. One of the vertices in the seed is designated as a hook (vertex h). When a copy of the seed is adjoined to the network, it is the seed's hook that latches into that larger graph. The hooking is accomplished by fusing together the hook and a latch (vertex) chosen from the network.

At step n, k_{n-1} copies of the seed are hooked into the graph, $G_{n-1} = (V_{n-1}, \mathcal{E}_{n-1})$, with vertex set V_{n-1} and edge set $\mathcal{E}_{n-1}$, that exists at time $n-1$. To complete the nth hooking step, we sample k_{n-1} latches from the graph G_{n-1}. The selection mechanism can take a number of forms, such as choosing distinct hooks as opposed to allowing repetitions.

We use the notation $|A|$ for the cardinality of a set A. We consider a *uniform* model that selects k_{n-1} *distinct* nodes in the network, with all $\binom{|V_{n-1}|}{k_{n-1}}$ subsets being equally likely. In the language of statistics, this boils down to *sampling without replacement*.

Figure 1 illustrates a seed and a network grown from it in three steps under the building sequence $k_n = n+1$. So, G_1 grows by choosing a latch from G_0 (the starred node in G_0), G_2 grows by choosing the two starred nodes from G_1, and G_3 grows by choosing the three starred nodes from G_2. The networks in Figure 1 have loops and multiple edges, as we do not restrict the study to *simple* graphs.

3.1. Notation

The notation Hypergeo(t, r, s) stands for the hypergeometric random variable associated with the random sampling of s objects out of a total of t objects, of which r objects are of a special type. So, the hypergeometric random variable counts the number of special objects in the sample.

It is customary to call the cardinality of the vertex set of a graph the *order of the graph* and reserve the term *size of the graph to* the cardinality of the set of edges in the graph. Let V_n be the set of vertices

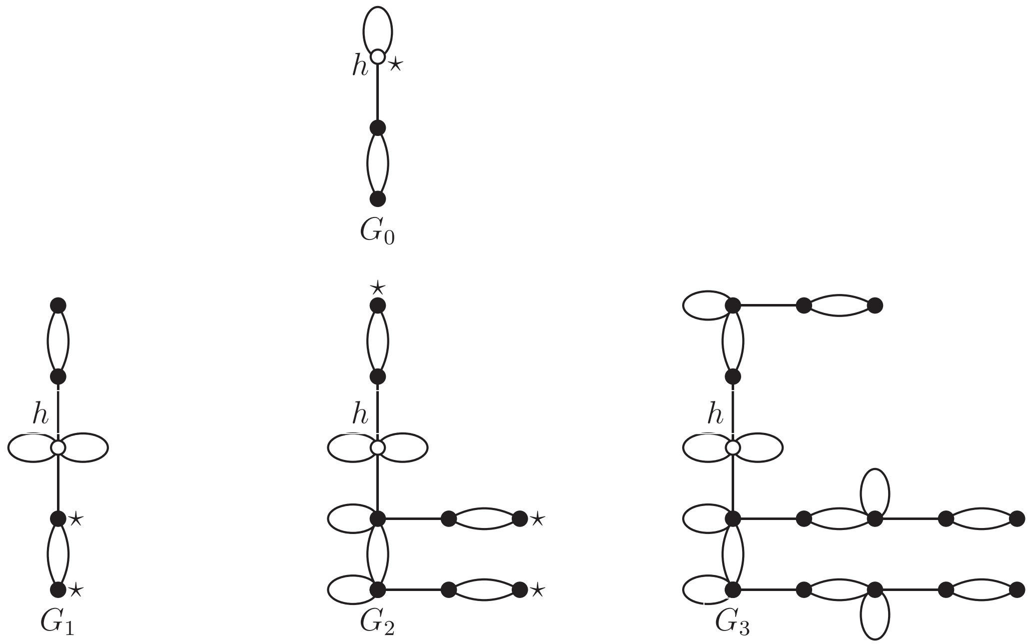

Figure 1. *A seed (top) with a hook and three networks grown from it (second row) under the building sequence $k_n = n + 1$. The white vertices in the network G_1, G_2, and G_3 represent the reference vertex.*

of the graph G_n, and $\mathcal{E}_n$ be the set of edges of that graph. Thus, the seed $G_0 = (V_0, \mathcal{E}_0)$ is a connected graph with the set V_0 of vertices and the set $\mathcal{E}_0$ of edges.

Let τ_n be the order of the graph at age n. Hence, the cardinality of the vertex set V_0 of the seed is $|V_0| = \tau_0$. The nth hooking step adds k_{n-1} copies of the seed at k_{n-1} distinct latches chosen uniformly at random from G_{n-1}. Each copy contributes $\tau_0 - 1$ new vertices to the network. The reason for subtracting 1 is the absorption of the hook. This gives the recurrence

$$\tau_n = \tau_{n-1} + k_{n-1}(\tau_0 - 1). \tag{1}$$

Unwinding this recurrence, we obtain

$$\tau_n = (\tau_0 - 1)\sum_{i=1}^{n} k_{i-1} + \tau_0. \tag{2}$$

We use the notation $\deg(v)$ to denote the degree of node v in a given graph, and we set $h^* = \deg(h)$.

3.2. Useful limits

By the regularity conditions, we can argue from (2) that

$$\frac{\tau_n}{k_{n-1}} = (\tau_0 - 1)\frac{1}{k_{n-1}}\sum_{i=0}^{n-1} k_i + \frac{\tau_0}{k_{n-1}} \to \frac{\tau_0 - 1}{a} + b =: \gamma.$$

Reorganize (1) as

$$1 = \frac{\tau_{n-1}}{\tau_n} + \frac{k_{n-1}}{\tau_n}(\tau_0 - 1).$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

to find the limit

$$\frac{\tau_{n-1}}{\tau_n} = 1 - \frac{k_{n-1}}{\tau_n}(\tau_0 - 1) \to 1 - \frac{(\tau_0 - 1)}{\gamma}.$$

4. A degree profile of the network

Various aspects of the degrees of nodes in a network are of interest in different contexts. For example, in the language of epidemiology, the degree of a node may be a useful representation of a highly infective person. From a health policy point of view, having knowledge about the degrees in conjunction with other graph parameters may help in identifying hot spots that trigger outbreaks and may be useful in controlling and mitigating the contagion. In the context of a social network, the degree of a node may represent the popularity and social skills of the person represented by the node.

Equally interesting are the global overall average degree in the entire graph (where we look at all the nodes), the local degree of a specific node during its temporal evolution, and the number of nodes of the smallest degree. We deal with the average behavior of each of these in a separate subsection. The different aspects of the degree complete a profile of the graph.

4.1. Evolution of the degree of a specific node

Suppose a node appears for the first time at step j. What will become of its degree at step n? At step j, several copies are added. To avoid a heavy notation identifying the time of appearance j, the copy number, which node within the copy to be tracked, and n, we use a simpler notation that needs only j and n, for after all nodes of the same degree in the seed have the same distribution over time.

Theorem 4.1. *Suppose $\{k_n\}_{n=0}^{\infty}$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^{\infty}$. Let $X_{j:n}$ be the degree of a node at time n that had appeared for the first time at step j. If initially its degree (in the seed) is δ, then we have*

$$\mathbb{E}[X_{j:n}] = \delta + h^* \sum_{i=j}^{n-1} \frac{k_i}{\tau_i} = \frac{h^* a}{(1-a)(\tau_0 - 1) + ab}(n-j) + o(n-j) + O(1).$$

Proof. Suppose a node v appears at time j for the first time. So, it belongs to one of the copies adjoined to the graph at that time. As the graph evolves, in any single step, the degree of v can increase, if it is one of the nodes selected as latches in that step; otherwise its degree stays put, and when it does increase, it goes up by $h^* = \deg(h)$, the degree of the hook in the seed. This gives rise to a recurrence:

$$X_{j:n} = X_{j:n-1} + h^* \mathbb{I}_{n-1}(v),$$

where $\mathbb{I}_{n-1}(v)$ is an indicator of the event of choosing v among the k_{n-1} latches of that step of growth. On average, we have

$$\mathbb{E}[X_{j:n}] = \mathbb{E}[X_{j:n-1}] + h^* \frac{\binom{\tau_{n-1}-1}{k_{n-1}-1}}{\binom{\tau_{n-1}}{k_{n-1}}} = \mathbb{E}[X_{j:n-1}] + h^* \frac{k_{n-1}}{\tau_{n-1}}.$$

Unwinding the recurrence, we obtain the exact average:

$$\mathbb{E}[X_{j:n}] = \delta + h^* \sum_{i=j}^{n-1} \frac{k_i}{\tau_i}.$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

By the limits in Subsection 3.2, we obtain

$$\mathbb{E}[X_{j:n}] = O(1) + o(n-j) + \frac{h^* a}{(1-a)(\tau_0 - 1) + ab}(n-j).$$

□

Remark 4.1. If $a = 0$ the $\mathbb{E}[X_{j:n}]$ is only $o(n-j)$.

Remark 4.2. Consider the case $a > 0$. The average in Theorem 4.1 indicates that the degree of a specific node experiences phases. The degree of a node in the early phase with $j = j(n) = o(n)$ grows linearly with its age in the network. When $j(n) \sim \rho n$, for $0 < \rho < 1$, we still get a linear growth, but the coefficient of linearity is attenuated to $(1-\rho)(h^* a/[(1-a)(\tau_0 - 1) + ab])$. At $\rho = 1$, we have $\mathbb{E}[X_{j:n}] = o(n)$.

Remark 4.3. If $a = 0$, we can only assert that $\mathbb{E}[X_{j:n}] = o(n-j) + \delta$. In this case, a finer analysis is needed to identify the leading order of the average degree of a node that appears at time j. For instance, in the case of a tree grown from the complete graph K_2, we have $\delta = 1$, $h^* = 1$, $k_n = 1$, and $a = 0$. The exact formula in this case yields

$$\mathbb{E}[X_{j:n}] = 1 + \sum_{i=j}^{n-1} \frac{1}{i+2} = H_{n+1} - H_{j+1} + 1.$$

Whence, we have the phases

$$\mathbb{E}[X_{j:n}] \sim \begin{cases} \ln n, & \text{if } j \text{ is fixed;} \\ \ln \dfrac{n}{j(n)}, & \text{if } j(n) \to \infty \text{ and } j(n) = o(n); \\ \ln \dfrac{1}{\rho}, & \text{if } j(n) \sim \rho n,\ 0 < \rho < 1; \\ 1, & \text{if } j(n) = n - o(n). \end{cases}$$

4.2. *The overall average degree*

The main result about the overall average degree in the graph is developed in this section. The result is expressed in terms of η, the number of edges in the seed graph.

Theorem 4.2. *Suppose $\{k_n\}_{n=0}^{\infty}$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^{\infty}$. Let Y_n be the degree of a randomly chosen node in the graph G_n at age n. We have*

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = \frac{2\eta}{\tau_0 - 1}.$$

Proof. Upon hooking k_{n-1} copies of the seed to k_{n-1} distinct nodes of $G_{n-1} = (V_{n-1}, \mathcal{E}_{n-1})$, we add ηk_{n-1} edges to the graph. Therefore, we have

$$|\mathcal{E}_n| = |\mathcal{E}_{n-1}| + \eta k_{n-1}.$$

This recurrence has the solution

$$|\mathcal{E}_n| = \eta \left(1 + \sum_{i=0}^{n-1} k_i\right).$$

Using the classical First Theorem of Graph Theory, we obtain

$$\sum_{v \in V_n} \deg(v) = 2|\mathcal{E}_n| = 2\eta \left(1 + \sum_{i=0}^{n-1} k_i\right).$$

Scaling the equation by τ_n, we get

$$\frac{1}{\tau_n} \sum_{v \in V_n} \deg(v) = \frac{2\eta}{\tau_n} \left(1 + \sum_{i=0}^{n-1} k_i\right).$$

Taking limits, and using Eq. (2), we obtain

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = 0 + \lim_{n\to\infty} 2\eta \frac{\sum_{i=0}^{n-1} k_i}{\tau_n} = \frac{2\eta}{\tau_0 - 1}.$$

□

Remark 4.4. The average degree in the seed is $2\eta/\tau_0$. For any building sequence, the asymptotic average degree in the graph is $2\eta/(\tau_0 - 1)$, only slightly higher than the average degree in the initial seed. This should be anticipated because the additions introduce a number of copies of the seed, each of which has the degree properties of the seed with the hook eliminated.

4.3. Nodes of the smallest degree

We study only the nodes of the smallest degree. Let d^* be the smallest degree in the seed. Note that the smallest admissible degree in the graph is d^*. After the network grows, the smallest degree in it may be d^* or higher. Let X_n be the number of nodes of degree d^* at time n. Thus, X_0 is the number of nodes of degree d^* in the seed. Later graphs can have more nodes of degree d^*. The seed in Figure 1 has $d^* = 2$, and $X_0 = 1$, $X_1 = 2$, $X_2 = 3$, and $X_3 = 3$.

4.3.1. Stochastic recurrence

In the evolution at step n, we hook k_{n-1} copies of the seed to the graph G_{n-1}. Let A_0 be the event $\deg(h) = d^*$ and $\mathbb{I}_{A_0}$ be an indicator that assumes value 1, if $\deg(h) = d^*$, otherwise, it assumes the value 0. A latch of degree d^* in the sample will have a higher degree (namely, its degree goes up to $d^* + \deg(h)$) in G_n. So, we lose such vertices in the count of X_n. If the hook degree is d^*, every hooked copy contributes only $X_0 - 1$ vertices of degree d^*.

For the case when $k_{n-1} = 1$ and the latch is ℓ, the change from X_{n-1} to X_n for the four cases can be seen as shown in the table below:

	$\deg(\ell) = d^*$	$\deg(\ell) \neq d^*$
$\deg(h) = d^*$	$(X_{n-1} - 1) + (X_0 - 1)$	$X_{n-1} + (X_0 - 1)$
$\deg(h) \neq d^*$	$(X_{n-1} - 1) + X_0$	$X_{n-1} + X_0$

Thus, for any value of k_{n-1}, the count X_n therefore satisfies a (conditional) stochastic recurrence:

$$\begin{aligned} X_n &= X_{n-1} + (X_0 - 1 - \mathbb{I}_{A_0})\, \text{Hypergeo}(\tau_{n-1}, X_{n-1}, k_{n-1}) \\ &\quad + (X_0 - \mathbb{I}_{A_0})(k_{n-1} - \text{Hypergeo}(\tau_{n-1}, X_{n-1}, k_{n-1})). \end{aligned} \tag{3}$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

4.3.2. The average proportion of nodes of degree d^*

Take (conditional) expectation of (3) to get

$$\begin{aligned}\mathbb{E}[X_n \mid G_{n-1}] &= X_{n-1} + (X_0 - 1 - \mathbb{I}_{A_0})\frac{X_{n-1}}{\tau_{n-1}}k_{n-1} \\ &\quad + (X_0 - \mathbb{I}_{A_0})\left(k_{n-1} - \frac{X_{n-1}}{\tau_{n-1}}k_{n-1}\right) \\ &= \left(1 - \frac{k_{n-1}}{\tau_{n-1}}\right)X_{n-1} + (X_0 - \mathbb{I}_{A_0})k_{n-1}. \end{aligned} \tag{4}$$

Theorem 4.3. *Suppose $\{k_n\}_{n=0}^{\infty}$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^{\infty}$, starting from a seed with X_0 nodes of the smallest degree d^*. Let X_n be the number of vertices of this degree in the graph after n steps of evolution according to the building sequence. We have*

$$\mathbb{E}[X_n] = (X_0 - \mathbb{I}_{A_0})\sum_{i=1}^{n} k_{i-1}\prod_{j=i+1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + X_0\prod_{j=1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right).$$

Subsequently, the average proportion converges to a limit independent of the limits a and b; namely we have the convergence

$$\mathbb{E}\left[\frac{X_n}{\tau_n}\right] \to \frac{X_0 - \mathbb{I}_{A_0}}{\tau_0}.$$

Proof. Taking a double expectation of (4) yields

$$\mathbb{E}[X_n] = \left(1 - \frac{k_{n-1}}{\tau_{n-1}}\right)\mathbb{E}[X_{n-1}] + (X_0 - \mathbb{I}_{A_0})k_{n-1}. \tag{5}$$

This recurrence equation is of the standard linear form

$$y_n = g_n y_{n-1} + h_n, \tag{6}$$

with solution

$$y_n = \sum_{i=1}^{n} h_i \prod_{j=i+1}^{n} g_j + y_0 \prod_{j=1}^{n} g_j. \tag{7}$$

So, the sought solution for the average of the number of nodes of degree d^* (for $n \geq 1$) is

$$\mathbb{E}[X_n] = (X_0 - \mathbb{I}_{A_0})\sum_{i=1}^{n} k_{i-1}\prod_{j=i+1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + X_0\prod_{j=1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right).$$

The strategy for the asymptotic part of the statement is two-fold: We prove the existence of a limit (under any building sequence) for the proportion from the exact solution. We then find the value of the limit from the recurrence under the mild regularity conditions imposed on the building sequence.

First, express the expected proportion as

$$\mathbb{E}\left[\frac{X_n}{\tau_n}\right] = (X_0 - \mathbb{I}_{A_0})\sum_{i=1}^{n} \frac{k_{i-1}}{\tau_n}\prod_{j=i+1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + \frac{X_0}{\tau_n}\prod_{j=1}^{n}\left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right); \tag{8}$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

at $i = n$, the first product does not exist, and is taken to be 1, as usual. Let

$$c_n = \sum_{i=1}^{n} \frac{k_{i-1}}{\tau_n} \prod_{j=i+1}^{n} \left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) = \sum_{i=1}^{n-1} \frac{k_{i-1}}{\tau_n} \prod_{j=i+1}^{n} \left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + \frac{k_{n-1}}{\tau_n};$$

We manipulate this to turn it into a recurrence as follows:

$$\begin{aligned}
c_{n+1} &= \sum_{i=1}^{n} \frac{k_{i-1}}{\tau_{n+1}} \prod_{j=i+1}^{n+1} \left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + \frac{k_n}{\tau_{n+1}} \\
&= \frac{\tau_n}{\tau_{n+1}} \left(1 - \frac{k_n}{\tau_n}\right) \sum_{i=1}^{n} \frac{k_{i-1}}{\tau_n} \prod_{j=i+1}^{n} \left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right) + \frac{k_n}{\tau_{n+1}} \\
&= \left(\frac{\tau_n - k_n}{\tau_{n+1}}\right) c_n + \frac{k_n}{\tau_{n+1}} \\
&= \left(\frac{\tau_{n+1} - (\tau_0 - 1)k_n - k_n}{\tau_{n+1}}\right) c_n + \frac{k_n}{\tau_{n+1}} \\
&= \left(\frac{\tau_{n+1} - \tau_0 k_n}{\tau_{n+1}}\right) c_n + \frac{k_n}{\tau_{n+1}}
\end{aligned}$$

Rearrange the recurrence in the form

$$c_{n+1} - \frac{1}{\tau_0} = c_n - \frac{\tau_0 k_n}{\tau_{n+1}} c_n + \frac{k_n}{\tau_{n+1}} - \frac{1}{\tau_0} = \left(c_n - \frac{1}{\tau_0}\right)\left(\frac{\tau_{n+1} - \tau_0 k_n}{\tau_{n+1}}\right),$$

leading to the inequality

$$\begin{aligned}
\left|c_{n+1} - \frac{1}{\tau_0}\right| &\le \left|c_n - \frac{1}{\tau_0}\right| \left|\frac{\tau_{n+1} - \tau_0 k_n + k_n}{\tau_{n+1}}\right| \\
&= \left|c_n - \frac{1}{\tau_0}\right| \frac{\tau_n}{\tau_{n+1}} \\
&\le \left|c_{n-1} - \frac{1}{\tau_0}\right| \frac{\tau_n}{\tau_{n+1}} \times \frac{\tau_{n-1}}{\tau_n} \\
&\vdots \\
&\le \left|c_0 - \frac{1}{\tau_0}\right| \frac{\tau_n}{\tau_{n+1}} \times \frac{\tau_{n-1}}{\tau_n} \times \cdots \times \frac{\tau_0}{\tau_1}.
\end{aligned}$$

Noting that the sum in c_n is empty at $n = 0$, we have $c_0 = 0$ and the bounds simplify to $0 \le |c_n - 1/\tau_0| \le 1/\tau_n$. So, both inferior and superior limits of $|c_n - 1/\tau_0|$ are equal to 0, which furnishes the existence of a limit for c_n equal to $1/\tau_0$, too.

As for the remainder part

$$r_n := \frac{X_0}{\tau_n} \prod_{j=1}^{n} \left(1 - \frac{k_{j-1}}{\tau_{j-1}}\right),$$

in (8), it clearly converges to 0, as τ_n is increasing, and the product is bounded from above by 1. Plugging in the limits $\lim_{n\to\infty} c_n = 1/\tau_0$ and $\lim_{n\to\infty} r_n = 0$ in (8), we reach the conclusion that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{X_n}{\tau_n}\right] = \frac{X_0 - \mathbb{I}_{A_0}}{\tau_0}.$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

□

Remark 4.5. In the case when the hook is not of the smallest degree d^*, we have $\mathbb{E}[X_n/\tau_n] \to X_0/\tau_0$. The initial proportion of nodes of the smallest degree in the seed is preserved on average in larger subsequent graphs.

Remark 4.6. In the case when the hook is of the smallest degree d^*, we have $\mathbb{E}[X_n/\tau_n] \to (X_0 - 1)/\tau_0$. The long-term proportion of nodes of the smallest degree is less than the proportion of nodes of degree d^* in the seed.

Remark 4.7. In the case when the hook is the only node of the smallest degree d^* in the seed, we have $X_0 = 1$, and $\mathbb{E}[X_n/\tau_n] \to 0$, for all $n \geq 1$. Indeed, the degree d^* disappears after the first latching at the initial hook and never reappears.

Remark 4.8. The limit in Theorem 4.3 is more than just an ultimate value in the take-all case. In this case, it is the actual value for each $n \geq 0$, which can be seen from the recurrence. The only term that does not vanish is the last term in sum, yielding $(X_0 - I_{A_0})k_{n-1}/\tau_n = (X_0 - I_{A_0})/\tau_0$.

5. Distances in the network

We measure node distances in G_n relative to a reference point (vertex). We take the reference to be the hook of G_0. We look at two (related) kinds of distances: The total path length and the average distance in the graph. Let the nodes of the nth graph be labeled with the numbers $1, 2, \ldots, \tau_n$, with 1 being reserved for the reference vertex and the rest of the nodes are arbitrarily assigned distinct numbers from the set $\{2, \ldots, \tau_n\}$. The depth of a node in the network is its distance from the reference vertex (i.e., the length of the shortest path from the node to the reference vertex measured in the number of edges). We denote the depth of the ith node in the nth network by $\Delta_{n,i}$ The *total path length* is the sum of all depths; namely it is

$$T_n = \sum_{i=1}^{\tau_n} \Delta_{n,i}.$$

For instance, the networks G_0, G_1, G_2, and G_3 in Figure 1 have total path lengths $T_0 = 3, T_1 = 6, T_2 = 18$, and $T_3 = 45$, respectively.

5.1. *Average total path length*

As the network grows, at step n, a sample of size k_{n-1} latches is chosen from G_{n-1} to grow G_{n-1} into G_n. Suppose these latches are $\ell_1, \ldots, \ell_{k_{n-1}}$ at depths $d_1, \ldots, d_{k_{n-1}}$. In view of the absorption of the hooks of the added graphs, a copy's hook fused at the jth latch adds $\tau_0 - 1$ nodes, which appear in G_n at depths equal to their distance from the hook of the copy translated by an additional distance from the latch to the reference vertex. So, collectively, the vertices of the copy hooked to ℓ_j increase the total path length by $T_0 + (\tau_0 - 1)d_j$. We have a conditional recurrence:

$$\mathbb{E}[T_n \mid G_{n-1}, d_1, \ldots, d_{k_{n-1}}] = T_{n-1} + \sum_{j=1}^{k_{n-1}} (T_0 + (\tau_0 - 1)d_j).$$

Averaging over the graphs and the choices of the latches within, we get

$$\mathbb{E}[T_n] = \mathbb{E}[T_{n-1}] + T_0 k_{n-1} + (\tau_0 - 1) \sum_{j=1}^{k_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j}]. \tag{9}$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

Lemma 5.1.

$$\sum_{j=1}^{k_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j}] = \frac{k_{n-1}}{\tau_{n-1}}\,\mathbb{E}[T_{n-1}].$$

Proof. Condition on the event $\ell_1 = i_1, \ldots, \ell_{k_{n-1}} = i_{k_{n-1}}$, to get

$$\sum_{j=1}^{k_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j}] = \sum_{j=1}^{k_{n-1}} \sum_{1\le i_1<i_2<\cdots<i_{k_{n-1}}\le \tau_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j} \mid d_1 = i_1, \ldots d_{k_{n-1}} = i_{k_{n-1}}]$$
$$\times\, \mathbb{P}(d_1 = i_1, \ldots, d_{k_{n-1}} = i_{k_{n-1}}).$$

The subsets of size k_{n-1} latches that appear in a sample of vertices from G_{n-1} are all equally likely, and we get

$$\sum_{j=1}^{k_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j}]$$
$$= \frac{1}{\binom{\tau_{n-1}}{k_{n-1}}} \sum_{j=1}^{k_{n-1}} \mathbb{E}\left[\left(\sum_{1\le i_1<i_2<\cdots<i_{k_{n-1}}\le \tau_{n-1}} \Delta_{n-1,\ell_j}\right)\middle|\, \ell_1 = i_1, \ldots \ell_{k_{n-1}} = i_{k_{n-1}}\right].$$

Let us write out the inner sum in expanded form:

$$\begin{aligned}
&\Delta_{n-1,1} + \Delta_{n-1,2} + \cdots + \Delta_{n-1,k_{n-1}}\\
&\quad + \Delta_{n-1,1} + \Delta_{n-1,2} + \cdots + \Delta_{n-1,k_{n-1}-1} + \Delta_{n-1,k_{n-1}+1}\\
&\quad \vdots\\
&\quad + \Delta_{n-1,\tau_{n-1}-k_{n-1}+1} + \Delta_{n-1,\tau_{n-1}-k_{n-1}+2} + \cdots + \Delta_{n-1,\tau_{n-1}}.
\end{aligned}$$

Upon a reorganization collecting similar terms, we get

$$\binom{\tau_{n-1}-1}{k_{n-1}-1}(\Delta_{n-1,1} + \Delta_{n-1,2} + \cdots + \Delta_{n-1,\tau_{n-1}}) = \binom{\tau_{n-1}-1}{k_{n-1}-1} T_{n-1}.$$

Plugging this expression in the expectation, we proceed to

$$\sum_{j=1}^{k_{n-1}} \mathbb{E}[\Delta_{n-1,\ell_j}] = \frac{\binom{\tau_{n-1}-1}{k_{n-1}-1}}{\binom{\tau_{n-1}}{k_{n-1}}}\mathbb{E}[T_{n-1}] = \frac{k_{n-1}}{\tau_{n-1}}\mathbb{E}[T_{n-1}].$$

□

Theorem 5.1. *Suppose $\{k_n\}_{n=0}^{\infty}$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^{\infty}$, starting from a seed of total path length T_0. Let T_n be the total path length after n steps of evolution according to the building sequence. We have*

$$\mathbb{E}[T_n] = T_0\tau_n\left(\sum_{i=1}^{n} \frac{k_{i-1}}{\tau_i} + \frac{1}{\tau_0}\right).$$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

Proof. By Lemma 5.1 and the recurrence (9), we have a recurrence for the average total path length:

$$\begin{aligned}\mathbb{E}[T_n] &= \mathbb{E}[T_{n-1}] + T_0 k_{n-1} + \frac{(\tau_0 - 1)k_{n-1}}{\tau_{n-1}}\mathbb{E}[T_{n-1}]\\ &= \left(1 + \frac{(\tau_0 - 1)k_{n-1}}{\tau_{n-1}}\right)\mathbb{E}[T_{n-1}] + T_0 k_{n-1}.\end{aligned}$$

Again, the recurrence is of the standard form (6) with the solution (7). In the specific case at hand, this solution is

$$\mathbb{E}[T_n] = T_0 \sum_{i=1}^{n} k_{i-1} \prod_{j=i+1}^{n} \left(1 + \frac{(\tau_0 - 1)k_{j-1}}{\tau_{j-1}}\right) + T_0 \prod_{j=1}^{n} \left(1 + \frac{(\tau_0 - 1)k_{j-1}}{\tau_{j-1}}\right).$$

The recurrence (1) on the order of the graph simplifies the solution into telescopic products

$$\mathbb{E}[T_n] = T_0 \sum_{i=1}^{n} k_{i-1} \prod_{j=i+1}^{n} \frac{\tau_j}{\tau_{j-1}} + T_0 \prod_{j=1}^{n} \frac{\tau_j}{\tau_{j-1}} = T_0 \tau_n \left(\sum_{i=1}^{n} \frac{k_{i-1}}{\tau_i} + \frac{1}{\tau_0}\right).$$

□

5.2. Average depth

Theorem 5.1 provides a benchmark for the calculation of the average depth. Let the depth of a randomly selected node in the nth network be D_n.

Corollary 5.1.

$$\mathbb{E}[D_n] = T_0 \sum_{i=1}^{n} \frac{k_{i-1}}{\tau_i} + D_0.$$

Proof. Given a specific development leading to G_n, the average depth in that graph is

$$\mathbb{E}[D_n \mid G_n] = \frac{\Delta_{n,1} + \cdots + \Delta_{n,\tau_n}}{\tau_n} = \frac{T_n}{\tau_n}.$$

Upon taking expectation, it follows that $\mathbb{E}[D_n] = \mathbb{E}[T_n]/\tau_n$. The form given in the statement ensues from Theorem 5.1. □

Corollary 5.2. *Under the regularity conditions* (R1)–(R3)*, we have the asymptotic equivalent*

$$\mathbb{E}[D_n] = \frac{aT_0}{\tau_0 - 1 + ab} n + o(n), \quad \text{as } n \to \infty.$$

Remark 5.1. Corollary 5.2 is more useful when

$$\lim_{n\to\infty} \frac{k_n}{\sum_{i=0}^{n} k_i} = a \neq 0.$$

When $a = 0$, as in the case of trees for example, one needs to sharpen the argument to find the leading asymptotic term, as we do in some specific cases below.

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

5.3. Distances under specific building sequences

At one extreme, there is the sequence of least possible growth ($k_n = 1$). At the other extreme, we have a take-all model ($k_n = \tau_n$) in which all the nodes of G_n are taken as latches for τ_n copies of the seed.

In the case of $k_n = k$, for fixed $k \in \mathbb{N}$, of nearly the least growth, the average depth is

$$\mathbb{E}[D_n] = T_0 \sum_{i=1}^{n} \frac{k}{\tau_i} + D_0.$$

The limit a in regularity condition (R2) is 0, and Corollary 5.2 only tells us that $\mathbb{E}[D_n] = o(n)$. However, we can sharpen the asymptotic equivalence from the specific construction of the case.

Here, we have $\tau_i = (\tau_0 - 1)i + \tau_0$, which gives

$$\mathbb{E}[D_n] = \frac{T_0 k}{\tau_0 - 1} \sum_{i=1}^{n} \frac{1}{i + \frac{\tau_0}{\tau_0 - 1}} + D_0.$$

In terms of the generalized harmonic numbers[1]

$$H_n(x) = \frac{1}{1+x} + \frac{1}{2+x} + \cdots + \frac{1}{n+x},$$

the depth in the near-least-growth is compactly expressed as

$$\mathbb{E}[D_n] = \frac{T_0 k}{\tau_0 - 1} H_n\left(\frac{\tau_0}{\tau_0 - 1}\right) + D_0 \sim \frac{T_0 k}{\tau_0 - 1} \ln n, \quad \text{as } n \to \infty.$$

Remark 5.2. The case $k = 1$ and $\tau_0 = 2$ grows a recursive tree. The seed is a rooted tree on two vertices, in which $T_0 = 1$ and $D_0 = \frac{1}{2}$. In this case, the average depth becomes

$$\mathbb{E}[D_n] = H_n(2) + \frac{1}{2} = H_{n+2} - 1 \sim \ln n, \quad \text{as } n \to \infty,$$

which recovers a known result [19].

Remark 5.3. At the other end of the spectrum, there is the take-all model, in which $k_i = \tau_i$, leading at step n to a graph of order $\tau_n = \tau_0^{n+1}$. Here, the limit a is $(\tau_0 - 1)/\tau_0$ and the limit b is 0. According to Corollary 5.2, we have $\mathbb{E}[D_n] \sim D_0 n$, as $n \to \infty$. This asymptotic estimate can be sharpened as the case is amenable to exact calculation:

$$\mathbb{E}[D_n] = T_0 \sum_{i=1}^{n} \frac{\tau_{i-1}}{\tau_i} + D_0 = T_0 \sum_{i=1}^{n} \frac{1}{\tau_0} + D_0 = D_0(n+1).$$

6. Eccentricity

The eccentricity $C(v)$ of a node v in a graph $\mathcal{G}$ is the distance between v and a vertex farthest from v in $\mathcal{G}$. The eccentricity is instrumental in constructing a notion of the diameter of a graph (extreme distances). We use the eccentricity of the hook and the various latches selected in G_{n-1} to determine the diameter of the graph G_n.

The eccentricity is technically defined as follows. If Q is a path in a graph $\mathcal{G}$, we denote its length by $|Q|$ (the number of edges in it). There can be several paths joining two vertices u and v in $\mathcal{G}$, and the

[1] Customarily, $H_n(0)$ is denoted by H_n.

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

distance between u and v, denoted by $d(u, v)$, is the length of the shortest such path. That is, with $\mathcal{P}(u, v)$ denoting the collection of paths between u and v, the distance between these two nodes is given by

$$d(u, v) = \min_{Q \in \mathcal{P}(u,v)} |Q|.$$

The eccentricity $C(v)$ of a vertex v in a graph with vertex set $\mathcal{V}$ is:

$$C(v) = \max_{u \in \mathcal{V}} d(v, u) = \max_{u \in \mathcal{V}} \min_{Q \in \mathcal{P}(v,u)} |Q|.$$

For instance, the eccentricity in Figure 1 of the reference vertex of G_0 is 2, of the reference vertex in G_1 is 2 as well, but of the reference vertex in G_2 is 4 and becomes 6 in G_3.

6.1. Eccentricity of a node in G_n

The k_{n-1} nodes selected as latches from the graph $G_{n-1} = (V_{n-1}, \mathcal{E}_{n-1})$ are vertices that play a key role in designing the network at stage n and onward and contribute significantly in determining the diameter of the graph at the next stage.

As a node's eccentricity changes over time, its value at step n in G_n may be different from its value at step $n-1$ in G_{n-1}. We need an eccentricity notation reflecting the possible change over time. For that we use $C_n(v)$ to speak of the eccentricity of a vertex v in G_n.

If $v \in V_n$ is a vertex in a copy of G_0 latched at a vertex $\ell_i \in V_{n-1}$, we express that by saying $v \in V_0^{co_i}$, otherwise we say $v \in V_{n-1}$. We now introduce some notation:

1. $\mathbb{L}_n = \{\ell_1, \ell_2, \cdots \ell_{k_n}\}$ is the set of latches selected in the graph G_n to produce the graph G_{n+1}.
2. $\widetilde{C}_n(v) = C_n(v) \mid G_{n-1}, \mathbb{L}_{n-1}$. This is the *conditional* eccentricity $C_n(v)$ of the node v in the graph G_n, given G_{n-1} and the k_{n-1} latches in it.
3. For any $v \in V_{n-1}$, we define $d_v^{\#} = \max_{\ell_j \in \mathbb{L}_{n-1}} d(v, \ell_j)$. So, $d_v^{\#}$ is the maximum distance from v to the nodes in $\mathbb{L}_{n-1}$.

Also, in what follows we use the notation $\mathbb{I}_C$ to indicate a predicate (condition) C. So, it is 1, when C holds, and is 0, otherwise.

Theorem 6.1. *Suppose $\{k_n\}_{n=0}^{\infty}$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^{\infty}$. Let v be a node in the graph G_n. Conditional upon the choice of the latches $\ell_1, \ldots, \ell_{k_{n-1}}$ in G_{n-1}, the eccentricity $C_n(v)$ is given by*

$$\widetilde{C}_n(v) = \begin{cases} \max\{C_{n-1}(v), d_v^{\#} + C_0(h)\}, & \text{if } v \in V_{n-1}; \\ \max\{d(v, h) + C_{n-1}(\ell_i), & \\ \quad \max\{C_0(v), (d(v, h) + d_{\ell_i}^{\#} + C_0(h))\mathbb{I}_{\{k_{n-1}>1\}}\}\}, & \text{if } v \in V_0^{co_i}. \end{cases}$$

Proof. The graph G_n is obtained by attaching a copy of the seed G_0 at each of the latches $\ell_1, \ell_2, \cdots, \ell_{k_{n-1}}$ selected in the graph G_{n-1}.

We denote the vertex set of the rth copy of the seed, for $r = 1, \ldots, k_{n-1}$, by $V_0^{co_r}$. We now compute the distance from a node v to a vertex u in G_n by considering the four cases:

	$u \in V_{n-1}$	$u \in V_0^{co_j}$
$v \in V_{n-1}$	$d(v, u)$	$d(v, \ell_j) + d(h, u)$
$v \in V_0^{co_i}$	$d(v, h) + d(\ell_i, u)$	$d(u, v)\mathbb{I}_{i=j} + (d(v, h) + d(\ell_i, \ell_j) + d(h, u))\mathbb{I}_{i \neq j}$

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

For a given $v \in V_n$, the maximum (over the range of u and j) in each block is

	$u \in V_{n-1}$	u in a copy
$v \in V_{n-1}$	$C_{n-1}(v)$	$d_v^\# + C_0(h)$
$v \in V_0^{\mathrm{co}_i}$	$d(v,h) + C_{n-1}(\ell_i)$	$\max\left\{C_0(v), \left(d(v,h) + d_{\ell_i}^\# + C_0(h)\right)\mathbb{I}_{\{k_{n-1}>1\}}\right\}$

The result now follows. □

Remark 6.1. Suppose a vertex ℓ is chosen as a latch from G_{n-1}. From Theorem 6.1, we pick up the top line and write

$$\widetilde{C}_n(\ell) = \max\{C_{n-1}(\ell), d_\ell^\# + C_0(h)\}.$$

If $k_{n-1} = 1$, then $d_\ell^\# = 0$, in which case we have $\widetilde{C}_n(\ell) = \max\{C_{n-1}(\ell), C_0(h)\}$.

7. Diameter of the graph G_n

The diameter of a connected graph with vertex set $\mathcal{V}$ is the longest distance between any two nodes in it [4]. That is, the diameter is the maximum eccentricity, $\max_{v\in\mathcal{V}} C(v)$. For example, the diameters of the graphs G_0, G_1, G_2, and G_3 in Figure 1 are, respectively, 2, 4, 6, and 10.

We now introduce some additional notation:

1. $\widetilde{Д_n} = Д_n \mid G_{n-1}, \mathbb{L}_{n-1}$. This is the conditional diameter $Д_n$ of the graph G_n given G_{n-1} and the k_{n-1} latches in $\mathbb{L}_{n-1}$ (see Subsection 6.1 for the definition of $\mathbb{L}_{n-1}$).
2. Only for $k_n > 1$, we define $q_n = \max_{\ell,\tilde{\ell}\in\mathbb{L}_n} d(\ell, \tilde{\ell}) = \max_{\ell\in\mathbb{L}_n} \ell^\#$. Thus, q_n computes the maximum distance between any two latches in G_n.
3. $\alpha_n = \max_{\ell\in\mathbb{L}_n} C_n(\ell)$. Thus, α_n is the maximum eccentricity of a latch in G_n.

Theorem 7.1. *Suppose $\{k_n\}_{n=0}^\infty$ is a building sequence of the family of graphs $\{G_n\}_{n=1}^\infty$. The conditional diameter $\widetilde{Д_n} = Д_n \mid G_{n-1}, \mathbb{L}_{n-1}$ of a graph of age n is given by*

$$\widetilde{Д_n} = \max\{Д_{n-1}, (2C_0(h) + q_{n-1})\mathbb{I}_{\{k_{n-1}>1\}},\ C_0(h) + \alpha_{n-1}\}.$$

Proof. The (conditional) diameter $\widetilde{Д_n}$ of G_n may remain the same as the diameter $Д_{n-1}$ of G_{n-1},[2] unless we can find longer paths in G_n. The latter case arises, if

(a) There is a pair x and y of latches in G_{n-1}, and a pair of vertices (say u in the copy latched at x and v in the copy latched at y), such that $d(u,x)+d(x,y)+d(y,v) > Д_{n-1}$. The case can be, only if $k_{n-1} > 1$. The longest such distance is obtained by maximizing over x, y, u and v to obtain $2C_0(h) + q_{n-1}$.[3]

(b) Or, we can find a vertex u far enough from a latch ℓ in G_{n-1} and another vertex v in the copy latched at ℓ such that $d(u,\ell) + d(\ell, v) > Д_{n-1}$. The longest such distance is obtained by maximizing over ℓ, u and v to obtain $\alpha_{n-1} + C_0(h)$.

The longest distance in the graph is the maximum of the three possibilities discussed. □

Remark 7.1. Consider the case where, at stage n (for each $n \geq 1$), we pick among the k_{n-1} latches two, say $\ell, \tilde{\ell}$ in G_{n-1}, such that $d(\ell, \tilde{\ell})$ is the diameter of G_{n-1}. Note that this selection mechanism is no longer random in the sense discussed in all the preceding sections. Let us call the diameter of the graph so constructed $Д^*_{n-1}$. This is only possible if $k_{n-1} > 1$, for each n. By arguments similar to what we used in the proof of Theorem 7.1, we get $Д^*_n = 2C_0(h) + Д^*_{n-1}$. Unwinding we get $Д^*_n = 2nC_0(h) + Д_0$.

[2] In the graph G_2 in Figure 1, if we pick the three latches at distances 2,3,4 from the top vertex, the diameter of the graph so obtained in step 3 will be equal to $Д_2$, the diameter of G_2.

[3] This situation occurs in the graph G_3 in Figure 1.

https://doi.org/10.1017/S0269964822000523 Published online by Cambridge University Press

Remark 7.1 shows that, under this special hooking mechanism, the diameter $Д_n$ at step n only requires the knowledge of the seed graph and n. It does not take into consideration how many latches were picked at stages 1 through $n-1$ as long as there are two latches $\ell, \tilde{\ell}$ picked at each stage such that $d(\ell, \tilde{\ell})$ is maximum.

Acknowledgment. The authors would like to sincerely thank Dr. Shaimaa M. Abd-Elaal, Egyptian Ministry of Health and Population, for many fruitful discussions that gave them a perspective on applications of the model proposed.

Competing interest. The authors declare no conflict of interest.

References

[1] Bahrani, M. & Lumbroso, J. (2018). Split-decomposition trees with prime nodes: Enumeration and random generation of cactus graphs. *2018 Proceedings of the Fifteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, New Orleans, Louisiana, pp. 143–157.
[2] Bhutani, K., Kalpathy, R., & Mahmoud, H. (2021). Average measures in polymer graphs. *International Journal of Computer Mathematics: Computer Systems Theory* 6(1): 37–53. doi:10.1080/23799927.2020.1860134
[3] Brown, G. & Shubert, B. (1984). On random binary trees. *Mathematics of Operations Research* 9: 43–65.
[4] Chartrand, G. & Zhang, P (2012). *A first course in graph theory*. Mineola, NY: Dover Publications.
[5] Chen, C. & Mahmoud, H. (2016). Degrees in random self-similar bipolar networks. *Journal of Applied Probability* 53: 434–447.
[6] De La Briandais, R. (1959). File searching using variable length keys. *Proceedings of the Western Joint Computer Conference*. San Francisco, CA: AFIPS, pp. 295–298.
[7] Desmarais, C. & Holmgren, C. (2019). Degree distributions of generalized hooking networks. *2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, San Diego, California, pp. 103–110.
[8] Desmarais, C. & Holmgren, C. (2020). Normal limit laws for vertex degrees in randomly grown hooking networks and bipolar networks. *The Electronic Journal of Combinatorics* 27(2): P2.45.
[9] Desmarais, C. & Mahmoud, H. (2021). Depths in hooking networks. *Probability in the Engineering and Informational Sciences* 1–9. doi:10.1017/S0269964821000164
[10] Drmota, M (2009). *Random trees: An interplay between combinatorics and probability*. Vienna: Springer-Verlag.
[11] Drmota, M., Gittenberger, B., & Panholzer, A. (2008). The degree distribution of thickened trees. *DMTCS Proceedings, Fifth Colloquium on Mathematics and Computer Science*, vol. AI, pp. 149–162.
[12] Fredkin, E. (1960). Trie memory. *Communications of the ACM* 3: 490–499.
[13] Gopaladesikan, M., Mahmoud, H., & Ward, M. (2014). Building random trees from blocks. *Probability in the Engineering and Informational Sciences* 28: (1) 67–81.
[14] Knuth, D. (1998). *The art of computer programming, vol. 3: Sorting and searching*, 2nd ed. Reading, MA: Addison-Wesley.
[15] Mahmoud, H (1992). *Evolution of random search trees*. New York: John Wiley & Sons.
[16] Mahmoud, H. (2019). Local and global degree profiles of randomly grown self-similar hooking networks under uniform and preferential attachment. *Advances in Applied Mathematics* 111: 101930.
[17] Mahmoud, H. (2019). A spectrum of series-parallel graphs with multiple edge evolution. *Probability in the Engineering and Informational Sciences* 33(4): 487–499.
[18] Resnick, S. & Samorodnitsky, G. (2016). Asymptotic normality of degree counts in a preferential attachment model. *Advances in Applied Probability* 48(A): 283–299.
[19] Smythe, R. & Mahmoud, H. (1995). A survey of recursive trees. *Theory of Probability and Mathematical Statistics* 51: 1–27.
[20] Szpankowski, W (2001). *Average case analysis of algorithms on sequences*. New York: Wiley-Interscience.
[21] van der Hofstad, R (2016). *Random graphs and complex networks*, vol. 1. Cambridge: Cambridge University Press.

Cite this article: Bhutani KR, Kalpathy R and Mahmoud H (2024). Random multi-hooking networks. *Probability in the Engineering and Informational Sciences* **38**, 100–114. https://doi.org/10.1017/S0269964822000523