Hub-authority web-graphs.

Colin Cooper*

June 26, 2003

Abstract

We give precise results on the age dependent distribution of degree, and asymptotic degree sequence for the hub-authority model of web-graphs. Although these asymptotics exhibit a power law, the parameters of the power law vary according to the exact ratio of in-degree and out-degree. We also show how the age dependent degree distribution can be used to estimate the hub and authority rankings of the HITS algorithm on such web-graphs.

1 Introduction

At the present moment there is considerable research into the structure of large-scale real networks, and in modeling these networks as the outcomes of discrete random processes. A general introduction to this topic can be found in Hayes [11] or Watts [15]. In particular, there is a strong interest in the structure of the Internet and World Wide Web (**www**). Experimental studies by, Albert, Barabasi and Jeong [1],[2] Broder et al [6] and Faloutsos, Faloutsos and Faloutsos [10] of the structure of the **www** have demonstrated an inverse power law for the proportion of vertices with a given degree.

The simplest (undirected) models of such structures require a graph process which evolves randomly by the addition of new vertices and/or edges at each time step t and whose expected proportional degree sequence follows a power law. One method of producing graph processes with a power law degree sequence is to introduce an element of preferential attachment (or copying) into the way that a new vertex attaches its edges to the existing graph. Models of such processes variously called web-graphs or scale-free graphs have recently been studied by many authors. Rather than cite specific details we refer the reader to the survey by Bollobás and Riordan [4] gives an excellent overview of the topic and many detailed structural results.

Clearly the model of the **www** which represents pages as nodes and embedded links as directed edges is intrinsically directed. It is therefore more natural to model the **www** as a directed

*Department of Computer Science, King's College, University of London, London WC2R 2LS, UK. *e-mail* `ccooper@dcs.kcl.ac.uk`

web-graph process. We will refer to the digraph of the **www** as the **www**-Graph to distinguish it from the simple web-graph processes discussed below.

Directed web-graph models are less amenable to analysis than undirected models, and there has been less progress in this direction. In this paper, we consider the *hub-authority* model of directed web-graphs. Our reasons for considering the hub-authority model were based on the popularity of this approach to ranking pages by some search algorithms. A *hub* is a vertex with a large out-degree (points to many other pages) and an *authority* is a vertex with a large in-degree (pointed to by many pages). The HITS algorithm of Kleinberg [12], its modification in the Clever system [7] and the SALSA model [14] derive a hub and an authority weight for selected vertices of the **www**-Graph. Similarly, the page-rank heuristic of Google [5] is a measure of importance based on a weighted summary of the edges of the fan-in of the selected vertex. The hub-authority view of the **www**-Graph is extremely pervasive and has passed inter-alia into the domain of e-business and economic modeling. The Google query `'hub authority web graph'` returned about 5000 pages many of which are directly relevant.

The main thrust of this paper is as follows. We define a web-graph process which corresponds to the hub-authority view of the **www**-Graph. We derive the age specific degree distribution of vertices in this hub-authority model. Thus if v is a vertex born at step v and $\boldsymbol{d}(v,t) = (d^-(v,t), d^+(v,t))$ is the in- and out-degree of v at step t, we give $\mathbf{Pr}(\boldsymbol{d}(v,t) = (r,s))$. From this we derive the power law asymptotic $\mathbf{Pr}(\boldsymbol{d}(v,t) = (r,s)) \sim r^{-a}s^{-b}$. We find that the power law parameters a, b vary as a function of the values of r, s. These results are summarized in Theorem 1. Finally we show how to use the degree distribution to calculate the expected hub-authority weightings of the HITS algorithm on a web-graph.

For the hub-authority process, the results of Theorem 1 are predominantly new. However, some results correspond to the work of Bollobás, Borgs, Chayes and Riordan [3] on scale-free hub-authority processes. The discussion after Theorem 1 gives the exact correspondence between this paper and [3]. Limitations of space mean that we are unable to include every detail of the proof of Theorem 1 in this paper. We have included, as an appendix, the central parts of the proof ((b)(i), (c)). We refer the reader to [8] for (a), (b)(ii) and for the concentration part of (c) which follow naturally.

2 Definition of Hub-Authority web-graph process

There is an initial connected digraph $G(0)$ of finite size. At each step $t = 1, 2, ...$, an independent choice is made between an OLD and a NEW procedure, according to a Bernoulii random variable. Let α be the probability that the NEW procedure is followed at step t, and $1 - \alpha$ be the probability the OLD procedure is followed.

NEW **procedure.** If the NEW procedure is chosen at step t, a new vertex v_t is added with degree (m^-, m^+). Thus v_t has m^+ edges directed towards, and m^- edges directed away from $G(t-1)$. The number of edges (m^-, m^+) of each type, is sampled from a *finite* distribution $\boldsymbol{p}$. That is to say, $\boldsymbol{p} = (p_r^-, p_s^+)$, $r, s = 0, 1, ..., L$, where $L < \infty$ and $p_m \geq 0$ constant. Let $\overline{m}^- = \mathbf{E}\, m^-$ and $\overline{m}^+ = \mathbf{E}\, m^+$, and let $\overline{m} = \overline{m}^- + \overline{m}^+$. We do not require that the marginal distributions $\boldsymbol{p}^-, \boldsymbol{p}^+$ of $\boldsymbol{p}$ are independent. In fact an obvious extreme case is where they are mutually exclusive. If we insist that $(p_0^-, p_0^+) = (0,0)$, then the web-graph is connected but the results given here still hold even if this is not the case.

Each new edge $e_i, i = 1, ..., m^+$ directed out of v_t makes an independent decision to select its terminal vertex w in $G(t-1) = (V(t-1),\ E(t-1))$ by copying or uniformly with probabilities $A_1,\ A_2 = 1 - A_1$ respectively. Thus

$$p_A(w,t) = A_1 \frac{d^-(w,t-1)}{|E(t-1)|} + A_2 \frac{1}{|V(t-1)|}. \tag{1}$$

Similarly, each new edge $e_i, i = 1, ..., m^-$ directed into v_t independently selects an initial vertex w as initial vertex by copying or uniformly with probabilities $D_1,\ D_2 = 1 - D_1$ respectively. Thus

$$p_D(w,t) = D_1 \frac{d^+(w,t-1)}{|E(t-1)|} + D_2 \frac{1}{|V(t-1)|}. \tag{2}$$

OLD **procedure.** If the OLD procedure is followed then M random edges are inserted into $G(t-1)$ according to a finite $\boldsymbol{q} = (q_1, ..., q_M, ..., q_{L'})$. Let $\overline{M} = \mathbf{E}\, M$.

The choice of initial vertex (resp. terminal vertex) of each inserted edge is made independently from uar/copy rules similar to (1),(resp. (2)) with parameters B_1, B_2 (resp. C_1, C_2). Thus following (1) let

$$\begin{aligned} p_B(v,t) &= B_1 \frac{d^+(v,t-1)}{|E(t-1)|} + B_2 \frac{1}{|V(t-1)|} && (3)\\ p_C(v,t) &= C_1 \frac{d^-(v,t-1)}{|E(t-1)|} + C_2 \frac{1}{|V(t-1)|}. && (4) \end{aligned}$$

It will be shown that the directed degree distribution of the Hub-Authority Process depends on two parameters $\eta,\ \nu$, given by

$$\begin{aligned} \eta^- &= \frac{\alpha \overline{m}^- A_1 + (1-\alpha)\overline{M} C_1}{\alpha \overline{m} + (1-\alpha)\overline{M}} & \nu^- &= \frac{\alpha \overline{m}^- A_2 + (1-\alpha)\overline{M} C_2}{\alpha} && (5)\\ \eta^+ &= \frac{\alpha \overline{m}^+ D_1 + (1-\alpha)\overline{M} B_1}{\alpha \overline{m} + (1-\alpha)\overline{M}} & \nu^+ &= \frac{\alpha \overline{m}^+ D_2 + (1-\alpha)\overline{M} B_2}{\alpha}. && (6) \end{aligned}$$

3 The Hub-Authority Process: Statement of results

Assertions made in this paper are in probability. Usually we qualify this explicitly with the abbreviation **whp** (with probability tending to 1 as $t \to \infty$).

It is convenient to classify vertices according their time of birth relative to the current step t. We say an *early* vertex, is a vertex born before $v_0(t) = o(t)$, where the precise value of $v_0(t)$ is given in (7) below. A *late* vertex, is one born at or after time v_0. Let

$$v_0(t) = \lceil \left(t^{\frac{4\eta}{4\eta+1}} \right) \log^5 t \rceil \tag{7}$$

For presentation of asymptotic results we use $a(t) \sim b(t)$ to mean $\lim_{t\to\infty} a(t)/b(t) = 1$. In a different context, we also use $X \sim D$ to denote a random variable X with distribution D. We adopt the usual definition of the generalized binomial coefficient, namely for real $a, b \geq 0$,

$$\binom{a+b}{a} = \frac{\Gamma(a+b+1)}{\Gamma(a+1)\,\Gamma(b+1)}, \tag{8}$$

and $\Gamma(x) = (x-1)\Gamma(x-1)$.

We take an indirect approach in which we derive the distribution of degree $d(v,t)$ of a fixed vertex v over all processes which include this vertex. We then use concentration of degree sequence to infer the structure of a typical process.

Let $\xi(m) = \xi^-(m^-), \xi^+(m^+)$ be given by

$$\xi = m + \nu/\eta, \tag{9}$$

where η, ν are given by (5), (6).

To save on notation, we let $\boldsymbol{d}(v,t) = (d^-(v,t), d^+(v,t)) = (r,s)$ denote the in- and out-degrees of vertex v at step t, and assume η, ν have the appropriate superscript.

Theorem 1. *Let* $r = o(t^{\eta^-/(2(1+4\eta^-))})$, $s = o(t^{\eta^+/(2(1+4\eta^+))})$,

(a) *Degree of late vertices: The distribution of in- and out-degree of any vertex are effectively un-correlated.*
Let $\boldsymbol{d}(v,v) = (m^-, m^+)$ *be the initial degree of vertex* v, *then*

$$\begin{aligned}&\mathbf{Pr}(\boldsymbol{d}(v,t) = (m^- + r, m^+ + s) \mid \boldsymbol{d}(v,v) = (m^-, m^+)) = (1+o(1))\times\\ &\quad \mathbf{Pr}(d^-(v,t) = m^- + r \mid \boldsymbol{d}(v,v) = (m^-, m^+))\,\mathbf{Pr}(d^+(v,t) = m^+ + s \mid \boldsymbol{d}(v,v) = (m^-, m^+)).\end{aligned}$$

(b) *Expected value of in-degree, distribution of in-degree and proportional in-degree sequence of late vertices (resp. out-degree):*

(i) *Negative binomial degree distribution: For* $v \geq v_0$ *and* $r \leq K \log t (t/v)^\eta$

$$\begin{aligned}&\mathbf{Pr}(d^-(v,t) = m^- + r \mid \boldsymbol{d}(v,v) = (m^-, m^+))\\ &\qquad = (1+o(1))\binom{r+\xi^- - 1}{r}\left(\frac{v}{t}\right)^{\eta^-\xi^-}\left(1 - \left(\frac{v}{t}\right)^{\eta^-}\right)^r.\end{aligned}$$

A tail inequality also applies to r *(see [8]).*

(ii) *Let* $N^-{}_t(l \mid m)$ *denote the proportion of vertices of* $G(t)$ *of in-degree* $l+m$ *with initial in-degree* m. *Then* **whp** $N^-{}_t(l \mid m) \sim \widehat{n}^-(l \mid m)$, *where*

$$\widehat{n}^-(l \mid m) = \frac{1}{\eta^-}\frac{\Gamma(\xi^- + 1/\eta^-)}{\Gamma(\xi^-)}\frac{\Gamma(l+\xi^-)}{\Gamma(l+1+\xi^-+1/\eta^-)}. \tag{10}$$

In particular, as $l \to \infty$,

$$\widehat{n}^-(l \mid m) \sim \frac{1}{\eta^-}\frac{\Gamma(\xi^-+1/\eta^-)}{\Gamma(\xi^-)}\, l^{-\left(1+\frac{1}{\eta^-}\right)}. \tag{11}$$

(c) *Joint proportional degree sequence:*
Let $N_t(r, s \mid m^-, m^+)$ be the proportion of vertices at step t with in-degree r, out-degree s, and initial degree (m^-, m^+). Then, **whp** *$N_t(r, s \mid m^-, m^+) \sim n(r, s \mid m^-, m^+)$ where:*

$$n(r, s \mid m^-, m^+) = \frac{\alpha 1(m^->0)+\gamma 1(m^+>0)}{\alpha+\gamma} \binom{r+\xi^- -1}{r} \binom{s+\xi^+ -1}{s} \frac{1}{\eta^-} \int_0^1 x^a (1-x)^r (1-x^b)^s dx, \tag{12}$$

where $b = \eta^+/\eta^-$ and $a = \eta^+/\eta^- \ \xi^+ + 1/\eta^- + \xi^- - 1$. Let $C_{r,s} = \Theta(1)$, constant. If $r \to \infty$ and $s = O(r^b)$ then

$$n(r, s \mid m^-, m^+) = C_{r,s} r^{-\left(1+\frac{\eta^+}{\eta^-}\xi^+ + \frac{1}{\eta^-}\right)} s^{-(1-\xi^+)}. \tag{13}$$

If $s \to \infty$ and $r = O(s^{1/b})$ then

$$n(r, s \mid m^-, m^+) = C_{r,s} r^{-(1-\xi^-)} s^{-\left(1+\frac{\eta^-}{\eta^+}\xi^- + \frac{1}{\eta^+}\right)}. \tag{14}$$

If $s = \Theta(r^b)$ the expressions (13), (14) are equivalent.

Clearly Theorem 1(c) is the most interesting result. When $\eta^- = \eta^+ = \eta$ we can derive more precise results, namely

$$n(r, s \mid m^-, m^+) \sim C \frac{r^{\xi^- -1} s^{\xi^+ -1}}{(r+s)^{\xi^- + \xi^+ + \frac{1}{\eta}}}.$$

This reduces to (13), (14) as required.

We now discuss the relationship between Theorem 1 and the results of [3]. Their model although formulated somewhat differently (as a directed scale-free graph), is equivalent to a restricted version of the Hub-Authority Process. The results they obtain are the marginal power laws for in- and out-degree (see (11) above) and the joint distribution (see (13), (14)) in the case where s is constant and $r \to \infty$ (or vice versa).

The model of [3] uses a different parameterization. The parameters $\alpha, \beta, \gamma, \delta$ of their model are not the same as ours. Our parameter α is $\alpha + \gamma$ of [3]. A single edge is inserted at each step so, if NEW-IN occurs then $\mathbf{Pr}(m^- = 1) = 1$ always, and if NEW-OUT occurs then $\mathbf{Pr}(m^+ = 1) = 1$. These events are mutually exclusive. Thus $\mathbf{Pr}((0,1)) = \alpha/(\alpha+\gamma)), \mathbf{Pr}((1,0)) = \gamma/(\alpha+\gamma))$. The parameter $\beta = 1 - (\alpha+\gamma)$ is the probability of inserting an edge between old vertices and δ is a uar parameter.

The translation to our model is:

$$A_1 = \frac{1}{1+\delta_{in}(\alpha+\gamma)}, \qquad D_1 = \frac{1}{1+\delta_{out}(\alpha+\gamma)}, \qquad C_1 = A_1, \qquad B_1 = D_1.$$

Thus calculating η, ν for their model we find $\eta^- = c_1$, $\eta^+ = c_2$ and $\xi^+(0) = \nu^+/\eta^+ = \delta_{out}$, where c_1, c_2 define the power law parameters for that paper.

4 Applications to the HITS algorithm

Before proceeding, we briefly review the HITS algorithm of Kleinberg [12] for ranking queries. Let A be the adjacency matrix of (a selected subgraph of) the **www**-Graph. Thus $A_{i,j} = 1$ if page i points to page j. The HITS algorithm iteratively calculates the vectors $\boldsymbol{x}, \boldsymbol{y}$. Here $\boldsymbol{x} = (x_j)$ and x_j gives the authority weight of vertex j, similarly $\boldsymbol{y} = (y_i)$ and y_i gives the hub weight of vertex i. The weight is used to rank the vertices. The vectors $\boldsymbol{x}, \boldsymbol{y}$ can be calculated as $\boldsymbol{x} = A^T A\boldsymbol{x}$ and $\boldsymbol{y} = AA^T\boldsymbol{y}$, thus if AA^T is known we can calculate $\boldsymbol{y}$ etc.

We now explain how to use the age specific degree distribution obtained in this paper to calculate the expected value of the entries of the matrix AA^T. This can be used to obtain the expected authority and hub vectors $\boldsymbol{x}, \boldsymbol{y}$ of the web-graph process. This information could be compared with experimental data from the **www**-Graph to see if it exhibits properties similar to a hub-authority web-graph process.

Let $B = \mathbf{E}\, AA^T$, and let $B_{i,k}$ be the (i,k)-th entry of B. Thus

$$B_{i,k} = \mathbf{E} \sum_j 1(i \to j)1(k \to j),$$

where $1(i \to j)$ is the indicator for the event that vertex i points to vertex j.

As an example we consider a very simple model of hub-authority web-graphs with a new vertex added at each step with in-degree and out-degree m (ie: $m^+ = m^- = m$) then $B_{i,k} \sim$. This means that $\nu = 0$ and $\eta^- = \eta^+ = 1/2$ and $2m$ edges are added at each step. This models is the hub-authority equivalent of the original model proposed by [2].

Let $i < k$ and consider $B_{i,k} = \mathbf{E} \sum_j 1(i \to j)1(k \to j)$. There are three cases to consider: $j \le i < k$, $i \le j \le k$ and $i < k \le j$. The second two cases are essentially uncorrelated and contribute $O(1)/\sqrt{ik}$ to the total. We turn our attention to the first case. Let $d^-(j,i) = r$, $d^-(j,k) = s$ be the in-degree of vertex j at steps i, k respectively. Thus

$$\mathbf{E}\, 1(i \to j)1(k \to j) = \frac{r}{2mi}\mathbf{Pr}(d^-(j,i) = r)\frac{s}{2mk}\mathbf{Pr}(d^-(j,k) = s \mid d^-(j,i) = r),$$

where

$$\begin{aligned} \mathbf{Pr}(d^-(j,i) = r) &\sim \binom{r+m-1}{r}\left(\frac{j}{i}\right)^{\frac{m}{2}}\left(1-\left(\frac{j}{i}\right)^{\frac{1}{2}}\right)^r \\ \mathbf{Pr}(d^-(j,k) = s \mid d^-(j,i) = r) &\sim \binom{r+s-1}{s}\left(\frac{k}{i}\right)^{\frac{r}{2}}\left(1-\left(\frac{k}{i}\right)^{\frac{1}{2}}\right)^s. \end{aligned}$$

The first distribution is from Theorem 1 (b). Conditional on $d^-(j,i) = r$ the growth of the in-degree of vertex j restarts at $i+1$ with initial value $m^- = r$, giving the second distribution.

Firstly note that

$$\sum_{j \le i} \mathbf{Pr}(d^-(j,i) = r) \sim 2i\frac{(m+1)!r!}{(m+r+2)!},$$

and next that

$$\sum_s \mathbf{Pr}(d^-(j,k) = s \mid d^-(j,i) = r) \sim r \left(\frac{k}{i}\right)^{\frac{1}{2}}.$$

Thus

$$\begin{aligned} B_{i,j} &= \sum_{j \le i} \mathbf{E}\, 1(i \to j) 1(k \to j) \\ &\sim \frac{1}{\sqrt{ik}} \frac{m+1}{2m} \sum_r \frac{r^2}{(m+r+2)(m+r+1)(m+r)}. \end{aligned}$$

The upper limit on r at step i is $O(i^{1/2})$ (see eg: [8],[9]). Potential errors creep in if we estimate r above $\Theta(i^{1/6})$ using the negative binomial distribution, however the approximation is accurate below $\Theta(i^{1/6})$ and of the correct order up to $O(i^{1/2})$. Thus the summation is of the form $c \log i$ where $1/6 \le c \le 1/2$ and there exists a positive constant $a = c(m+1)/(2m)$ such that

$$B_{i,j} = a \frac{\log i}{\sqrt{ik}}.$$

The eigen-vector $\boldsymbol{y}$ can be obtained by standard techniques.

References

[1] R. Albert, A. Barabasi and H. Jeong, *Diameter of the world wide web.* Nature 401, 103-131, (1999).

[2] A. Barabási and R. Albert, *Emergence of scaling in random networks*, Science 286, 509-512, (1999).

[3] B. Bollobás, C. Borgs, J. Chayes and O. Riordan, *Directed scale-free graphs.* SODA (2003).

[4] B. Bollobás, O. Riordan, Mathematical results on scale-free graphs, *Handbook of Graphs and Networks*, S. Bornholdt, H. Schuster (eds), Wiley-VCH, Berlin (2002).

[5] S. Brin and L. Page, *The anatomy of a large scale hypertextual web search engine*, Computer Networks 30, 107-117 (1998).

[6] A. Broder, R. Kumar, F.Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, *Graph structure in the web,* Proc. 9th WWW Conf. 309-320 (2000).

[7] Members of the Clever Project. *Hypersearching the web.* Scientific American, June 1999.

[8] C. Cooper, *The degree sequence of directed web-graphs.* Submitted to Combinatorics, Probability and Computing (2002).

[9] C. Cooper and A. Frieze, *A general model of web graphs,* Proceedings of ESA 2001.

[10] M. Faloutsos, P. Faloutsos and C. Faloutsos, *On power-law relationships of the internet topology*, SIGCOMM 1999, Comput. Commun. Rev. 29, 251, (1999).

[11] B. Hayes, *Graph theory in practice: Part II.* American Scientist 88, 104-109. (2000). `http://www.sigmaxi.org/amsci/issues/Comsci00/compsci2000-03.html`

[12] J. Kleinberg. *Authoritative sources in a hyperlinked environment.* Proc. 9th ACM-SIAM Symposium on Discrete Algorithms, (1998), Journal of the ACM 46 (1999).

[13] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal. *The web as a graph.* Proceedings of the 19th ACM Symposium on Principles of Database Systems, (2000), 1-10.

[14] R. Lempel and S. Moran. *The stochastic approach for link-structure analysis (SALSA).* ACM Transactions on Information Systems 19(2) 131-160 (2001).

[15] D. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton University Press (1999).

5 Appendix: Proof of Theorem 1

Let $|E(t)|$ (resp. $|V(t)|$) be the total number of edges (resp. vertices) of $G(t)$. Thus

$$\begin{aligned}\mathbf{E}\ |E(t)| =& t(\alpha\overline{m} + (1-\alpha)\overline{M}) + |E(0)|,\\ \mathbf{E}\ |V(t)| =& t\alpha + |V(0)|.\end{aligned}$$

The values of $|E(t)|$, $|V(t)|$ are the sum of independent bounded random variables and are sharply concentrated in t. The inequality

$$h(t) = \mathbf{E}\ h(t)\left(1{+}O\left(\sqrt{\tfrac{K(1+\log t)}{t}}\right)\right) \qquad h(t) = |E(t)|,\ |V(t)| \qquad t \geq 1 \tag{15}$$

holds simultaneously with probability $1 - {t_0}^{-K'}$ for all $t \geq t_0$ and $t_0 > 0$. Let $\boldsymbol{C}(t)$ denote the set of processes $G(t)$ for which (15) holds. We refer to such processes as *good*.

We derive the marginal distribution of in- or out-degree, conditional on $\boldsymbol{d}(v,v) = (m^-, m^+)$. It follows from minor alterations to the proof (see [8]) that joint distribution $\boldsymbol{d}(v,t)$ is asymptotic to the product of the marginals, even in the extreme case where the existence of edges directed into or out of a new vertex are mutually exclusive events. For the calculations of the marginal degree distributions, we replace $d^-(v,t)$ by $d(v,t)$, and supress superscripting on the parameters etc.

5.1 1 (b)(i): Time dependent marginal degree distribution

Theorem 2. *Let η, ν, ξ be given by (5), (6), (9), and let v_0 be given by (7). Let $l = o((t/\log t)^{1/4})$. Let*

$$\pi_{l,m}(v,t) = \binom{l+\xi-1}{l}\left(\frac{v}{t}\right)^{\eta\xi}\left(1-\left(\frac{v}{t}\right)^{\eta}\right)^{l}. \tag{16}$$

Let $\delta^*(l,v)$ be given by

$$\delta^*(l,v) = O\left((m+l)\sqrt{\frac{\log v}{v}} + \frac{(m+l)^2}{v}\right). \tag{17}$$

Let $v \geq v_0$. Suppose that, at step v a new vertex of initial degree m is added, then

$$\mathbf{Pr}(d(v,t) = \; m+l \mid d(v,v)=m) = \pi_{l,m}(v,t)\,(1+l\delta^*(l,v))\,. \tag{18}$$

Proof We first give an outline description of the proof, and then fill in the details.

Let $\mu_t = (1_\alpha(t), m(t), M(t))$ give the choice of procedure (NEW or OLD), and number of edges inserted at step t, and let $\boldsymbol{\mu}(t) = (\mu_1, ..., \mu_t)$ be the history of these choices. Let $\boldsymbol{G}(\boldsymbol{\mu}(t)) = \{G(t) : \boldsymbol{\mu}(t)\}$. All graphs in $\boldsymbol{G}(\boldsymbol{\mu}(t))$ have the same number of edges $|E(T)|$ and vertices $|V(T)|$, for all $T \leq t$.

Let $\boldsymbol{\mu}(t)$ be such that $\mu_v = (1, m, 0)$. Thus a new vertex labeled v, of initial degree m is added at step v. Let $d(v,t)$ denote the degree of the vertex v at the end of step t.

Suppose $d(v,t) = m+l$ and that $\boldsymbol{\tau} = (\tau_1, ..., \tau_l)$ gives the steps τ_j at which $d(v, \tau_j)$ changes from $m+j-1$ to $m+j$. If more than one edge hits v during step τ_j the value τ_j is repeated the appropriate number of times in $\boldsymbol{\tau}$. We define $\tau_0 = v$ and $\tau_{l+1} = t$.

The set $\boldsymbol{C}(t)$ of good sequences is defined in (15) above. Let $I(t)$ be given by

$$I(t) = \{\boldsymbol{\mu}(t) \in \boldsymbol{C}(t) : 1_\alpha(v) = 1,\ d(v,v) = m\}.$$

By working within $I(t)$ we can replace $|E(T)|$, $|V(T)|$ in (1-4) by their approximate expected values to remove conditioning on previous history of these variables at each step $T = v, ..., t$. The approximation error we incur at each step is $(1 + O(\sqrt{\log T/T}))$.

The probability $\mathbf{Pr}(d(v,t) = m+l,\ \boldsymbol{\tau} \mid d(v,v) = m)$ is obtained as follows:

$$\mathbf{Pr}(d(v,t) = m+l,\ \boldsymbol{\tau} \mid d(v,v) = m) = \prod_{j=1,...,l+1} \mathbf{Pr}(\tau_j \mid \tau_{j-1},\ d(v,v) = m).$$

Here,$\mathbf{Pr}(\tau_j \mid \tau_{j-1})$ means $d(v,T) = m+j-1$ for $\tau_{j-1} \leq T < \tau_j$ and $d(v,\tau_j) = m+j$. Also, $\mathbf{Pr}(\tau_{l+1} \mid \tau_l)$ is taken to mean that the degree of v remains unaltered at $m+l$ from step τ_l up to and including $\tau_{l+1} = t$. Finally we sum over $\boldsymbol{\tau}$ to get $\mathbf{Pr}(d(v,t) = m+l \mid d(v,v) = m)$. We now give the detailed analysis.

Let $X(t)$ count the number of edges selecting vertex v at step t. The distribution of $X(t)$ is given by

$$X(t) \sim (1_\alpha(t)\mathrm{Bin}(m(t), p_A(t))) + (1_\beta(t)[\mathrm{Bin}(M(t), p_B(t)) + \mathrm{Bin}(M(t), p_C(t))])\,,$$

where the parameters are described in Section 2, and $\mathrm{Bin}(m,p)$ is the Binomial distribution. Conditional on $d(v,t) = m+j$ and $I(t)$, and recalling the definitions (1-4), (5) and (6) we have

$$\mathbf{E}\ X(t+1 \mid d(v,t) = m+j, I(t)) = \frac{\eta(m+j)+\nu}{t}\left(1+O\left(\sqrt{\tfrac{\log t}{t}}\right)\right).$$

Let

$$\epsilon(t,j) = O(\sqrt{\log t/t} + (m+j)/t), \tag{19}$$

then

$$\begin{aligned}
\mathbf{Pr}(X(t+1) =& 0 \mid d(v,t) = m+j, I(t)) && (20)\\
=&\mathbf{E}_{\ \mu_{t+1}}\left(1_\alpha(1-p_A)^{m(t+1)} + 1_\beta[(1-p_B)^{M(t+1)}(1-p_C)^{M(t+1)}]\right) && (21)\\
=&1 - \left(\frac{\eta(m+j)+\nu}{t}\,(1+\epsilon(t,j))\right). && (22)
\end{aligned}$$

If $\tau_{j+1} = \tau_j$ or $\tau_j + 1$ let $\Psi_j = 1$. If $\tau_{j+1} \geq \tau_j + 2$, let $\Psi_j = \mathbf{Pr}(X(T) = 0, \tau_j < T < \tau_{j+1})$, thus

$$\Psi_j = \prod_{\tau_j < T < \tau_{j+1}} 1 - \left(\frac{\eta(m+j)+\nu}{T}\,(1+\epsilon(T,j))\right). \tag{23}$$

Let $\delta(t,j) = O((m+j)\sqrt{\log t/t} + (m+j)^2/t)$, then

$$\begin{aligned}
\Psi_j &= \exp\left(\left(-(\eta(m+j)+\nu)\sum_{\tau_j<T<\tau_{j+1}}\frac{1}{T}\right) + \delta(\tau_j,j)\right)\\
&= \left(\frac{\tau_j}{\tau_{j+1}}\right)^{\eta(m+j)+\nu}(1+\delta(\tau_j,j)).
\end{aligned}$$

Let $\Phi_l = 1$ and for $0 \leq j \leq l-1$, let $\Phi_j = \mathbf{Pr}(X(t+1) = 1 \mid d(v,t) = m+j, I(t))$. Thus

$$\Phi_j = \frac{\eta(m+j)+\nu}{t}\,(1+\epsilon(t,j)).$$

Also, $\mathbf{Pr}(X(t+1) \geq 2) = O(\Phi_j\ (m+j)/t)$. Thus, when the entries in $\boldsymbol{\tau}$ are not distinct, the correction arising from these events can be absorbed into $\delta(\tau_j,j)$.

Let $F(\boldsymbol{\tau})$ denote $\mathbf{Pr}(d(v,t) = m+l$ and $\boldsymbol{\tau} \mid d(v,v) = m, I(t))$ then

$$F(\boldsymbol{\tau}) = \prod_{j=1,\ldots,l+1}\mathbf{Pr}(\tau_j \mid \tau_{j-1},\ d(v,v) = m, I(\tau_j)) = \prod_{j=0}^{l}(\Psi_j\ \Phi_j).$$

Ignoring for the moment the multiplicative error terms, we see that

$$F(\boldsymbol{\tau}) \sim \left(\frac{v}{\tau_1}\right)^{\eta m+\nu}\frac{\eta m+\nu}{\tau_1}\left(\frac{\tau_1}{\tau_2}\right)^{\eta(m+1)+\nu}\frac{\eta(m+1)+\nu}{\tau_2}\cdots\left(\frac{\tau_{l-1}}{\tau_l}\right)^{\eta(m+l-1)+\nu}\frac{\eta(m+l-1)+\nu}{\tau_l}\left(\frac{\tau_l}{t}\right)^{\eta(m+l)+\nu}.$$

Cancelling repeated values of τ_j from the numerator and denominator of $F(\boldsymbol{\tau})$ gives

$$F(\boldsymbol{\tau}) = \frac{\Gamma(l+\xi)}{\Gamma(\xi)}\left(\frac{v}{t}\right)^{\eta\xi}\prod_{j=1}^{l}\frac{\eta\ \tau_j^{\eta-1}}{t^\eta}(1+\delta(\tau_j,j)). \tag{24}$$

Thus

$$\mathbf{Pr}(d(v,t)=m+l \mid d(v,v)=m) = \left(1+O\left(\tfrac{1}{t^K}\right)\right)\sum_{\boldsymbol{\tau}} F(\boldsymbol{\tau}) + O\left(\tfrac{1}{t^K}\right).$$

Now

$$\begin{aligned}\sum_{\boldsymbol{\tau}}\prod_{j=1}^{l}\frac{\eta\tau_j^{\eta-1}}{t^\eta}(1+\delta(\tau_j,j)) &= \frac{1}{l!}\left(\sum_{v<\tau\le t}\frac{\eta\tau^{\eta-1}}{t^\eta}(1+\delta(\tau,l))\right)^l \\ &= \frac{1}{l!}\left(1-\left(\frac{v}{t}\right)^\eta(1+\delta^*)\right)^l,\end{aligned}$$

where

$$\delta^*(l,v) = O\left((m+l)\sqrt{\frac{\log v}{v}} + \frac{(m+l)^2}{v} + (m+l)\sqrt{\frac{\log t}{t}}\left(\frac{t}{v}\right)^\eta + \frac{(m+l)^2}{t}\left(\frac{t}{v}\right)^\eta\right). \quad (25)$$

Thus, somewhat crudely,

$$\mathbf{Pr}(d(v,t)=m+l \mid d(v,v)=m) = \pi_{l,m}(v,t)(1+l\delta^*).$$

This gives the value of (18) and thus completes the proof of the theorem.

5.2 1(c): Asymptotic degree distribution

Let $N_t(r,s) = N_t(r,s \mid m^-,m^+)$ denote the number of vertices of $G(t)$ with in-degree r, out-degree s, and initial in-degree m^- and out-degree m^+. Thus

$$\mathbf{E}\, N_t(r,s \mid m^-,m^+) = \sum_{v=1}^{t}\mathbf{Pr}(d^-(v,t)=m^-+r, d^+(v,t)=m^++s \mid m^-,m^+).$$

By calculations similar to Theorem 2(b) the proportion $N_t(r,s)$ of such vertices is given **whp** by $(1+o(1))\mathbf{E}\, N_t(r,s)/(\alpha t)$. By 1(a), (b)(i), we have

$$\begin{aligned}\mathbf{E}\, D_t(r,s) &= O(v_0) + (1+o(1))\sum_{v=v_0}^{t}\mathbf{Pr}(d^-(v,t)=m^-+r \mid m^-)\mathbf{Pr}(d^+(v,t)=m^++s \mid m^+) \\ &= Ctr^{-(1-\xi^-)}s^{-(1-\xi^+)}I(r,s),\end{aligned}$$

where $I(r,s)$ is given by

$$I(r,s) = \int_0^1 x^a(1-x)^r(1-x^b)^s dx,$$

where $b=\eta^+/\eta^-$ and $a=\eta^+/\eta^-\ \xi^+ + 1/\eta^- + \xi^- - 1$.

Assuming at least one of $r,s\to\infty$ we consider three cases: $s=o(r^b)$, $s=\Theta(r^b)$ and $r=o(s^{1/b})$.

Case $s=o(r^b)$.

Let $B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx$ denote the standard Beta integral. Thus $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$. The main measure of $\int_0^1 x^a(1-x)^r dx$ comes from the interval $[0, \omega/r]$ where $\omega \to \infty$ arbitrarily slowly.

Let $s = (r/\omega^2)^b$, thus when $x = \omega/r$, $sx^b = 1/\omega^b$ and so $(1-x^b)^s = (1-o(1))$. Thus

$$I(r,s) = (1-o(1))B(a+1, r+1) = Cr^{-(a+1)}.$$

Case $r = o(s^{1/b})$**.** This is similar to the above case. Making the transformation $y = x^b$ we find

$$I(r,s) = \frac{1}{b}(1-o(1))B((a+1)/b, s+1) = Cs^{-\left(\frac{a+1}{b}\right)}.$$

Case $s = \Theta(r^b)$**.** Clearly $(1-x^b) \le 1$ so

$$I(r,s) \le B(a+1, r+1) = Cr^{-(a+1)}.$$

To obtain a lower bound we argue as follows: As $s = \Theta(r^b)$, we can write $s = fr^b$ where f is constant. When $x_0 = c/r$, where $f = c^{1-b}$, then x_0 satisfies $rx = sx^b$. Assuming that $sx^b \le rx$ for $x \le x_0$, and using $e^{-y/(1-y)} \le 1-y$, we have

$$\begin{aligned} I(r,s) &\ge \int_0^{x_0} x^a e^{-3rx} dx + \int_{x_0}^{\omega} x^a e^{-3sx^b} dx \\ &= \frac{1}{(3r)^{a+1}} \int_0^{3c} y^a e^{-y} dy + \frac{1}{b(3s)^{(a+1)/b}} \int_{3c}^{\omega'} y^{\frac{a+1}{b}-1} e^{-y} dy. \end{aligned}$$

Thus

$$I(r,s) \ge C_1 r^{-(a+1)} + C_2 s^{-(a+1)/b}.$$

However, as $r = s^{1/b}c^{1-1/b}$ this is equivalent to $I(r,s) \ge Cr^{-(a+1)}$.