# DISCOVERY OF NODAL ATTRIBUTES THROUGH A RANK-BASED MODEL OF NETWORK STRUCTURE

ADAM DOUGLAS HENRY AND PAWEŁ PRAŁAT

Abstract. The structure of many real-world networks co-evolves with the attributes of individual network nodes. Thus, in empirical settings it is often necessary to observe link structures as well as nodal attributes; however, it is sometimes the case that link structures are readily observed whereas nodal attributes are very difficult to measure. This paper investigates whether it is possible to assume a model of how networks co-evolve with nodal attributes, and then apply this model to infer unobserved nodal attributes based on a known network structure. We find that it possible to do so in the context of a previously-studied "rank" model of network structure, where nodal attributes are represented by externally-determined ranks. In particular, we show that node ranks may be reliably estimated by examining node degree in conjuction with the average degree of first- and higher-order neighbors.

## 1. Introduction

Research on complex networks has allowed us to better describe and model the interconnections amongst agents embedded within social, natural, and physical systems [8, 16, 21]. Despite advances in the rigorous modeling of self-organizing networks (see, for example [3] and [4]), we are still in need of better models of how nodal attributes both influence and coevolve with network structure. Such models are crucial given the well-known role of nodal attributes in driving link structure. In the Web graph, for instance, it has been posited that new linkages are formed with probability proportional to node degree [2], however it is also likely that link formation is determined by the content of nodes—for example, through homophily processes linkages may form with higher probability if the lexical content of two webpages are very similar to one another (one possible application of a "spatial" preferential attachment model; see [1, 10]). Models that explicitly account for the interplay between link structures and nodal attributes are therefore relevant to research on the Web graph, but also apply more broadly to the study of complex social networks [13].

In this paper, we investigate how models of complex networks may be used to reliably guess nodal attributes based on observations of network structure alone. Given the importance of nodal attributes in driving the dynamics of many real-world systems, it is important to understand how we may estimate such variables when empirical observations are difficult or impossible. For example, in some social networks (such as

1991 *Mathematics Subject Classification.* Primary: 05C80. Secondary: 05C07.

*Key words and phrases.* random graphs, web graphs, protean graphs, social networks, degree distribution, differential equations method, power law graphs, scale-free networks.

Paweł Prałat gratefully acknowledges support from MPrime and Ryerson University.

online social networks) it may be quite to observe interactions, but the actual prestige, beliefs, or emotions of individuals within the network (which may partly determine or be determined by structure) are not easily measured. This research illustrates how one may a mathematical model of the coevolution of network structure and nodal attributes, and then "reverse" this model develop reliable estimates of individual nodal attributes based based on knowledge of the link structure only.

We illustrate this approach by using an existing model of network self-organization propsed in [11], which in turn is a generalization of the model proposed in [14]. This model fits well into the larger enterprise of modeling the co-evolution of networks and nodal attributes, illustrated for example by [7, 11, 12, 14, 18, 19, 20]. In the model used here, linkages are chosen based in part on nodes' externally determined rank relative to other nodes. Usage of this ranking scheme allows us to model attributes that change stochastically as nodes enter and exit the system over time. In doing so, we build upon prior work on ranking models [5, 11, 14], and suggest some new directions for future research and applications of these models.

Our paper builds upon earlier research presented at the 2010 WAW conference [6]. As with the earlier paper, our central purpose is to illustrate the feasibility of reversing mathematical models of network self-organization to estimate nodal attributes. We investigate this core question in the context of ranking model by showing how one can reliably infer the externally-determined rank of nodes based on observations of node degree, and the average degree of first-order neighbors, the average degree of second-order neighbors, and so on. While we originally illustrated the possibility of inferring ranks based on observed degree through computational simulation only, this paper demonstrates how one may rigorously infer node attributes based on structural properties alone. We turn first to a discussion of the ranking model. We then consider the problem of inferring ranks based on observations of node degree, and suggest some areas for future research.

## 2. Model

In this section, we formally define a ranking model that reflects the attributes of nodes within a hypothetical system. This model, called the Protean graph, was introduced by Łuczak and Prałat in [14] and then generalized to other ranking schemes by Janssen and Prałat in [11]. This model not only specifies the process by which attributes are assigned to individual nodes, but also specifies the way in which these attributes shift over time as nodes enter or exit the system. The model focuses on systems where the total number of nodes is large but fixed (for example, if at each time step an node is removed uniformly at random and immediately replaced by a new one). This type of behavior is most consistent for well-established systems. Such stochastic systems are also usually more challenging to model than, say, systems that are "young" or "middle-aged" and hence growing over time, with nodes being added to the system at a faster rate than they are removed.

The only modification in the model we study in this paper from the model studied in [11] is that an initial rank is biased towards average values. However, the proofs in [11] may be easily adjusted to the new ranking function used here. We omit proofs

in this paper, stating the results only. The second minor difference is that in [11] the model was defined as a graph process, whereas in this paper the model is split into two parts. In the first part, we focus on the behavior of the rank function only, and in the second part we introduce a graph process overlaid on the rank function. Separating these two pieces of the model aids tractability and illustrates the independence of the ranking and graph evolution processes.

All results presented here are asymptotic (that is, with $n$ tending to infinity). We say that an event holds *asymptotically almost surely* (*aas*), if it holds with probability tending to one as $n \to \infty$. We will sometimes use the stronger notion of *wep* in favour of the more commonly used *aas*, since it simplifies some of our proofs. We say that an event holds *with extreme probability* (*wep*), if it holds with probability at least $1 - \exp(-\Theta(\log^2 n))$ as $n \to \infty$. Thus, if we consider a polynomial number of events that each holds *wep*, then *wep* all events hold. To combine this notion with asymptotic notations such as $O()$ and $o()$, we follow the conventions in [22].

**2.1. A rank model of content.** At each time $t$, we have exactly $n$ objects (such as nodes, who may represent agents in a social system) in a set $V_t$. Moreover, at each time $t$, each object $v \in V_t$ has rank $r_t(v) \in [n]$ (we use $[n]$ to denote the set $\{1, 2, \ldots, n\}$). In order to obtain a proper ranking, the rank function $r_t : V_t \to [n]$ is a bijection for all $t$, so every object has a unique rank. In agreement with the common use of the word "rank," high rank refers to an object $v$ for which $r_t(v)$ is small: the highest ranked object is ranked number one, so has rank equal to 1; the lowest ranked object has rank $n$. The initialization and update of the ranking is done according to a *ranking scheme*. Various ranking schemes may be considered, and each might lead to different behaviors. We first define the general model, and then list a few natural ranking schemes.

The model produces a sequence $\{(V_t, r_t)\}_{t=0}^{\infty}$ of sets $V_t$ of $n$ objects and ranking functions $r_t$, where $t$ denotes time. To initialize the model, let $V_0$ be any set of $n$ objects and let $r_0$ be any initial rank function $r_0 : V_0 \to [n]$ which is consistent with the ranking scheme. For $t \geq 1$ we form $(V_t, r_t)$ from $(V_{t-1}, r_{t-1})$ according to the following rules:

(i) Choose uniformly at random an object $u_t \in V_{t-1}$ and delete it.
(ii) Add a new object $v$. (We refer to the time step $t$ in which object $v$ was added as time in which $v$ was born.)
(iii) Assign an initial rank to $v$, update $V_t$ and the ranking function $r_t : V_t \to [n]$ according to the ranking scheme.

One can define a number of different ranking schemes. In this paper, we focus on the random initial rank scheme but the concept of the ranking by age will also be important. Therefore, let us define the following two schemes. In order to distinguish them, we will use $a_t$ for the ranking by age and $r_t$ for the random initial rank.

(i) **Ranking by age**: The newly added object $v$ obtains an initial rank $n$; its rank decreases by one each time an object with smaller rank is removed. Formally, for each $v \in V_{t-1} \setminus \{u_t\}$, $a_t(v) = a_{t-1}(v) - \gamma$, where $\gamma = 1$ if the rank of the object deleted in step $t$ is smaller than $a_{t-1}(v)$, and 0 otherwise.

(ii) **Random initial rank**: The object added at time $t$ obtains an initial rank $R_t$ which is randomly chosen from $[n]$ according to a prescribed distribution. Ranks of all objects are adjusted accordingly. Formally, for each $v \in V_{t-1} \setminus \{u_t\}$, $r_t(v) = r_{t-1}(v) + \delta - \gamma$, where $\delta = 1$ if $r_{t-1}(v) > R_t$ and 0 otherwise, and $\gamma = 1$ if where the rank of $u_t$, the object deleted in step $t$, is smaller than $r_{t-1}(v)$, and 0 otherwise.

The coupon collector problem can give us insight into when all objects from the initial set $V_0$ will be deleted. Namely, let $L = n(\log n + \omega(n))$ where $\omega(n)$ is any function tending to infinity with $n$. It is a well-known result that *aas* after $L$ steps all original objects will have been deleted.

The behavior of the age rank function $a_t(v)$ was studied in [11].

**Theorem 2.1** ([11])**.** *Let $a_t(v)$ be the age rank of object $v$ at time $t$. Then* wep*, for every $t$ in the range $0 \le t \le t_f = \frac{1}{2}n \log n - 2n \log\log n$, we have*

$$a_t(v) = n \exp(-t/n)(1 + O(\log^{-1/2} n))$$

*conditional upon the object $v$ surviving until time $t_f$.*

We consider the case where the rank $R_i$ of the object $v$ added at time $i$ is chosen at random from $[n]$. The ranks of existing objects are adjusted accordingly. We make the assumption that all initial ranks are chosen according to the same distribution. In particular, we fix a continuous bijective function $F : [0, 1] \to [0, 1]$, and for all integers $1 \le k \le n$, we let

$$\mathbb{P}(R_i \le k) = F\left(\frac{k}{n}\right).$$

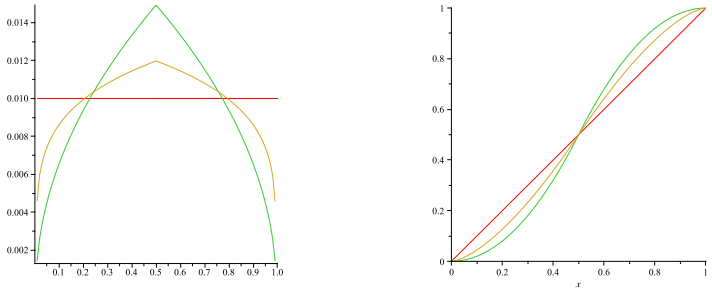
Thus, $F$ represents the limit, for $n$ going to infinity, of the cumulative distribution functions of the variables $R_i$. To simplify the calculations while exploring a wide array of possibilities for $F$, we assume $F$ to be of the form

$$F(x) = \begin{cases} (2x)^s/2 & \text{if } 0 \le x \le 1/2 \\ 1 - (2(1-x))^s/2 & \text{if } 1/2 < x \le 1 \end{cases}, \text{ where } s \ge 1.$$

In [11], a similar function was analyzed, namely $F(x) = x^s$. We modify the function slightly to obtain the behaviour we desire, namely, that average initial ranks occur more often comparing to both extreme cases. As noted earlier, the proofs are easily adapted to a new function.

This distribution has the advantage of allowing us to generalize our results to a broad class of realistic initial distributions, ranging from situations where initial ranks are distributed uniformly at random (when $s = 1$) to situations where agents enter the system with a mediocre rank with higher probability (when $s > 1$; in this case the highest probability rank is $n/2$). For example, see Figure 1 (a)-(b) to see the differences in distributions across $s = 1.0$, $s = 1.2$, and $s = 1.5$. This functional form is reasonable because it reflects the notion that many types of attributes follow a Normal distribution in social systems; it tends to be unlikely that new agents will be "born" into the system with a very low rank or a very high rank. If rank represents a type of dynamic fitness

where agents compete for better ranks, entering agents are unlikely to have poor ranks because then they may not be able to enter the system at all, and they are also unlikely to enter with very good ranks, which are obtained only through a history of competition in the system. Our functional form for $F(x)$ reflects these possibilities.



(a) $f(x) = \mathbb{P}\left(x \le R/n \le x + \frac{1}{100}\right)$ (b) $F(x) = \mathbb{P}(R/n \le x)$.

FIGURE 1. Different distributions: $s = 1.0, 1.2, 1.5$.

The case $s = 1$ represents the uniform distribution of the $R_i$. The random variable $r_t(v)$ is sharply concentrated around the initial rank $R_i$. The following result was obtained in [11].

**Lemma 2.2.** *Suppose that object $v$ obtained an initial rank $R \ge \sqrt{n}\log^2 n$. Then,* wep

$$r_t(v) = R(1 + O(\log^{-1/2} n))$$

*to the end of its life.*

In the case $s > 1$, the initial rank is biased towards the middle range ranks. The rank function exhibits more complex behaviour in this case. Due to the symmetry of the function $F(x)$, without loss of generality we can assume that an initial rank is at most $\frac{n}{2}$. For ranks close to $\frac{n}{2}$ we clearly cannot predict the behaviour; the final rank can be bigger or smaller than the initial rank. However, if the initial rank is separated a bit from the middle rank, then we get a concentration.

**Theorem 2.3.** *Suppose that an object $v$ obtained an initial rank*

$$r_0(v) = R < \frac{n}{2} - \sqrt{n}\log^2 n$$

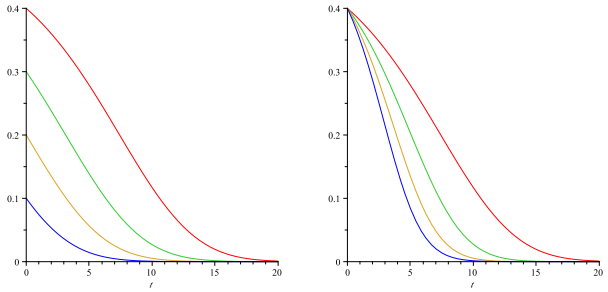
*at time 0. Then* wep*, for every $t$ in the range $0 \le t \le t_f = \frac{1}{2}n\log n - 2n\log\log n$ conditional upon the object $v$ surviving until time $t$,*

$$r_t(v) = \frac{n}{2}\left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)e^{(s-1)t/n} + 1\right)^{\frac{1}{1-s}}(1 + O(\log^{-1/2} n)) \tag{1}$$

*provided*

$$\frac{n}{2}\left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)e^{(s-1)t/n} + 1\right)^{\frac{1}{1-s}} \ge \sqrt{n}\log^2 n\,.$$

Figure 2 (a),(b) presents the behaviour of different initial ranks for one specific value of $s = 1.2$ as well as the behaviour of one specific initial rank $R = 0.4n$ for different values of $s$. Both rank and time are scaled by $n$.



(a) $R = 0.1, 0.2, 0.3, 0.4$ (b) $s = 1.2, 1.3, 1.4, 1.5$

FIGURE 2. (a) The behaviour for different initial ranks ($s = 1.2$); (b) The behaviour for different values of $s$ ($R = 0.4$).

**2.2. A rank model of network structure.** In this subsection, we introduce the network on top of the process discussed in previous sections. We need two more parameters, the *attachment strength* $\alpha \in (0, 1)$ and *initial degree* $d \in \mathbb{N}$. This time, the model produces a sequence $\{(G_t, r_t)\}_{t=0}^{\infty}$ of graphs $G_t = (V_t, E_t)$ on $n$ vertices and ranking functions $r_t : V_t \to [n]$. To initialize the model, let $G_0$ be any graph on $n$ vertices and let $r_0$ be any initial rank function $r_0 : V_0 \to [n]$ which is consistent with the ranking scheme. For $t \geq 1$ we form $G_t$ from $G_{t-1}$ according to the following rules:

(i) Choose uniformly at random a vertex $u_t \in V_{t-1}$ and delete it.
(ii) Add a new vertex $v_t$ together with $d$ edges from $v_t$ to existing vertices chosen randomly with weighted probabilities. The edges are added in $d$ substeps. In each substep, one edge is added, and the probability that $v_i$ is chosen as its endpoint (the link probability), is proportional to $r_{t-1}(v_i)^{-\alpha}$.
(iii) Assign an initial rank to $v_t$, update $V_t$ and the ranking function $r_t : V_t \to [n]$ according to the ranking scheme.

Similarly as before, we need to wait until time $L$ when all vertices from the initial graph $G_0$ are deleted. At this point we reach the stationary distribution and our task is to analyze the behaviour of a random graph $G_L$. As we already mentioned, *aas* $L \leq n(\log n + \omega(n))$ where $\omega(n)$ is any function tending to infinity with $n$.

In [11], it was shown that the uniform distribution for the initial rank (that is, the specific case of $s = 1$ in our model) generates a power-law degree distribution with exponent $1 + 1/\alpha$. One can adopt the proofs from [11] and show that this is also the case for $s > 1$.

## 3. The discovery of nodal attributes through structure

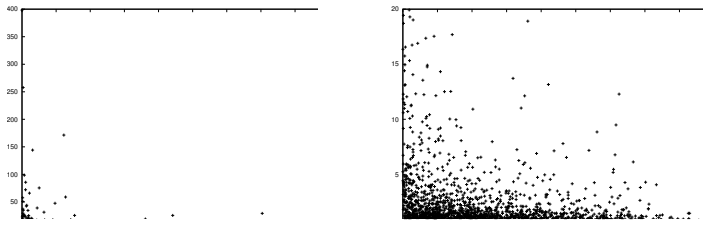
While the Web graph is a useful platform for social networks research, the notion that networks self-organize as a function of nodal attributes suggests the need to observe

both structure as well as attributes of the nodes. While structures may be observed directly, for example through hyperlink data, in many cases attributes (ranks) of agents embedded in the network will be latent, unobserved variables. However, given a realistic model of the process by which the network was generated, it is possible to infer likely attributes of nodes. Here we begin with a heuristic argument for how nodal attributes may be inferred from structure, followed by a more rigorous analysis of this problem. The simulations presented in the paper were performed by using a program written in C/C++, which may be found at [24].

**3.1. A heuristic argument.** We focus on time $L$ when all vertices from the initial graph $G_0$ are deleted and the stationary distribution of the Markov chain is reached. Consider the degree $\deg(v, L)$ of a given node. In this section, for simplicity let us focus on 'in-degree' $\deg^-(v, L)$ of $v$; that is, the number of neighbors of $v$ that are older than this vertex.

Note that for a vertex of high degree, the number of younger neighbors is negligible comparing to the total degree. Moreover, despite of the fact that precise ages of vertices are difficult to estimate in practice, given any edge $uw$ it is known which vertex is older. Thus, it makes sense to focus on in-degrees here.

Given the model outlined here, this degree (or in-degree) is a function of two factors related to content: first, the length of time the node has been in the system, and second, the initial rank assigned to the node when it was born into the network. Agents with smaller initial ranks tend to have larger degrees, and older vertices also tend to have larger degrees. Despite this correlation, the true relationship is quite complicated and it would seem to be a lost cause to try to infer only one of these attributes (age or rank) based on degree only. Figure 3 presents the relation between age and degree for vertices of degree at least $d/2$ when networks are simulated according to the model described here ($n = 20,000$, $d = 100$, $s = 1.5$, and $\alpha = 0.8$). Young vertices have small degree (there is no time to accumulate neighbors, even if the initial rank is good) but old vertices can still have small degree (because they have an unattractive rank).




(a) (b)

FIGURE 3. In-degree of $v$ vs. the age rank of $v$ (rescaled)

It turns out, however, that it becomes feasible to estimate these properties when we broaden our focus from the degree of a single agent to properties of their second neighborhood. Consider, for example, the following coefficient defined for a vertex $v$ of

non-zero degree that is proportional to the average degree of neighbors of $v$:

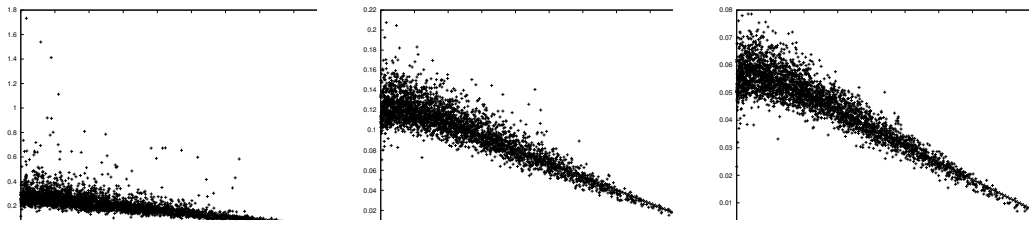
$$b_2(v) = \frac{\sum_{u \sim v} \deg^-(u)}{\deg^-(v)}.$$

We put $b_2(v) = 0$ if $\deg^-(v) = 0$. Clearly, old nodes have more old neighbors compared to younger nodes. In other words, there is a correlation between the age of $v$ and ages of its neighbors. On the other hand, ranks are generated independently, so a distribution of ranks of the neighbors of $v$ should be similar to the distribution we use in the model. The more neighbors $v$ has, the stringer the correlation should be. Older vertices should have larger $b_2(v)$ coefficients. See Figure 4(a) for the relation for vertices of degree at least $d/2$.

This process can be carried even further to develop more finely-tuned estimates of nodes' unobserved attributes. We can take a look at third, fourth, and higher-order neighborhoods by defining, recursively, for $k \geq 3$

$$b_k(v) = \frac{\sum_{u \sim v} b_{k-1}(u)}{\deg^-(v)},$$

provided that $\deg^-(v) > 0$; otherwise, $b_k(v) = 0$. Again, in this case older vertices should have larger coefficients and the error should decrease for, say, $k = 3$ and $k = 4$. See Figure 4(b-c) for the results for $b_3(x)$ and $b_4(x)$.



(a) $b_2(x)$ (b) $b_3(x)$ (c) $b_4(x)$

FIGURE 4. $b_k(x)$ vs. the age rank of $v$ (rescaled)

Even upon casual examination, these scatterplots reveal a strong, nearly linear, relationship between the average degree of neighbors (or higher-order $b_k$ coefficients) and the age of node. Inferring rank rather than age may be accomplished in a similar way.

One way of viewing the increasing predictive power of these structural characteristics is to perform an simple OLS linear regression with node age as the dependent variable (the unobserved variable to be inferred in real-world applications) and degree or various $b_k$ measures as possible independent variables. Using this approach, we find that predicting age as a linear function of average degree becomes more precise as we move to higher-order neighborhoods. For example, the $R^2$ statistic when age is predicted using degree only is 0.01, meaning that node degree explains only 1% of the variance in actual node age. When the average degree of neighbor ($b_2$) is used as an independent variable, a linear model explains 35% of the variance on age ($R^2 = 0.35$). $R^2$ jumps to 0.77 for $b_3$, and 0.83 for $b_4$.

These regression models provide at least heuristic evidence that one can achieve fairly accurate predictions of age when one examines the degree of neighbors, and neighbors of neighbors, and so forth. And while these linear models are suggestive of strong patterns, the scatterplots also make it clear that the accuracy with which we can predict age depends on the degree of the node. In particular, it seems that for low-degree nodes age may be predicted with fair accuracy (in particular because, having just entered the system, the number of relationships is a more direct result of initial rank) while the relationship between $b_k$ and age for high-degree nodes is less precise.

It is also interesting to note that going from the second neighborhood to the third neighborhood provides a smaller marginal benefit in terms of predictive power, as measured by the $R^2$ values. While examination of the third neighborhood provides the strongest inferences regarding age, of course there will be an upper bound on the depth of neighborhoods that may be examined, plus there is likely to be an optimal neighborhood to examine in terms of maximizing the predictive power of this method. These issues, along with the strength of predictions that may be made for small- versus high-degree agents, are discussed presently.

**3.2. A rigorous approach to inferring nodal attributes.** In order to obtain an estimation for $b_2(v)$ (or $b_k(v)$, in general) we need to calculate the expected (in- and total) degree of a vertex $v_i$ with age-rank $a_L(v_i) = i = xn$. In the proof of Theorem 5.6 (see the Appendix) we provide a useful estimation for the case where the expected value is tending to infinity together with $n$. This was enough to prove that the degree distribution follows a power-law, despite the fact that almost all vertices have a constant expected degree. When the expected degree does not grow with $n$ we cannot, of course, expect the degree to be well concentrated around this value. However, when a large number of vertices are considered, the average degree must be very close to what we expect (by the Chernoff bound). Hence, if the degree of a vertex $v$ (for which we aim to have a well-estimated parameter) is large enough, then the value of $b_2(v)$ is well concentrated (recall that $b_2(v)$ is, in fact, an average degree of a neighbor of $v$).

Suppose that the initial rank of $v_i$ is at most $n/2 - \sqrt{n}\log^2 n$. A more careful estimation than the one used in Theorem 5.6 will give us an asymptotic value of the expected degree of $v_i$ at time $L$. To get a better estimation for the number of initial neighbors of $v_i$ that are not deleted at time $L$, we note that $i-1$ vertices are older than $v_i$, and hence we expect that $v_i$ has $\frac{i-1}{n-1}d = (1+o(1))xd$ older neighbors. Therefore,

$$\mathbb{E}\deg(v_i, L) \sim xd + d(1-\alpha)2^{\alpha}\int_x^1 \left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{-\alpha}{1-s}} dy,$$

where the notation $a \sim b$ stands for $a = (1+o(1))b$. In the formula for $\deg^-(v_i, L)$, there is no $xd$ term. The rest remains exactly the same.

Suppose now that the initial rank of $v_i$ is at least $n/2 + \sqrt{n}\log^2 n$. Due to the symmetry of ranks, the rank of $v_i$ behaves exactly as $n - r_t(v_i^*)$ where $v_i^*$ is a vertex that obtained the initial rank of $n - R \leq n/2 - \sqrt{n}\log^2 n$. Using Theorem 2.3 as before,

we get that *wep* $v_i$ had the following rank when $v_j$ $(j = yn)$ was born:

$$r_{t_j}(v_i) \sim n - \frac{n}{2}\left(\left(\left(\frac{2(n-R)}{n}\right)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{1}{1-s}},$$

and so

$$\mathbb{E}\deg(v_i, L) \sim xd + d(1-\alpha)\int_x^1 \left(1 - \frac{1}{2}\left(\left(\left(\frac{2(n-R)}{n}\right)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{1}{1-s}}\right)^{-\alpha} dy.$$

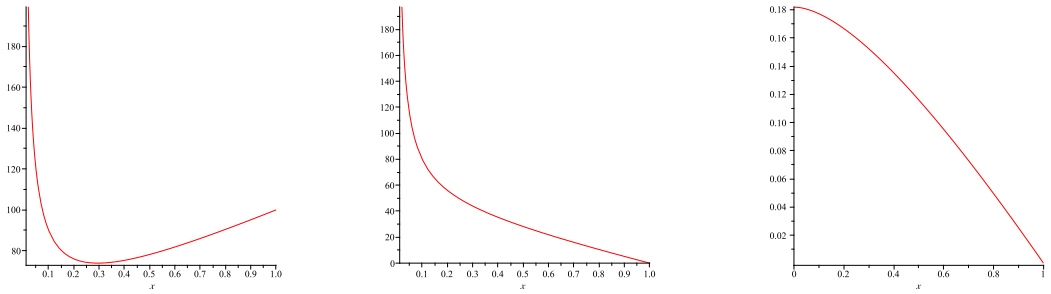
(Again, in the formula for $\deg^-(v_i, L)$ there is no $xd$ term; the rest remains exactly the same.) Since $\mathbb{P}(R \le zn) = F(z) = (2z)^s/2$ for $z \le 1/2$ and $F(z) = 1 - F(1-z)$, we get immediately the following result. Let

$$\begin{aligned} g(x) &= d(1-\alpha)s2^{\alpha+s-1}\int_0^{1/2} z^{s-1}\int_x^1 \left(\left((2z)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{-\alpha}{1-s}} dydz \\ &+ d(1-\alpha)s2^{s-1}\int_0^{1/2} z^{s-1}\int_x^1 \left(1 - \frac{1}{2}\left(\left((2z)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{1}{1-s}}\right)^{-\alpha} dydz. \end{aligned}$$

**Theorem 3.1.** *Let $0 < \alpha < 1$, $d \in \mathbb{N}$, $i = i(n) \in [n]$, and let $v_i$ be the vertex whose age rank at time $L$ equals $a_L(v_i) = i = xn$. Then, the expected (in-)degree of $v_i$ is given by*

$$\begin{aligned} \mathbb{E}\deg(v_i, L) &\sim xd + g(x) \\ \mathbb{E}\deg^-(v_i, L) &\sim g(x). \end{aligned}$$

It appears that there is no closed formula for $\mathbb{E}\deg(v_i, L)$ and $\mathbb{E}\deg^-(v_i, L)$ but it may easily be estimated numerically. Figure 5(a-b) presents the expected degree for the set of parameters used before (that is, $d = 100$, $s = 1.5$, and $\alpha = 0.8$). The computations presented in the paper were performed by using Maple$^{\text{TM}}$ [15]. The worksheets may be found at [24].



(a) $xd + g(x)$ $(\mathbb{E}\deg(v_i, L))$ (b) $g(x)$ $(\mathbb{E}\deg^-(v_i, L))$ (c) $b_2(x)$ (asymptotic formula)

FIGURE 5. (a-b) Degree distribution for $d = 100$, $s = 1.5$, and $\alpha = 0.8$ as a function of age-rank (rescaled); (c) $b_2(x)$ vs. the age rank of $v$ (rescaled) for the following set of parameters: $d = 100$, $s = 1.5$, and $\alpha = 0.8$

Let us note that the behaviour we experience here is similar to the one for the Protean graph [14] which may be viewed as the father of rank-based models. Old vertices have very large expected value, since they have non-negligible probability of getting small initial rank (which implies a good rank during the whole process and thus large expected degree). Young vertices still have relatively large expected degree: most of initial neighbors are sill in the network with high probability and so the expected degree is close to $d$. In the worst position are middle-aged vertices, who are not old enough to attract new neighbors but have also had many initial neighbors dissapear. The expected in-degree is, of course, a decreasing function: the older the vertex, the more time to accumulate a high degree.

Now, with Theorem 3.1 in hand we can easily derive a formula for $b_2(v_i)$. Suppose that $\deg^-(v_i, L) \geq \log^4 n$ (or $\deg(v_i, L) \geq \log^4 n$, since the out-degree is negligible) with $i = xn$, and the initial rank of $v_i$ is $R$. It follows from (3) that

$$\mathbb{E}\left(\sum_{u\sim v_i} \deg^-(u, L)\right) \sim d(1-\alpha)2^\alpha \left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)^{\frac{-\alpha}{1-s}} x^{-\alpha} \int_x^1 y^\alpha \cdot g(y)dy$$

Since the in-degree of $v_i$ is well concentrated around its expectation *wep*, we are ready to state the last result of this paper. Figure 5(c) then presents the value of $b_2(v_i)$ based on the asymptotic formula.

**Theorem 3.2.** *Let $0 < \alpha < 1$, $d \in \mathbb{N}$, $i = i(n) \in [n]$, and let $v_i$ be the vertex whose age rank at time $L$ equals $a_L(v_i) = i = xn$ and $\deg^-(v_i, L) \geq \log^4 n$. Then,* wep

$$b_2(v_i) \sim \frac{\mathbb{E}\left(\sum_{u\sim v_i} \deg^-(u, L)\right)}{\mathbb{E}\deg^-(v_i, L)} \sim \frac{1+\alpha}{1-x^{1+\alpha}} \int_x^1 y^\alpha \cdot g(y)dy.$$

## 4. Conclusion

This paper outlined a model of network self-organization that is driven by the ranks of individual nodes, in terms of arbitrary attributes that are inherently individual phenomenon. This model may be applied to a broad array of social systems, where nodal attributes (and rank) may represent wealth, power, beliefs, skills, or any other node-level variables that are likely to play an important role in networking behavior. This is a stochastic model involving the formation and deletion of network ties, and adjustment in ranks, as nodes dynamically enter and exit the system over time.

This research builds upon prior work in ranking and associated models of network self-organization, and continues the enterprise of linking these network models to enhance our understanding of the dynamics of real-world social systems. An important area for future research is to carefully consider how network structure evolves as a function of nodal attributes. Of course, this not only requires models of networks per se, but also requires models of attributes of individuals and how these attributes are manifest in network structure.

On the other hand, nodal attributes are often exceedingly difficult to observe and can be a limiting factor on our ability to study complex, self-organizing networks. However, given realistic models of how network structure is driven by these attributes, it is possible to estimate individual attributes based on structure only. This research

provides a platform for more research, emphasizing analytical proof, that exploits the potential "reversibility" of mathematical models to infer latent, unobserved variables that are crucial to the further development of network science.

## References


[1] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat, *A spatial web graph model with local influence regions*, Internet Mathematics **5** (2009), 175–196.

[2] A.-L. Barabási and R. Albert, *Emergence of Scaling in Random Networks*, Science, vol. **286**, no. 5439, pp. 509–512, Oct. 1999.

[3] A. Bonato, *A Course on the Web graph*, American Mathematical Society Graduate Studies Series in Mathematics, Providence, Rhode Island, 2008.

[4] F.R.K. Chung and L. Lu, *Complex graphs and networks*, American Mathematical Society, U.S.A., 2004.

[5] S. Fortunato, A. Flammini, and F. Menczer, *Scale-free network growth by ranking*, Phys. Rev. Lett. **96**(21): 218701 (2006).

[6] A. Henry and P. Prałat, *Rank-Based Models of Network Structure and the Discovery of Content*, Proceedings of the 8th Workshop on Algorithms and Models for the Web Graph (WAW 2011), Lecture Notes in Computer Science **6732**, Springer, 2011, 62–73.

[7] A.D. Henry, P. Prałat, and C.Q. Zhang, *Emergence of segregation in evolving social networks*, Proceedings of the National Academy of Sciences, vol. **108**, no. 21, pp. 8605–8610, May. 2011.

[8] M.O. Jackson, *Social and Economic Networks*, Princeton University Press, Princeton, NJ, 2008.

[9] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[10] J. Janssen, P. Prałat, and R. Wilson, *Estimating node similarity from co-citation in a spatial graph model*, Proceedings of the 2010 ACM Symposium on Applied Computing (SAC) – Special Track on Self-organizing Complex Systems, 2010, 1329–1333.

[11] J. Janssen and P. Prałat, *Protean graphs with a variety of ranking schemes*, Theoretical Computer Science **410** (2009), 5491–5504.

[12] J. Janssen and P. Prałat, *Rank-based attachment leads to power law graphs* SIAM Journal on Discrete Mathematics **24** (2010), 420–440.

[13] D. Lazer, *The co-evolution of individual and network*, The Journal of Mathematical Sociology, vol. **25**, no. 1, 69–108, Mar. 2001.

[14] T. Łuczak and P. Prałat, *Protean graphs*, Internet Mathematics **3** (2006), 21–40.

[15] M.B. Monagan, K.O. Geddes, K.M. Heal, G. Labahn, S.M. Vorkoetter, J. McCarron, and P. DeMarco, Maple 10 Programming Guide, Maplesoft, Waterloo ON, Canada, 2005.

[16] M.E.J. Newman, *The Structure and Function of Complex Networks*, SIAM Review, vol. **45**, no. 2, 167–256, 2003.

[17] B. Pittel, J. Spencer and N. Wormald, *Sudden emergence of a giant k-core in a random graph*, J. Combinatorial Theory, Series B **67** (1996), 111–151.

[18] P. Prałat, *A note on the diameter of protean graphs*, Discrete Mathematics **308** (2008), 3399–3406.

[19] P. Prałat, *Connectivity threshold and recovery time in rank-based models for complex networks*, Discrete Mathematics **311** (2011), 932–939.

[20] P. Prałat and N. Wormald, *Growing protean graphs*, Internet Mathematics **4** (2009), 1–16.

[21] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.

[22] N.C. Wormald, Random graphs and asymptotics. Section 8.2 in *Handbook of Graph Theory*, J.L. Gross and J. Yellen (eds), pp. 817–836. CRC, Boca Raton, 2004.

[23] N. Wormald, *The differential equation method for random graph processes and greedy algorithms* in Lectures on Approximation and Randomized Algorithms, eds. M. Karoński and H. J. Prömel, PWN, Warsaw, pp. 73–155, 1999.

[24] `http://www.math.ryerson.ca/~pralat/`, accessed June 20, 2011.

## 5. Appendix

**5.1. Introduction to the Differential Equations method.** The general setting that is used in the DEs method [23] is a sequence of random processes indexed by $n$ (which in our case is the number of objects in any set $V_t$). The aim is to find properties of the random process in the limit as $n \to \infty$. The conclusion we aim for is that variables defined on a random process are well concentrated, which informally means that with high probability they are very close to certain deterministic functions. These functions arise as the solution to a system of ordinary first-order differential equations. One of the important features of this approach is that the computation of the approximate behavior of processes is clearly separated from the proof that the approximation is correct.

To show that the random variables in a process usually approximate the solution of differential equations, we need to use large deviation inequalities. These inequalities are often used to give an upper bound on the probability that a random variable deviates very far from its expected value. In a typical situation with a random process, the aim is to show that the random variable $Y_t$ of interest is sharply concentrated. In fact,

$$Y_t - Y_0 = \sum_{i=1}^{t} (Y_i - Y_{i-1}).$$

If the differences $Y_i - Y_{i-1}$ are independent, then the Chernoff bound is very useful (see for example Theorem 2.8 [9]).

**Theorem 5.1** ([9])**.** *Let $X$ be a random variable that can be expressed as a sum $X = \sum_{i=1}^{n} X_i$ of independent random indicator variables where $X_i \in \mathrm{Be}(p_i)$ with (possibly) different $p_i = \mathbb{P}(X_i = 1) = \mathbb{E}X_i$. Then the following holds for $t \geq 0$:*

$$\begin{aligned} \mathbb{P}(X \geq \mathbb{E}X + t) &\leq \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right), \\ \mathbb{P}(X \leq \mathbb{E}X - t) &\leq \exp\left(-\frac{t^2}{2\mathbb{E}X}\right). \end{aligned}$$

*In particular, if $\varepsilon \leq 3/2$, then*

$$\mathbb{P}(|X - \mathbb{E}X| \geq \varepsilon \mathbb{E}X) \leq 2\exp\left(-\frac{\varepsilon^2 \mathbb{E}X}{3}\right).$$

When the differences are not independent but there is a large degree of independence, results can often be obtained by making use of analogous bound given for martingales.

**Definition 5.2.** *A **martingale** is a sequence $X_0, X_1, \ldots$ of random variables defined on the random process such that*

$$\mathbb{E}(X_{n+1} \mid X_0, X_1, \ldots, X_n) = X_n.$$

*In most applications, the martingale satisfies the property that*

$$\mathbb{E}(X_{n+1} \mid X_0, X_1, \ldots, X_n) = \mathbb{E}(X_{n+1} \mid X_n) = X_n.$$

As a simple example, consider the following "random walk." Toss a coin $n$ times. Let $S_n$ be the difference between the number of heads and the number of tails after $n$ tosses. $S_n$ is a martingale. Indeed,

$$\mathbb{E}(S_{n+1} \mid S_n) = S_n + \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = S_n.$$

Clearly, the expected value of $S_n$ is zero. Thus, it is natural to expect that $S_n$ stays relatively close to zero. The following well-known Hoeffding-Azuma inequality serves as a tool to investigate this.

**Lemma 5.3.** *Let $X_0, X_1, \ldots$ be a martingale. Suppose that there exist constants $c_k > 0$ such that*

$$|X_k - X_{k-1}| \leq c_k$$

*for each $k \leq n$. Then, for every $t > 0$,*

$$\begin{aligned} \mathbb{P}(X_n \geq \mathbb{E}X_n + t) &\leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right), \\ \mathbb{P}(X_n \leq \mathbb{E}X_n - t) &\leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right). \end{aligned}$$

This is often applied with $t$ growing much faster than $\sqrt{n}$ and the $c_k$ all small non-zero integers. In the martingale discussed above $c_k = 1$ for all $k$. Hence,

$$\mathbb{P}(|S_n| \geq \alpha\sqrt{n}) \leq 2\exp\left(\frac{(\alpha\sqrt{n})^2}{2n}\right) = 2\exp\left(\alpha^2/2\right),$$

which is arbitrarily small for $\alpha$ large enough.

Finally, let us mention that the Hoeffding-Azuma inequality can be generalized in many ways: analogous inequality holds for supermartingales ($\mathbb{E}(X_{n+1}|X_n) \leq X_n$) as well as submartingales ($\mathbb{E}(X_{n+1}|X_n) \geq X_n$). Our proofs use the supermartingale method of Pittel et al. [17], as described in [23, Corollary 4.1]. We will use the following useful lemma.

**Lemma 5.4.** *Let $G_0, G_1, \ldots, G_L$ be a random process and $X_t$ a random variable determined by $G_0, G_1, \ldots, G_t$, $0 \leq t \leq L$. Suppose that for some real $\beta$ and $\gamma$,*

$$\mathbb{E}(X_t - X_{t-1} \mid G_0, G_1, \ldots, G_{t-1}) < \beta$$

*and*

$$|X_t - X_{t-1} - \beta| \leq \gamma$$

*for $1 \leq t \leq L$. Then for all $\varepsilon > 0$,*

$$\mathbb{P}\big(\textit{For some } t \textit{ with } 0 \leq t \leq L : X_t - X_0 \geq t\beta + \varepsilon\big) \leq \exp\Big(-\frac{\varepsilon^2}{2L\gamma^2}\Big).$$

### 5.2. Proof of Theorem 2.3.

*Proof.* The conditional expected change in $r_t(v)$ in time step $t+1$, conditional on object $v$ surviving until time $t+1$ and that $r_t(v) \leq \frac{n}{2}$, is given by:

$$\mathbb{E}(r_{t+1}(v) - r_t(v) \mid G_t) = -\frac{r_t(v)-1}{n-1} + 2^{s-1}\left(\frac{r_t(v)}{n}\right)^s.$$

Defining a real function $z(x)$ to model the behaviour of $r_{xn}(v)/n$, this suggests the differential equation

$$z'(x) = -z(x) + 2^{s-1}z(x)^s,$$

with the initial condition $z(0) = R/n$. The general solution is

$$z(x) = \frac{1}{2}\left(Ce^{(s-1)x}+1\right)^{\frac{1}{1-s}}, \quad C \in \mathbb{R},$$

and the particular solution is

$$z(x) = \frac{1}{2}\left(\left(\left(\frac{R}{n}\right)^{1-s} - 1\right)e^{(s-1)x} + 1\right)^{\frac{1}{1-s}}.$$

We need to transform $r_t(v)$ into something close to a martingale. Since the general solution can be written in the form

$$\log\left(\left(\frac{1}{2z(x)}\right)^{s-1} - 1\right) - (s-1)x = C \in \mathbb{R},$$

we should consider the following real-valued function

$$H(r,t) = \log\left(\left(\frac{n}{2r_t}\right)^{s-1} - 1\right) - (s-1)\frac{t}{n} \tag{2}$$

We hope that a good concentration for $H(r,t)$ implies a good behavior for $r_t(v)$.

Clearly, we cannot control the process when the rank of $v$ drops below some small value (say, $\sqrt{n}\log^2 n$). However, up to this point of the process we should expect good concentration for these random variables. But how can we assume that the random variable is above the threshold when we are about to investigate its shape? The following concept, simple but very powerful, comes to the rescue. Define the stopping time as follows:

$$T = \min\left\{t \geq 0 : r_t(v) < \frac{1}{2}\sqrt{n}\log^2 n \ \vee \ r_t(v) > \frac{n}{2} - \frac{1}{2}\sqrt{n}\log^2 n \ \vee \ t = t_f\right\}.$$

A stopping time is any random variable $T$ with values in $\{0,1,\dots\} \cup \{\infty\}$ for which it can be determined whether $T = \hat{t}$ for any time $\hat{t}$ from knowledge of the process up to and including time $\hat{t}$. The name can be misleading, since a process does not *stop* when it reaches a stopping time. The key observation is that if a supermartingale $(X_i)$ is stopped at a stopping time (that is, $(X_i)$ becomes static for all time after the stopping time), then the result is a supermartingale.

Let $\mathbf{w}_t = (r_t(v), t)$, and consider the sequence of random variables $(H(\mathbf{w}_t) : 0 \leq t \leq t_f)$. Let us stress it again, $H$ is chosen so that $H(\mathbf{w})$ is close to a constant along

every trajectory $\mathbf{w}$ of the differential equation. It is easy to check that the second-order partial derivatives of $H$ are $O(r^{-2}) = O(n^{-1}\log^{-4} n)$ along the trajectory $\mathbf{w}_t$, provided $T > t$. Therefore, with $i \wedge T$ denoting $\min\{i, T\}$, using the fact that the rank changes by at most one in each step, we have

$$\begin{aligned} |H(\mathbf{w}_{(t+1)\wedge T}) - H(\mathbf{w}_{t\wedge T})| &= O(1/r_{t\wedge T}(v)) = O(1/\sqrt{n}\log^2 n) \\ \mathbb{E}(H(\mathbf{w}_{(t+1)\wedge T}) - H(\mathbf{w}_{t\wedge T}) \mid G_{t\wedge T}) &= O(1/n\log^4 n). \end{aligned}$$

Now we may apply Lemma 5.4 to the sequence $(H(\mathbf{w}_{t\wedge T}) : 0 \le t \le t_f)$, and symmetrically to $(-H(\mathbf{w}_{t\wedge T}) : 0 \le t \le t_f)$, with $\varepsilon = 1/\log^{1/2} n$, $\beta = O(1/n\log^4 n)$, and $\gamma_t = O(1/\sqrt{n}\log^2 n)$ to show that *wep*

$$|H(\mathbf{w}_{t\wedge T}) - H(\mathbf{w}_{t_0})| = O(\log^{-1/2} n).$$

As $H(\mathbf{w}_0) = \log((2R/n)^{1-s} - 1)$, it follows from the definition of $H$ that *wep*

$$\begin{aligned} r_t(v) &= \frac{n}{2}\left(e^{H(\mathbf{w}_0)}e^{(s-1)t/n} + 1\right)^{\frac{1}{1-s}}(1 + O(\log^{-1/2} n)) \\ &= \frac{n}{2}\left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)e^{(s-1)t/n} + 1\right)^{\frac{1}{1-s}}(1 + O(\log^{-1/2} n)) \end{aligned}$$

for every $0 \le t \le T$.

To complete the proof we need to show that *wep*, $T = t_f$. The events asserted by (1) hold *wep* up until time $T$, as shown above. Thus, in particular, *wep* $r_T(v) > (1 + o(1))\sqrt{n}\log^2 n$ and $r_T(v) < n/2 - (1 + o(1))\sqrt{n}\log^2 n$, which implies that $T = t_f$ *wep*. □

**5.3. Power law for $s > 1$.** Let $Z_k$ denote the number of vertices of degree $k$, and $Z_{\ge k} = \sum_{l \ge k} Z_l$.

**Theorem 5.5.** *Let $0 < \alpha < 1$ and $d \in \mathbb{N}$, $\log^4 n \le k \le n^{\alpha/2}\log^{-3\alpha} n$. Then* wep

$$Z_{\ge k} = (1 + o(1))\left(\frac{d(1-\alpha)}{k(1+\alpha)}\right)^{1/\alpha} n.$$

The proof is a consequence of the following result.

**Theorem 5.6.** *Let $0 < \alpha < 1$, $d \in \mathbb{N}$, $i = i(n) \in [n]$, and let $v_i$ be the vertex whose age rank at time $L$ equals $a_L(v_i) = i = xn$. Let $R$ be the initial rank of $v_i$.*

(i) *Assume that $\sqrt{n}\log^2 n < R < \frac{n}{2} - \sqrt{n}\log^2 n$. Then the expected degree of $v_i$ is given by*

$$\mathbb{E}\deg(v_i, L) = (1 + O(\log^{-1/2} n))\frac{d(1-\alpha)2^\alpha}{1+\alpha}\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)^{\frac{-\alpha}{1-s}}\left(x^{-\alpha} - x\right),$$

*provided $x = o(1)$ or $R/n = o(1)$; otherwise $\mathbb{E}\deg(v_i, L) = O(1)$.*

(ii) *Assume that $R > \frac{n}{2} + \sqrt{n}\log^2 n$. Then $\mathbb{E}\deg(v_i, L) = O(1)$.*

(iii) *Moreover, if* $\mathbb{E}\deg(v_i, L) \geq \log^4 n$, *then* wep

$$\deg(v_i, L) = \mathbb{E}\deg(v_i, L) + O(\sqrt{\mathbb{E}\deg(v_i, L)}\log n),$$

*and if* $\mathbb{E}\deg(v_i, L) < \log^4 n$, *then* wep $\deg(v_i, L) = O(\log^4 n)$.

*Proof.* For (i), let us consider vertices $v_i$ and $v_j$ with age-ranks $a_L(v_i) = i$ and $a_L(v_j) = j$, respectively, and let $i = xn$ and $j = yn$ $(i < j)$. Suppose that $v_i$ obtained an initial rank of $R$. Let $t_i$ and $t_j$ be the times that vertices $v_i$ and $v_j$ were born, respectively. By Theorem 2.1, *wep* $t_i = L-(1+O(\log^{-1/2} n))n\log(1/x)$, $t_j = L-(1+O(\log^{-1/2} n))n\log(1/y)$, and so $t_j - t_i = (1+O(\log^{-1/2} n))n\log(y/x)$. By Theorem 2.3, *wep* $v_i$ had the following rank when $v_j$ was born:

$$r_{t_j}(v_i) = \frac{n}{2}\left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{1}{1-s}}(1 + O(\log^{-1/2} n)).$$

Thus, the contribution to the degree of $v_i$ of vertices born after $v_i$ is the sum of independent indicator variables of the event that a vertex $v_j$ links to $v_i$ in a particular substep of time step $t_j$. The probability of this event is $r_{t_j}(v_i)^{-\alpha}/g_\alpha(n)$, where

$$g_\alpha(n) = \sum_{j=1}^{n} j^{-\alpha} = \frac{n^{1-\alpha}}{1-\alpha} + O(1).$$

Since every vertex has initial degree $d$, the contribution to the degree of $v_i$ by older vertices is $O(d)$. Combining this, we obtain the following expression for the expected degree:

$$\mathbb{E}\deg(v_i, L) = O(d)+(1+O(\log^{-1/2} n))d(1-\alpha)2^\alpha \int_x^1 \left(\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)\left(\frac{y}{x}\right)^{s-1} + 1\right)^{\frac{-\alpha}{1-s}} dy.$$

If $x = \Omega(1)$ and $R/n = \Omega(1)$ then the expected degree is a constant and the degree is smaller than $\log^2 n$ *wep*. Otherwise it simplifies to

$$\begin{aligned}\mathbb{E}\deg(v_i, L) &= (1 + O(\log^{-1/2} n))d(1-\alpha)2^\alpha \left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)^{\frac{-\alpha}{1-s}} x^{-\alpha}\int_x^1 y^\alpha dy \\ &= (1 + O(\log^{-1/2} n))\frac{d(1-\alpha)2^\alpha}{1+\alpha}\left(\left(\frac{2R}{n}\right)^{1-s} - 1\right)^{\frac{-\alpha}{1-s}}\left(x^{-\alpha} - x\right), \qquad (3)\end{aligned}$$

which finishes the proof of (i).

The proof of (ii) is easy: it follows from Theorem 2.3 that *wep* $v_i$ has rank at least $n/2$ during the whole process. Hence

$$\mathbb{E}\deg(v_i, L) \leq n \cdot (1 + o(1))d(1-\alpha)(n/2)^{-\alpha}/n^{1-\alpha} = O(1).$$

Finally, in order to get (iii) it is enough to notice that, provided $\mathbb{E}\deg(v_i, L) = \Omega(\log^4 n)$, *wep* $\deg(v_i, L) = \mathbb{E}\deg(v_i, L)(1 + O(\log^{-1/2} n))$, by the Chernoff bound. □

As we already mentioned, at this point Theorem 5.5 follows easily from the previous theorem. If $i = o(k^{-1/\alpha}n)$, then with probability $1/2 + o(1)$ vertex $v_i$ has degree at least $k$, but $o(k^{-1/\alpha}n)$ vertices of this type turn out to be negligible comparing to the total number of vertices of degree at least $k$. (Note that it follows from part (ii) that with probability $1/2 + o(1)$ vertex $v_i$ gets the initial rank at least $n/2 + \sqrt{n}\log^2 n$ and so has no chance to accumulate a large number of vertices, regardless of how old it is at time $L$.) Since the expected degree of $v_i$ is of order $(\frac{R}{n} \cdot \frac{i}{n})^{-\alpha}$, the probability that a vertex has degree at least $k$ is equal to $O((k^{-1/\alpha}n/i)^s)$. Hence, the expected number of vertices with $i \geq i_0 = Ck^{-1/\alpha}n$ ($C$ is a large constant) is equal to

$$\sum_{i=i_0}^{n} O((k^{-1/\alpha}n/i)^s) = O((k^{-1/\alpha}n)^s) \cdot O(i_0^{1-s}) = O(C^{1-s}k^{-1/\alpha}n).$$

This number, again, turns out to be negligible for $C \to \infty$ (that is, $i_0 \gg k^{-1/\alpha}n$).

Therefore, the only non-trivial probability of having degree at least $k$ is for vertices of age-rank $i = ck^{-1/\alpha}n$ for $c \in (0, \infty)$. The threshold $R_0 = R_0(k, i)$ on the initial rank which causes the vertex to have degree at least $k$ at time $L$ is the following:

$$R_0 = \frac{n}{2}\left(\left(\frac{d(1-\alpha)2^\alpha}{k(1+\alpha)}((i/n)^{-\alpha} - (i/n))\right)^{(1-s)/\alpha} + 1\right)^{1/(1-s)}.$$

To be precise, it follows from Theorem 5.5 that $R < (1 - \log^{-1/3} n)R_0$ implies that $\deg(v_i, L) \geq k$ whereas for $R > (1 + \log^{-1/3} n)R_0$ we get that $\deg(v_i, L) < k$. Hence, the expected number of vertices of degree at least $k$ is

$$\begin{aligned}
(1 + o(1))\sum_{i=1}^{n} F(R_0/n) &= (1 + o(1))\sum_{i=1}^{n}(2R_0/n)^s/2 \\
&= (1/2 + o(1))\sum_{i=1}^{n}\left(\left(\frac{d(1-\alpha)2^\alpha}{k(1+\alpha)}((i/n)^{-\alpha} - (i/n))\right)^{(1-s)/\alpha} + 1\right)^{s/(1-s)} \\
&= (1/2 + o(1))n\int_0^1 \left(\left(\frac{d(1-\alpha)2^\alpha}{k(1+\alpha)}(x^{-\alpha} - x)\right)^{(1-s)/\alpha} + 1\right)^{s/(1-s)} dx \\
&= (1/2 + o(1))n\left(\frac{d(1-\alpha)2^\alpha}{k(1+\alpha)}\right)^{1/\alpha},
\end{aligned}$$

and the result follows from the Chernoff bound.

School of Government and Public Policy, University of Arizona, Tucson, AZ 85721-0027, USA
*E-mail address*: adamdouglashenry@gmail.com

Department of Mathematics, Ryerson University, Toronto, ON, Canada, M5B 2K3
*E-mail address*: pralat@ryerson.ca