# Rank-Based Models of Network Structure and the Discovery of Content

Adam Douglas Henry and Paweł Prałat

[1] Division of Public Administration, West Virginia University, Morgantown, WV 26506-6322, USA
[2] Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310, USA

**Abstract.** Research on self-organizing networks, especially in the context of the Web graph, holds great promise to understand the complexity that underlies many social systems. We argue that models of social network structure should begin to consider how structure arises from the "content" of networks, a term we use to describe attributes of network actors that are independent of their structural position, such as skill, intelligence, or wealth. We propose a rank model of how content (operationalized as attribute rank relative to other individuals) may change amongst agents over time within a stochastic system. We then propose a model of network self-organization based on this rank model. Finally, we demonstrate how one may make inferences about the content of networks when attributes are unobserved, but network structures are readily measured. This approach holds promise to enhance our study of social interactions within the Web graph and in complex social networks in general.

## 1. Why network content matters

Research on the Web graph has been very influential in social science research regarding the structure and function of complex social networks. While the structure and emergence of networks has been a long-standing theme in disciplines such as sociology [1][2], political science [3][4], and economics [5], coupling theoretical models with rigorous models of network self-organization (e.g., [6] and [7]) is still an emerging area of research. Indeed, research on complex networks, especially in the context of the Web graph, has broad applicability in the social sciences and can help to inform methods to unpack the complexity that characterizes many social systems.

At the same time, the social sciences can also contribute to modeling work in mathematics and computer science, since it offers concrete theories about the factors that drive network relationships. Thus, social science theory can help to discipline researchers' focus on particular models that are likely to be more realistic in particular contexts. In research on the Web graph it is important to focus on social drivers of network structure since the Web is, after all, a self-organizing network created and manipulated by human beings. Interactions within the Web

graph are both a direct reflection of human behavior (e.g., when organizations decide to reference one another due to shared interests or resources), but also hold promise as indicators for latent forms of socially-relevant relations such as trust or agreement [8].

In this paper we argue that modeling work on social networks should take seriously the role of network content—meaning the inherent attributes of network actors [9]—in driving network self-organization. Some research has begun to do this by asking, for example, how network structures are influenced by the fitness of actors [10], strategies [11], or the spatial positioning of agents [12]. These types of "content" models of network structure are important supplements to classical modeling approaches that emphasize the importance of structural drivers such as node degree or other measures of centrality. This is because many social science theories are ultimately concerned with attributes of individual actors—why are some powerful and others marginalized, why do political organizations behave the way they do, and how are behaviors, norms, or beliefs learned from others within a network. Thus, to understand complex social networks one must consider structure, but also how structure is dependent upon, and co-evolves with, network content. This will allow researchers to move towards coherent theories of emergent behaviors within social networks.

We contribute to this endeavor in two ways. First, this paper posits a simple, mathematically tractable, yet reasonable model of network self-organization that accounts for the ways in which network content drives network structure. This is a contribution in of itself, and builds heavily upon earlier modeling work in this area by Łuczak, Prałat, Wormald (e.g., [13], [14], [15], [16]) and especially by Prałat and Janssen (e.g., [17] [18]). The model outlined here is a "rank" model where link formation probabilities are based on externally-determined prestige labels relative to other agents in the system; this general approach was first proposed by Fortunato, Flammini and Menczer in [19]. Thus, this paper is concerned with at least preliminary models of network self-organization.

Second, and more importantly, we investigate how this network model may be used to estimate network content—that is, the rank of nodes—based on observed structure alone. This is an important area for research, since in many applications of social network analysis we may know the structure of the network (for example, if networks amongst organizations are measured using hyperlink data), but attributes of actors remain a latent, unobserved variable. Our research builds on prior work to estimate node attributes from observed structure [12], although this research involved a different model and was focused on predicting distances between nodes rather than the attributes of the nodes themselves. We find that making inferences about node ranks is eminently doable, and this research establishes a baseline for methods of statistically inferring node attributes from network structure only. We illustrate the use of this approach through computational simulation, which provides a starting point for future work emphasizing mathematical proof.

The progression of this paper is as follows. We first discuss how the content of a network may be thought of in terms of the rank of vertices—while this is

just one of many possible approaches, it has direct applicability to models of the Web graph and of attendant social network structures as well. The model we consider is stochastic, involving the random entry into and exit from the system of vertices over time, and we present some essential results regarding the shifting of ranks over time using the differential equations method [20]. We then overlay a network model on top of the basic ranking model, which provides a starting point for thinking about how network relations are chosen based on rank. We then present the results of simulations that show how we may determine the content (rank) of vertices based on observations of the structure only. While simulations are used here justify the essential prediction method, future versions of this paper will present rigorous results through mathematical proof. Moreover, due to space limitations, proofs of theorems stated in this paper have been omitted but will be included in a future version.

## 2. A rank model of content

In this section, we formally define a ranking model that reflects the "content" of actors within a hypothetical social system. This model not only specifies the process by which attributes are assigned to individual agents, but also specifies the way in which these attributes shift over time as actors enter or exit the system. Our focus is on modeling systems where the total number of actors is large but fixed (for example, if at each time step an agent 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 agents being added to the system at a faster rate than they are removed.

### 2.1. Model overview

At each time $t$, we have exactly $n$ objects 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 a 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 can be considered, and might lead to different behavior. We first give 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 results are generally about the behavior of ranking functions, where the asymptotics are based on $n$ 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 [21].

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.

### 2.2. Ranking by age

To understand the influence of age, we need to understand the behavior of the age rank function $a_t(v)$ defined before (in short, $a_t(v) - 1$ equals the number of objects in $V_t$ that were born earlier than $v$). We assume (without loss of generality) that $v$ was born at time 0, so $a_0(v) = n$. For $t > 0$, $a_t(v)$ decreases

by one precisely when in time step $t+1$, the object $u$ which is deleted was older than $v$, so $a_t(u) < a_t(v)$. We obtain that

$$\mathbb{E}(a_{t+1}(v) - a_t(v) \mid G_t) = -\frac{a_t(v) - 1}{n-1},$$

conditional on the fact that $v$ is not deleted. To analyze this random variable, we use the differential equations method. Defining a real function $z(x)$ to model the behaviour of $a_{xn}(v)/n$, the above relation implies the following differential equation

$$z'(x) = -z(x)$$

with the initial condition $z(0) = 1$.

The general solution is $z(x) = \exp(-x+C)$, $C \in \mathbb{R}$ and the particular solution is $z(x) = \exp(-x)$. This *suggests* that a random variable $a_t(v)$ should be close to the deterministic function $n\exp(-t/n)$. The following theorem precisely states the conditions under which this holds. This theorem is proved in [17].

**Theorem 1.** *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 $t_f$.*

### 2.3. Randomly chosen initial rank

In this section, 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

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

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 $n/2$). See Figure 1 (a)

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 very 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.

**Case $s = 1$:** 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 [17].

**Lemma 1.** *Suppose that object $v$ obtained an initial rank $R \geq \sqrt{n}\log^2 n$. Then,* wep

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

*to the end of its life.*

**Case $s > 1$:** In this case, 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.** *Suppose that an object $v$ obtained an initial rank*

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

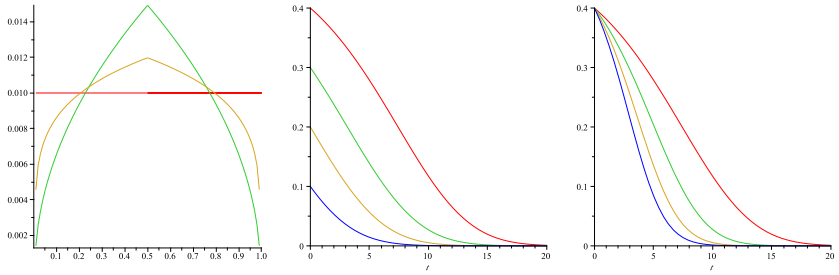
*at time* 0. *Then* wep*, for every $t$ in the range $0 \leq t \leq t_f = \frac{1}{2}n\log n - 2n\log\log n$ conditional upon the object $v$ surviving until time $t$,*

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

*provided*

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

Figure 1 (b),(c) 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 is scaled by $n$.)



(a) $s = 1.0, 1.2, 1.5$ (b) $R = 0.1, 0.2, 0.3, 0.4$ (c) $s = 1.2, 1.3, 1.4, 1.5$

**Fig. 1.** (a) Different distributions: $f(x) = \mathbb{P}\left(x \leq R/n \leq x + \frac{1}{100}\right)$; (b) The behaviour for different initial ranks ($s = 1.2$); (c) The behaviour for different values of $s$ ($R = 0.4$).

## 3. A rank model of network structure

In this section, 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.

In [17], it has been shown that the uniform distribution for the initial rank (that is, the specific case of $s = 1$ in our model) generates *wep* a power-law degree distribution with exponent $1 + 1/\alpha$. Here, we will show that it is also the case for $s > 1$. However, there is a constant factor difference. Let $Z_k$ denote the number of vertices of degree $k$, and $Z_{\geq k} = \sum_{l \geq k} Z_l$.

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

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

The proof is a consequence of the following result.

**Theorem 4.** *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 $a(v_i, L) = i = xn$. Let $R$ be the initial rank of $v_i$, and 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 $\mathbb{E}\deg(v_i, L) = O(1)$. 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 $\deg(v_i, L) = O(\log^4 n)$.

## 4. The discovery of content through structure

While the Web graph is a useful platform for social networks research, the notion that networks self-organize as a function of network content 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 most cases attributes (ranks) of agents embedded in the network are 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 network agents.

Consider, for example, the degree of a given node. Given the model outlined here, this 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, however, 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. Fig 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 neighbours, even if the initial rank is good) but old vertices can still have small degree (because they have an unattractive rank).

As noted above, networks generated according to this rank model are characterized by power-law degree distributions, which is readily observed within simulated networks. Fig 2 presents the cumulative degree distribution on a log-log scale: $y(x)$ is the number of vertices of degree at least $x$.

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

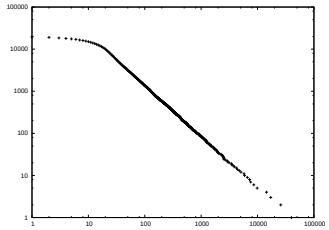
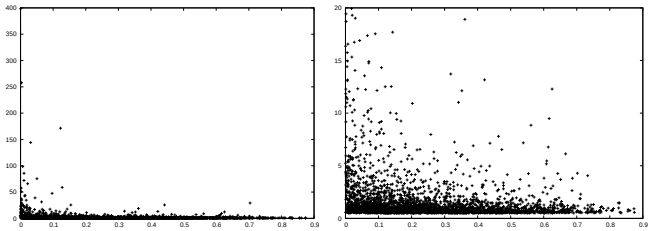


**Fig. 2.** Power-law degree distribution generated by the rank model



(a) (b)

**Fig. 3.** Degree of $v$ vs. the age rank of $v$ (rescaled)

for a vertex $v$ of non-zero degree that is proportional to the average degree of neighbours of $v$:

$$c_2(v) = \frac{\sum_{u \sim v} \deg(u)}{\deg(v)}.$$

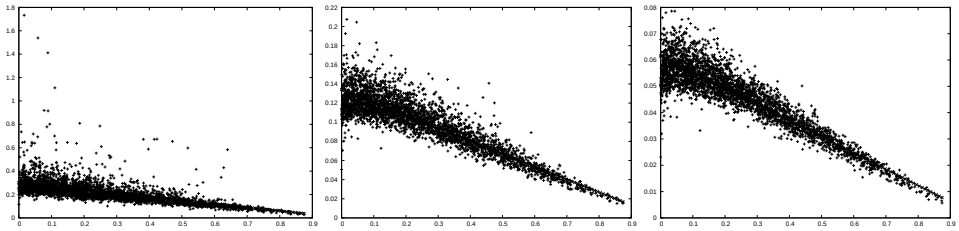
We put $c_2(v) = 0$ if $\deg(v) = 0$. Clearly old nodes have more old neighbours compared to younger nodes. In other words, there is a correlation between the age of $v$ and ages of its neighbours. On the other hand, ranks are generated independently, so a distribution of ranks of the neighbours of $v$ should be similar to the distribution we use in the model. The more neighbours $v$ has, the better correlation we should see. Older vertices should have larger coefficients $c_2(v)$'s. See Fig. 4(a) for the relation for vertices of degree at least $d/2$.

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

$$c_i(v) = \frac{\sum_{u \sim v} c_{i-1}(u)}{\deg(v)},$$

provided that $\deg(v) > 0$; otherwise, $c_i(v) = 0$. Again, in this case older vertices should have larger coefficients and the error should decrease for, say, $i = 3$ and $i = 4$. See Fig. 4(b-c) for the results for $c_3(x)$ and $c_4(x)$.

Even upon casual examination, these scatterplots reveal a strong, nearly linear, relationship between the average degree of neighbors (or higher-order



(a) $c_2(x)$ (b) $c_3(x)$ (c) $c_4(x)$

**Fig. 4.** $c_i(x)$ vs. the age rank of $v$

$c_i$ 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 $c_i$ 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 ($c_2$) 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 $c_3$, and 0.83 for $c_4$.

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 $c_i$ 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, will be sorted out through mathematical proof in a journal version of this paper.

## 5. Conclusion

This paper outlined a model of network self-organization that is driven by the "ranks" of individual agents in terms of an arbitrary attributes that are inherently individual phenomenon, such as wealth, power, beliefs, skills, or any other actor-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 actors 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 network content. 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. The World Wide Web provides an excellent platform for the study of such networks because it yields large-scale, high-quality network data that contains traces of real-world interactions amongst social or political agents.

On the other hand, network content is often exceedingly difficult to observe and can be a limiting factor on our ability to study complex, self-organizing social networks. However, given realistic models of how network structure is driven by content, it seems that we are able to make reasonable inferences regarding the attributes of individuals based on structure only. This research provides a platform for more research, emphasizing analytical proof, the exploits the potential "reversibility" of mathematical models to infer latent, unobserved variables that are crucial to the development of network theory in the social sciences.

## References


1. J.M. Podolny and K.L. Page, *Network Forms of Organization*, Annual Review of Sociology **24** (1998), 57–76.
2. S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.
3. M. Thatcher, *The Development of Policy Network Analysis: From Modest Origins to Overarching Frameworks*, Journal of Theoretical Politics **10(4)** (1998), 389–416.
4. E.O. Laumann and D. Knocke, *The Organizational State: Social Choice in National Policy Domains*, University of Wisconsin Press, Madison, WI, 1987.
5. M.O. Jackson, *Social and Economic Networks*, Princeton University Press, Princeton, NJ, 2008.
6. A. Bonato, *A Course on the Web graph*, American Mathematical Society Graduate Studies Series in Mathematics, Providence, Rhode Island, 2008.
7. F.R.K. Chung, L. Lu, *Complex graphs and networks*, American Mathematical Society, U.S.A., 2004.
8. A.D. Henry, *Measuring Social Networks Using the World Wide Web*, American Association for the Advancement of Science annual meeting, Chicago, Illinois, 2009.

9. K.M. Ingold, *How Network Structures Influence Policy Outputs*, Policy Studies Journal, forthcoming.
10. C. Borgs, J.T. Chayes, C. Daskalakis, and S. Roch, *First to Market is not Everything: an Analysis of Preferential Attachment with Fitness*, Proceedings of the 39rd annual ACM Symposium on the Theory of Computing (STOC), 135–144 (2007).
11. B. Skyrms and R. Pemantle, *A Dynamic Model of Social Network Formation*, Proceedings of the National Academy of Sciences **97(16)** (2000). 9340–9346.
12. 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.
13. T. Łuczak and P. Prałat, *Protean graphs*, Internet Mathematics **3** (2006), 21–40.
14. P. Prałat, *A note on the diameter of protean graphs*, Discrete Mathematics **308** (2008), 3399–3406.
15. P. Pralat, *Connectivity threshold and recovery time in rank-based models for complex networks*, to appear in Discrete Mathematics.
16. P. Prałat and N. Wormald, *Growing protean graphs*, Internet Mathematics **4** (2009), 1–16.
17. J. Janssen and P. Prałat, *Protean graphs with a variety of ranking schemes*, Theoretical Computer Science **410** (2009), 5491–5504.
18. J. Janssen and P. Prałat, *Rank-based attachment leads to power law graphs* SIAM Journal on Discrete Mathematics **24** (2010), 420–440.
19. S. Fortunato, A. Flammini, and F. Menczer, *Scale-free network growth by ranking*, Phys. Rev. Lett. **96**(21): 218701 (2006).
20. 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.
21. 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.
22. S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.
23. 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.