

## Emergence of tempered preferential attachment from optimization

Raissa M. D'Souza, Christian Borgs, Jennifer T. Chayes, Noam Berger, and Robert D. Kleinberg

*PNAS* published online Mar 29, 2007;  
doi:10.1073/pnas.0606779104

**This information is current as of March 2007.**

<b>Supplementary Material</b>	Supplementary material can be found at: <a href="http://www.pnas.org/cgi/content/full/0606779104/DC1">www.pnas.org/cgi/content/full/0606779104/DC1</a>  This article has been cited by other articles: <a href="http://www.pnas.org#otherarticles">www.pnas.org#otherarticles</a>
<b>E-mail Alerts</b>	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or <a href="#">click here</a> .
<b>Rights &amp; Permissions</b>	To reproduce this article in part (figures, tables) or in entirety, see: <a href="http://www.pnas.org/misc/rightperm.shtml">www.pnas.org/misc/rightperm.shtml</a>
<b>Reprints</b>	To order reprints, see: <a href="http://www.pnas.org/misc/reprints.shtml">www.pnas.org/misc/reprints.shtml</a>

Notes:

# Emergence of tempered preferential attachment from optimization

Raissa M. D'Souza<sup>\*†</sup>, Christian Borgs<sup>‡</sup>, Jennifer T. Chayes<sup>‡</sup>, Noam Berger<sup>§</sup>, and Robert D. Kleinberg<sup>¶</sup>

<sup>\*</sup>Department of Mechanical and Aeronautical Engineering, University of California, Davis, CA 95616; <sup>‡</sup>Microsoft Research, Redmond, WA 98052; <sup>§</sup>Department of Mathematics, University of California, Los Angeles, CA 90095; and <sup>¶</sup>Department of Computer Science, Cornell University, Ithaca, NY 14853

Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved February 2, 2007 (received for review August 6, 2006)

**We show how preferential attachment can emerge in an optimization framework, resolving a long-standing theoretical controversy. We also show that the preferential attachment model so obtained has two novel features, saturation and viability, which have natural interpretations in the underlying network and lead to a power-law degree distribution with exponential cutoff. Moreover, we consider a generalized version of this preferential attachment model with independent saturation and viability, leading to a broader class of power laws again with exponential cutoff. We present a collection of empirical observations from social, biological, physical, and technological networks, for which such degree distributions give excellent fits. We suggest that, in general, optimization models that give rise to preferential attachment with saturation and viability effects form a good starting point for the analysis of many networks.**

Preferential attachment (PA) has a long tradition as a model to describe social, biological, and technological networks. It is often assumed to be the underlying mechanism at work when power law distributions are observed in data sets. Although the basic PA model and simple variants have had remarkable success in describing various networks, some authors have suggested that these models are not fundamental (1–3), and/or that they do not have the subtle properties necessary to describe empirical observations in real networks (4–6), and often suggest optimization as an alternate mechanism which gives rise to power laws. In a recent technical work, we introduced a mathematical model which implicitly addresses some of these concerns (7, 8).

The purpose of this manuscript is threefold. First, we show how a form of PA can emerge as a consequence of a natural competition model of networks, to a large extent resolving a long-standing controversy between those who view PA as an axiom and those who argue that power laws should arise from a fundamental optimization framework.

Second, we show that the PA model so obtained has two novel features, which we call saturation and viability. The saturation leads to an eventual exponential cutoff of the PA power-law degree distribution. Therefore, we call our model tempered PA (TPA). Previously, saturation was put directly into the PA model (9) and later, into more general network models (10), to explain a broad range of empirical data that exhibit power-laws with exponential cutoffs. Here, we show that saturation and the corresponding cutoffs can arise naturally from an underlying optimization model. Our model exhibits another property, viability, which reflects the fact that not all attempts to create new nodes are successful. The saturation and the viability together determine the value of the power law. Despite the fact that we have not seen viability discussed in previous works, we believe it to be as fundamental a property of network growth as saturation.

The third purpose of this work is to fit empirical observations from a variety of networks. The simple competition model we propose has only a single parameter, and thus both the saturation and viability are functions of this one parameter. However, because these are, in general, two independent effects, it is natural to assume that saturation and viability would be char-

acterized by two independent parameters. Thus we consider the simplest two-parameter TPA model. We show that this model gives very good fits for a variety of social, biological, and technological networks. In particular, the two-parameter TPA model gives an excellent fit to previously unexplained high-quality data of Internet structure compiled by the Cooperative Association for Internet Data Analysis (CAIDA, [www.caida.org](http://www.caida.org)).

Interestingly, for almost all data sets analyzed, the fits obtained result in the power law portion of the degree distribution having exponent between 1 and 2 (see Table 1). However, the power law in the basic preferential attachment model (11, 12) is 3, and power laws in more general models including mixtures of preferential and uniform attachment achieve exponents greater than 2 (13–17). Our two-parameter TPA model with saturation and viability gives power laws between 1 and 3, and thus provides a much better fit to a large body of empirical observations.

It is not the intent of this work to propose that this simple two-parameter TPA model describes a significant fraction of all networks. Instead, we are merely using this model as a prototype of a class of models based on optimization that exhibit saturation and viability effects in addition to preferential attachment. We suggest that the full class of such models provides a good starting point for the analysis of many networks. This class of models may also provide a good starting point for the numerical simulation of network topologies. TPA, like the standard PA processes, requires only minimal computational resources: to simulate a system of size  $N$  nodes requires computational time linear in  $N$ .

## PA vs. Optimization

PA, also called proportional attachment or cumulative advantage, has a long and illustrious history. It is often assumed to be the underlying mechanism at work when power law distributions are observed in data sets. Although introduced by the mathematician Polya in 1923 (18), the concept of PA was popularized by Zipf and Simon in the 1940s and 1950s, respectively, as an explanation for the power law distribution observed empirically in the population sizes of cities (19), and the distribution of wealth amongst members of a society (20). Also in the 1950s, Mandelbrot proposed an alternate view, that power law distributions could arise as solutions to underlying optimization problems. Mandelbrot's focus was on statistical properties of written language, and on optimizing the amount of information transmitted per symbol (1). A series of exchanges between

---

Author contributions: R.M.D. designed research; R.M.D., C.B., J.T.C., N.B., and R.D.K. performed research; R.M.D. and C.B. analyzed data; and R.M.D., C.B., and J.T.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Abbreviations: PA, preferential attachment; TPA, tempered PA; AS, autonomous system.

<sup>†</sup>To whom correspondence should be addressed. E-mail: [raissa@cse.ucdavis.edu](mailto:raissa@cse.ucdavis.edu).

This article contains supporting information online at [www.pnas.org/cgi/content/full/0606779104/DC1](http://www.pnas.org/cgi/content/full/0606779104/DC1).

© 2007 by The National Academy of Sciences of the USA

**Table 1. Empirical observations of power-law with exponential decay distributions,  $p(x) \sim x^{-B} \exp(-x/C)$**

System with $p(x) \sim x^{-B} \exp(-x/C)$	Exponent, $B$	Cutoff, $C$
Full protein-interaction map of <i>Drosophila</i> (26)	1.20 ± 0.08	27 ± 4
High-confidence protein-interaction map of <i>Drosophila</i> (26)	1.26 ± 0.25	3.9 ± 0.8
Gene-flow/hybridization network of plants as function of spatial distance (27)	0.75	10 <sup>5</sup> m
Earthquake magnitude (28, 29)	1.35–1.7	≈ 10 <sup>21</sup> Nm
Avalanche size of ferromagnetic materials (30)	1.2–1.4	$L^{1.4}$ ( $L$ is sample size)
ArXiv coauthor network (25)	1.3 ± 0.1	53 ± 5
MEDLINE coauthor network (25)	2.1 ± 0.5	≈ 5,800
PNAS paper citation network (10)	0.49	4.21

Mandelbrot and Simon soon ensued, published in the journal *Information and Control*, and launched a still-unresolved public controversy about whether PA or optimization is the more appropriate explanation for emergence of power laws. In recent years, we have seen a flurry of activity on PA, especially in the context of networks, sparked by the model of network growth via PA introduced by Barabási and Albert in 1999 (11). In their model, nodes arrive sequentially and attach to existing nodes with probability proportional to the existing node's current degree. The Barabási–Albert model has been used extensively to describe growth of many classes of networks from social to technological ones. The complementary approach of studying the optimization origins of power laws although dormant for many years, has also recently gained renewed interest (2, 3). For a review of the historical debate, see ref. 21 and references therein; for PA, see ref. 22 and references therein.

Our model resolves the long-standing controversy, unifying both philosophical camps, showing that PA and optimization can be intimately related. Furthermore, it elucidates the potential origins of PA. It is surprising that, until now, despite the decades of study of PA and its widespread uses, no one has investigated or introduced an underlying optimization mechanism that gives rise to it.

### Emergence of Preferential Attachment

**General Framework.** Rather than assuming the PA premise that the “rich get richer,” we start by considering underlying tradeoffs between constrained resources, as in ref. 3. We first consider the framework at a high-level, illustrating its applicability with a specific idealized example, in rough analogy with growth of the Internet. We later introduce the precise model analyzed mathematically. Consider the Internet at the autonomous systems (AS) level, represented as a collection of nodes  $\{j\}$  linked to one another (the link structure is not important for now), further consider its growth via introduction of new ASes (i.e., new service providers). Each newly arriving entity, labeled  $t$ , has to connect to the existing physical network and also establish peering relationships with other existing nodes to route and transmit data packets. Node  $t$  will want to make connections that minimize the monetary startup costs required, while still guaranteeing good network performance for its users (this is the tradeoff). Monetary startup costs are not from laying fiber-optic cable to make the physical connection. There are currently thousands of miles of available cable that have already been laid, but are not yet in use (23) (i.e., “dark fiber”), and the Internet backbone currently contains a large amount of excess capacity (24). Startup costs are instead associated with renting “lit” fiber from existing carriers and negotiating peering agreements. We make a simplified model based on these heuristics: node  $t$  chooses to fully peer with one other node  $j$ , but to reach  $j$  it must make agreements to rent lit fiber from  $n_{ij}$  carriers (i.e.,  $n_{ij}$  transit

intervals). Formally, we can say that the new entity  $t$  wants to minimize the cost function

$$c_t = \min_j [\alpha n_{ij} + h_j]. \quad [1]$$

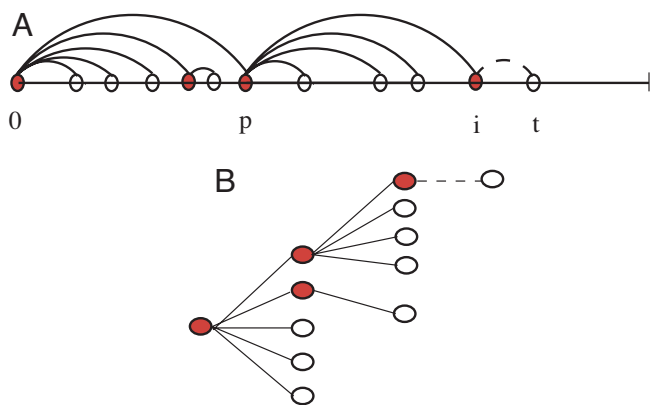
Here,  $n_{ij}$  reflects how many fiber transit intervals a data packet sent from  $t$  to  $j$  has to traverse.  $h_j$  is a measure of network distance from node  $j$  to the core of the network, hence reflects the network delay experienced by node  $j$ . Also,  $\alpha$  is a dimensionless constant which determines the relative weighting between the two metrics, with  $\alpha > 0$ . Node  $t$  wants to connect to  $j$  that has the best network performance, relative to the number of transit intervals traversed. Note that we can consider the general form of Eq. 1,

$$c_t = \min_j [\alpha \Lambda_{ij} + h_j], \quad [2]$$

where  $\Lambda_{ij}$  and  $h_j$  are two distinct metrics in competition, and extend the framework to different contexts. For instance, in a systems biology setting, the tradeoff might involve the ability to evolve quickly versus specialization/efficiency. The protein-interaction network of the system could be used as a basis for both metrics. For instance, ease of evolution could be measured by quantifying the extent of modularity of this network. Efficiency could be measured by quantifying the number of distinct modules involved in crucial control pathways.

**Formal Model Specification.** A precise mathematical model for Eq. 1 is constructed by considering a one-dimensional line, as shown in Fig. 1. We begin at time 0 with a single root node, labeled 0. New nodes arrive one at a time, at random positions along the unit line. Each arriving node  $t$  chooses to connect to the existing node  $j$ , which minimizes Eq. 1. Here,  $n_{ij}$  is the number of nodes already present in the interval between  $t$  and  $j$  (which can be interpreted as the number of transit intervals required to reach  $j$ ), and  $h_j$  (the network performance degradation experienced by  $j$ ) is measured by the hop-count of node  $j$  to the root node. To simplify our analysis, we add the constraint that  $x_j < x_t$ , where  $x_j$  is the physical distance of node  $j$  to the root (i.e., we require that nodes only connect to nodes that are closer to the root). Finally, ties are broken by connecting to the node with the shortest distance between  $x_j$  and  $x_t$ .

It is not hard to show that, given the cost function Eq. 1 and definitions above, node  $t$  will connect either to the node  $i$  directly to its left (which has cost  $c_i = h_i$  because  $n_{ti} = 0$ ), or to the parent of  $i$ , denoted  $p$  (which has cost  $c_p = \alpha n_{tp} + h_i - 1$ , because  $h_p = h_i - 1$ ). Thus, unlike other optimization models (3), ours requires only local information. Comparing  $c_i$  and  $c_p$ , if  $n_{tp} < 1/\alpha$ , node  $t$  attaches to  $p$ ; otherwise, it attaches to  $i$ . Denoting the degree of a node  $p$  as  $d_p$ , we see that when  $d_p < 1/\alpha$ , any new arrival that lands adjacent to  $p$ , or to one of its children, connects to  $p$  (node  $p$ 's attractiveness is proportional to its degree, and its children are



**Fig. 1.** The one-parameter TPA model. (A) An example of the growth process on the line, with  $1/\alpha = 4$ . Node  $t$  arrives and wants to minimize the cost function  $\alpha n_{ij} + h_j$ , where  $n_{ij}$  is the number of existing nodes in the interval between  $t$  and  $j$ , and  $h_j$  is the hop-count of node  $j$  to the root. The minimization can only be achieved by connecting to the adjacent node  $i$  (so that  $n_{ti} = 0$ ), or to the parent of  $i$  (which has  $h_p = h_i - 1$ ). If  $n_{tp} \geq 1/\alpha$  (as is the case illustrated),  $i$  gains the new attachment, otherwise  $p$  does. Thus as soon as there are  $1/\alpha$  children of node  $p$ , node  $p$ 's attractiveness saturates. (B) The equivalent network structure resulting from the growth process on the line.

“infertile”). If  $d_p \geq 1/\alpha$ , any new arrival that lands adjacent to  $p$ , or to one of its closest  $1/\alpha$  children, connects to  $p$ . However, if it lands to the right of a farther away child of  $p$ , that child acquires the attachment (node  $p$ 's attractiveness has saturated at  $1/\alpha$ , and all but the closest  $1/\alpha$  children have become “fertile”). Introducing the notation  $A = \lceil 1/\alpha \rceil$  (i.e., the smallest integer greater than or equal to  $1/\alpha$ ), we can summarize the probability of a fertile node  $p$  acquiring the next attachment as

$$P_{t \rightarrow p}(t) = \min(d_p, A) / \mathcal{N}(t), \quad [3]$$

where  $\mathcal{N}(t)$  is the total number of nodes present at the arrival of node  $t$ . Infertile children do not acquire attachments. A node may be initially infertile, but once enough attachments occur that there are at least  $(A - 1)$  children between the node and its parent  $p$ , that node becomes fertile.

The TPA process thus defined differs from the original PA model in two respects: (i) Each node produces  $A - 1$  infertile children before producing fertile ones. (ii) The preferential attachment of each fertile node saturates at degree  $A$ , beyond which the node does not become any more attractive. We call the second effect “saturation,” indicating that fertile nodes become more attractive only up to a certain threshold. Saturation effects clearly occur in real systems with resource constraints. We call the first effect “viability”: each node is either viable or not viable. In general, one expects a distribution of viability: not all children have an equal ability to procreate, not all social contacts have the same networking capabilities, not all businesses have the same abilities to successfully spinoff subsidiary companies, and not all ASes have the same attractiveness to their peers.

As one might expect, TPA leads to a different degree distribution from the standard PA model. Whereas the Barabási–Albert model produces a degree distribution that is a power law with exponent 3 on all scales, our one-parameter TPA model leads to a power law with exponent 2 up to the threshold  $A$ , and exponential decay for degrees above  $A$ . More explicitly, we find that the probability of observing a node of degree  $d$  is  $p(d) \propto d^{-2}$  for  $d \leq A$ , and  $p(d) \propto \exp(-\kappa d)$  for  $d > A$ , with  $\kappa = \log(1 + 1/A)$  (see refs. 7 and 8 for the derivation of this degree distribution).

Obviously, because our competition model contains only a single parameter, both the viability and the saturation thresholds of the corresponding preferential attachment model are func-

tions of this parameter. From the viewpoint of describing real systems, there is no reason to expect the viability and saturation to be determined by a common parameter, and thus it is natural to consider a system with two independent parameters, a viability threshold  $A_1$  and a saturation threshold  $A_2$ . Indeed, although it seems reasonable that saturation, which often reflects global resource limitations, may sometimes be characterized by a single cutoff, we believe that viability will have a distribution of values reflecting this distribution of viabilities in the population. Thus, in realistic models, we would expect nodes with a distribution of viabilities and saturations, with each distribution being described by one or more parameters. Nevertheless, for simplicity, we next focus on the model with two independent thresholds,  $A_1$  and  $A_2$ , which we will call two-parameter TPA.

Certain limiting cases of the two-parameter TPA model are worth noting. If  $A_1 = 1$  and  $A_2 = \infty$ , it is the standard model of PA. If  $A_1 = 1$  and  $A_2$  is finite, it is the standard model of PA with a cutoff. On the other hand, if  $A_1 = A_2 = 1$ , it is a uniform attachment model.

The mathematical analysis of the two-threshold model has been undertaken in refs. 7 and 8, where we have shown that the probability of observing a node of degree  $d$  is a power law up to  $A_2$ ,  $p(d) \propto d^{-\gamma}$ , with some exponent  $\gamma$  between 1 and 3, whose exact value depends on both  $A_1$  and  $A_2$ . More precisely,  $\gamma$  is decreasing in  $A_1$  and increasing in  $A_2$ , and is obtained as the largest eigenvalue of an explicit finite-dimensional matrix as given in refs. 7 and 8. For  $d > A_2$ , the degree distribution decays exponentially,  $p(d) \propto \exp(-\kappa d)$ , where  $\kappa$  is a simple function of  $\gamma$  and  $A_2$ .

Finally, if one relaxes the constraint of hard thresholds, i.e., if one allows certain distributions of saturations and viabilities, each again characterized by a single parameter, we would expect that the resulting degree distribution is well approximated by a somewhat smoothed version of that above, namely a power law multiplying an exponential decay, henceforth denoted PLED. Below, we show how such a distribution has been used to fit a wide variety of network data. Then, we look specifically at previously unexplained AS Internet data and show how it is fit remarkably well by the  $(A_1, A_2)$  two-parameter TPA distribution, and also by a simple PLED, again with two parameters.

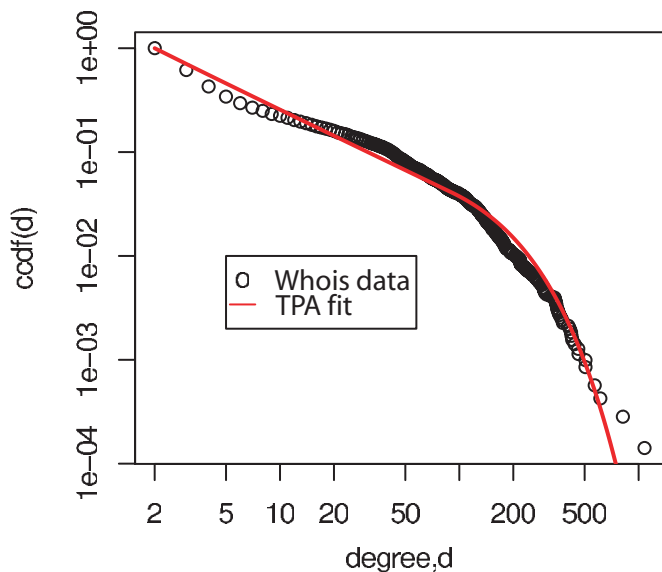
From our perspective, the most interesting open theoretical question is to produce a competition model which is provably equivalent to preferential attachment with independently tunable saturation and viability.

### Correspondence with Data

It has been found repeatedly that PLEDs may provide a much better statistical fit to data than do pure power laws. For instance, consider scientific collaboration and citation networks. The in-degree distribution of the network of paper citations for a 20+year collection of papers published in PNAS (10), and likewise the connectivity distributions of coauthor networks from astrophysics, condensed matter, high energy, and computer science databases (25), all are substantially better fit by power-laws with exponential tails than by pure power laws. There are numerous other such examples from a range of fields, spanning earth sciences, ecology, biology, and physics. For some examples of systems where PLEDs are found to provide the best fit to empirical data, see Table 1.

For power laws to continue indefinitely with no upper cutoff or finite-size effects is physically unrealistic if resources are constrained. To explain the exponential decay observed empirically in a range of networks, saturation similar to that in Eq. 3 has been put in at the axiomatic level (9, 10). For instance, the rate at which a paper gains citations, or the level of activity of an individual scientist, is considered to eventually saturate.

In physical systems, the exponential decay of a power-law often arises due to some finite-length scale in the problem, such

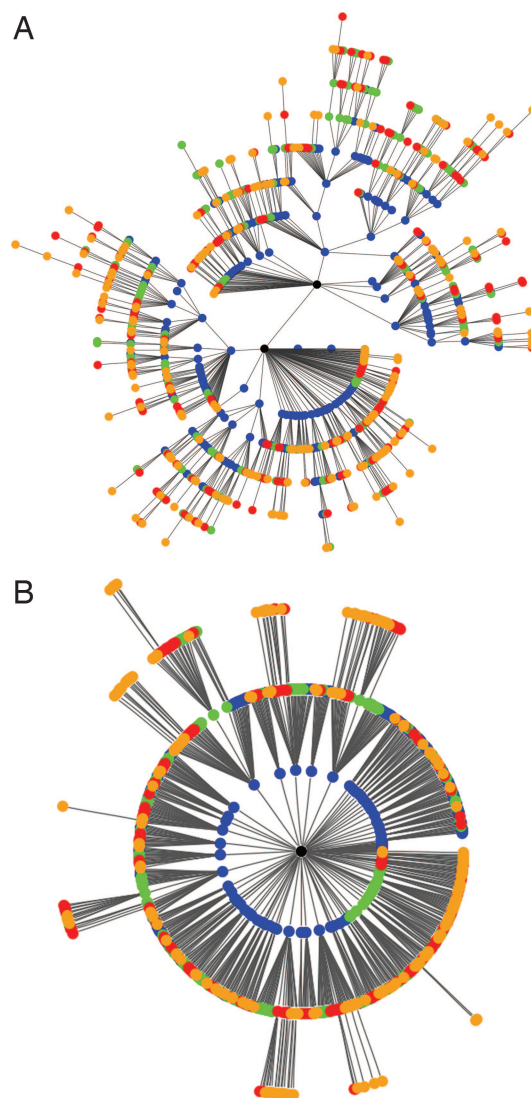


**Fig. 2.** CCDF of the “Whois” AS-level connectivity of the Internet. The circles are data compiled by CAIDA from the RIPE WHOIS database (5, 6). The line is the theoretical TPA CCDF with  $A_1 = 187$  and  $A_2 = 90$  (resulting in the power law portion of the distribution having exponent  $\gamma = 1.83$ ).

as the overall system size. PLEDs are a signature of Barkhausen noise in ferromagnetic materials, describing both the distribution of sizes and the duration of avalanches (30). Likewise, in experiments on explosive fragmentation, the distribution of fragment mass is shown to obey a PLED (31). Such a distribution also provides the best fit to the data of the magnitude of earthquakes spanning a >20-year period, as documented in the Harvard Centroid–Moment Tensor Database (28, 29).

We are beginning to understand the complex networked structure of biological systems. Several years ago, it was recognized that protein interaction networks appear to have heavy-tailed degree distributions (32). New technologies for gene-sequencing now allow for more precise determination of this distribution. For instance, consider the protein-interaction network for *Drosophila*. Giot *et al.* (26) constructed the full, genome-scale, protein interaction map (consisting of 20,439 unique interactions), as well as the high-confidence protein-interaction map, which isolates only the biologically relevant interactions (consisting of 4,780 unique interactions and 4,679 proteins) (26). The distribution of the number of interactions per protein is analyzed for both the full and the high-confidence map. The authors remark that “neither distribution may be adequately fit by a single power-law. Both may be fit, however, by . . .  $\text{Prob}(n) \sim n^{-\alpha} \exp^{-\beta n}$ .” The values for the parameters obtained by a fit to the data (with  $r^2 > 0.98$  in both cases) are given in Table 1.

As seen in Table 1, the exponents  $\gamma$  for the power-law portion tend to be between 1 and 2, although there are a few instances of  $\gamma < 1$ . Recall that the original PA model has  $\gamma = 3$  (11, 12). A variety of modified versions of PA have been considered. A combination of PA plus uniform attachment, yielding  $\gamma > 2$  was proposed by Dorogovtsev *et al.* (13) and rigorously analyzed in refs. 15 and 16. A “copying model” of network growth was proposed and rigorously shown to give  $\gamma > 2$  (17). A similar model, worth noting because it considers attachment either to a random node or its parent, also achieves  $\gamma > 2$  (33). Finally, simulations and heuristic calculations suggest that a modified copying model gives  $\gamma > 1$  (34, 35). Our two-parameter TPA model is the first to rigorously produce the regime  $1 \leq \gamma \leq 3$ .



**Fig. 3.** Comparison of PA and TPA graphs. Both graphs are grown to  $n = 1,000$  nodes. The oldest  $N/4$  nodes are colored blue, the next quarter green, then red, and finally orange. (A) PA graph. (B) TPA with  $A_1 = 17$ ,  $A_2 = 10$ , thus  $\gamma = 1.83$ . Note the effects of aging in TPA; in contrast to PA, very few of the newest nodes attach to the root. Also, due to the minimization of hop-count, the diameter of the TPA graph is much smaller than that of the PA graph. By varying the choice of parameters  $A_1$  and  $A_2$ , it is possible to achieve graphs of appearance intermediate between A and B, or even more extreme than B with respect to the diameter and the number of new nodes attaching to the root.

Furthermore, optimization is a valid scenario in both biological and technological contexts.

### Fitting the Internet AS-Level Data

We are interested in the topology of the Internet at the AS-level. There are a number of sources of such data, obtained using different methodologies, which reveal different features of the topology (for instance, “tangential” versus “radial” links). In particular, we concentrate on the topology view extracted from the RIPE WHOIS database (maintained by network operators), as compiled and characterized by CAIDA. See refs. 5 and 6 for access to background and further information on the validity of this data.

The WHOIS data does not fit the standard paradigm that the AS-level topology follows a power law degree distribution. The two standard views that do support the power law paradigm

(built from direct traceroute sampling or extracted from BGP tables) both rely on traceroute-like sampling (5). Traceroute sampling has been shown to produce biases that can make an underlying simple random graph appear to have a power law degree distribution (36, 37). Though recent heuristic arguments suggest that, despite these biases, it is possible to use traceroute-like probes to distinguish heavy-tailed degree distributions from an exponential distribution, it is not possible to determine the precise form of the distribution (38). In refs. 5 and 6, the RIPE WHOIS data are put forth as a valid view of the AS-level Internet and the lack of agreement with the power-law paradigm discussed at various points. We show that our TPA model with  $A_1 \neq A_2$  provides a remarkable fit to this data. Furthermore, a simple PLED provides a similar fit. For explicit details on the fitting procedure see ref. 39 and [supporting information \(SI\) Appendix](#).

Fig. 2 is a plot of the AS-level connectivity of the Internet constructed from the RIPE WHOIS data. The complementary cumulative distribution function (CCDF),  $\text{ccdf}(x) = 1 - \sum_{d=1}^x p(d)$ , is plotted, along with the CCDF of the best fit provided by the TPA model, obtained for viability  $A_1 = 187$  and saturation  $A_2 = 90$ , corresponding to an exponent of  $\gamma = 1.83$  for the power law portion of the distribution.  $R^2 = 0.97$  for this fit. We do not include vertices of degree one in Fig. 2. As noted in refs. 5 and 6, technically, nodes should be of degree at least two to qualify for an AS number; therefore, the degree  $d = 1$  data are unreliable. See refs. 5 and 6 for further discussion. See ref. 40 for a detailed discussion of the role of constraints and optimization in the structure of the Internet backbone.

For illustrative purposes we generate an image of a corresponding TPA graph. To render a small graph, we have scaled  $A_1$  and  $A_2$  down, while maintaining the value of  $\gamma = 1.83$  fixed. Fig. 3 compares this resulting TPA graph with a PA graph of the same number of nodes.

## Discussion and Further Work

We show that an underlying optimization mechanism can give rise to a form of preferential attachment (namely, the one-parameter TPA model). In addition to elucidating underlying causes of PA, the optimization also provides a mechanism for the emergence of saturation, leading to more realistic distributions of power laws with exponential tails. The most intriguing open question is to construct a simple competition model that gives

rise to tempered preferential attachment with independent saturation and viability. This more general two-parameter TPA model (with viability and saturation) provides a paradigm for achieving degree distributions with power law exponent  $1 \leq \gamma \leq 3$  and eventual exponential decay.

The competition framework is very general and we conjecture that, in addition to the metrics defined in Eq. 1, there are several other metrics which would give rise to a form of TPA in other contexts, such as in a system's biology setting or with regards to the economics of trade relations. The metrics presented herein were inspired by analogy with the AS level Internet.

From a practical viewpoint, the TPA model could be extremely useful due its potential widespread applicability, yet the low overhead required for numerical simulation. As shown above, the optimization model giving rise to the one-parameter TPA model is purely local, unlike typical optimization models (for instance, ref. 3) that require that all alternatives be evaluated before the global optimal can be determined. Numerical simulation of the one- or two-parameter TPA model is very similar to simulation of the standard PA process. Both TPA and PA require computational time linear in  $N$  to generate a network of size  $N$ , but for TPA we must keep track of a small amount of additional information (i.e., which vertices are fertile).

Here, our focus is on degree distributions. Although the degree distribution is an important characteristic of a network, the fine structure is equally important. We have yet to characterize the fine structures resulting via TPA. Our basic model is a tree, and thus has no clustering coefficient; this could be addressed by the modification that each newly arriving node makes  $m > 1$  distinct links to existing nodes.

Finally, similar to PA and the model of ref. 3, our model is an equilibrium one. The values of the controlling parameters ( $A_1$  and  $A_2$ ) are fixed initially and never change. We expect that to make the model more realistic, these thresholds should change as a function of time or with feedback. In addition, we are interested in considering models where viability is described by a distribution of values, rather than herein, where it is described by a single threshold.

We thank Chen-Nee Chuah for her critical reading of our manuscript and the anonymous referees for suggestions. This work benefited from several of the authors attending the Mathematical Sciences Research Institute "Probability, Algorithms and Statistical Physics" program and, in particular, the "Models of Real-World Random Networks" workshop.

- Mandelbrot B (1953) in *Communication Theory*, ed Jackson W (Butterworth, London), pp 486–502.
- Carlson JM, Doyle J (1999) *Phys Rev E* 60:1412–1427.
- Fabrikant A, Koutsoupias E, Papadimitriou CH (2002) *Lect Notes Comp Sci (ICALP 2002)* 2380:110–122.
- Li L, Alderson D, Willinger W, Doyle J (2004) *ACM SIGCOMM Comp Comm Rev* 34:3–14.
- Mahadevan P, Krioukov D, Fomenkov M, Huffaker B, Dimitropoulos X, claffy kc, Vahdat A (2005) [arxiv:cs.NI/0508033](http://arxiv.org/abs/cs.NI/0508033).
- Mahadevan P, Krioukov D, Fomenkov M, Huffaker B, Dimitropoulos X, claffy kc, Vahdat A (2006) *ACM SIGCOMM Comp Comm Rev* 36:17–26.
- Berger N, Borgs C, Chayes JT, D'Souza RM, Kleinberg RD (2004) *Lect Notes Comput Sci (ICALP 2004)* 3142:208–221.
- Berger N, Borgs C, Chayes JT, D'Souza RM, Kleinberg RD (2005) *Combinatorics Probability Comput* 14:697–721.
- Amaral LAN, Scala A, Barthélemy M, Stanley HE (2000) *Proc Natl Acad Sci USA* 97:11149–11152.
- Börner K, Maru JT, Goldstone RL (2004) *Proc Natl Acad Sci USA* 101:5266–5273.
- Barabási A-L, Albert R (1999) *Science* 286:509–512.
- Bollobás B, Riordan O, Spencer J, Tusnády G (2001) *Random Struct Algorithms* 18:279–290.
- Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) *Phys Rev Lett* 85:4633–4636.
- Krapivsky PL, Redner S, Leyvraz F (2000) *Phys Rev Lett* 85:4629–4632.
- Cooper C, Frieze A (2003) *Random Struct Algorithms* 22:311–335.
- Buckley PG, Osthus D (2004) *Discrete Math* 282:53–68.
- Kumar R, Raghavan P, Rajagopalan S, Sivakumar D, Tomkins A, Upfal E (2000) in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS)* (ACM Press, New York), pp 57–65.
- Eggenberger F, Pólya G (1923) *Vorgänge Zeitschrift Agnew Math Mech* 3:279–289.
- Zipf GK (1949) *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge, MA).
- Simon HA (1955) *Biometrika* 42:425–440.
- Mitzenmacher M (2004) *Internet Math* 1:226–251.
- Newman MEJ (2003) *SIAM Rev* 45:167–256.
- Hansen E (Jan 17, 2005) *CNET News*, <http://news.com.com/Google+wants+dark+fiber/2100-1034.3-5537392.html>.
- Nucci A, Taft N, Barakat C, Thiran P (2004) *IEEE J Select Areas Commun* 22:1692–1707.
- Newman MEJ (2001) *Proc Natl Acad Sci USA* 98:404–409.
- Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, et al. (2003) *Science* 302:1727–1736.
- Wilkinson MJ, Elliott LJ, Allaingillaume J, Shaw MW, Norris C, Welters R, Alexander M, Sweet J, Mason DC (2003) *Science* 302:457–459.
- Geller RJ, Jackson DD, Kagen YY, Mulargia F (1997) *Science* 275:1616–1617.
- Kagan YY (1994) *Physica D* 77:160–192.
- de Queiroz SLA, Bahiana M (2001) *Phys Rev E* 64:066127.
- Katsuragi H, Honjo H, Ihara S (2005) *Phys Rev Lett* 95:095503.
- Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási A-L (2000) *Nature* 407:651–654.
- Krapivsky PL, Redner S (2001) *Phys Rev E* 63:066123.
- Bhan A, Galas D, Dewey DG (2002) *Bioinformatics* 18:1486–1493.

35. Chung F, Lu L, Dewey TG, Galas DJ (2003) *J Comp Biol* 10:677–688.
36. Lakhina A, Byers JW, Crovella M, Xie P (2003) in *Proceedings of IEEE INFOCOM 2003* (IEEE, Washington, DC), pp 332–341.
37. Achlioptas D, Clauset A, Kempe D, Moore C (2005) in *Proceedings of the 37th ACM Symposium on Theory of Computing (STOC)* (ACM Press, New York), pp 694–703.
38. Dall'Asta L, Alvarez-Hamelin I, Barrat A, Vazquez A, Vespignani A (2006) *Theor Comput Sci* 355:6–24.
39. D'Souza RM, Borgs C, Chayes JT, Berger N, Kleinberg RD (2007) arXiv:cs.NI/0701198.
40. Doyle JC, Alderson DL, Li L, Low S, Roughan M, Shalunov S, Tanaka R, Willinger W (2005) *Proc Natl Acad Sci USA* 102:14497–14502.