



ELSEVIER

Available online at www.sciencedirect.com

Physica A ■ (■■■■) ■■■-■■■

PHYSICA A

www.elsevier.com/locate/physa

Modeling innovation by a kinetic description of the patent citation system

Gábor Csárdi^{a,b,*}, Katherine J. Strandburg^c, László Zolányi^{a,b}, Jan Tobochnik^d,
Péter Érdi^{a,b}

^aDepartment of Biophysics, KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, 29-33 Konkoly-Thege Road, Budapest, H-1121, Hungary

^bCenter for Complex Systems Studies, Kalamazoo College, 1200 Academy Street, Kalamazoo, MI 49006, USA

^cDePaul University—College of Law, 25 East Jackson Boulevard, Chicago, IL 60604, USA

^dDepartment of Physics and Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, MI 49006, USA

Received 4 April 2006; received in revised form 9 August 2006

Abstract

This paper reports results of a network theory approach to the study of the United States patent system. We model the patent citation network as a discrete time, discrete space stochastic dynamic system. From patent data we extract an attractiveness function, $A(k, l)$, which determines the likelihood that a patent will be cited. $A(k, l)$ shows power law aging and preferential attachment, the exponent of the latter is increasing since 1993, suggesting that patent citations are increasingly concentrated on a relatively small number of patents. In particular, our results appear consistent with an increasing patent “thicket”, in which more and more patents are issued on minor technical advances.

© 2006 Published by Elsevier B.V.

Keywords: Innovation; Patents; Citation network; Preferential attachment; Aging

1. Introduction

Innovation plays a key role in economic development and the patent system is intended (and required by the United States Constitution) to promote innovation. The patent system promotes innovation by giving inventors the power to exclude others from using their inventions during the patent term. The power to exclude is a double-edged sword, however, because it benefits the original inventor, but imposes costs on later innovators seeking to build on past inventions. Thus, the proper design of the patent system is an important matter—and a matter of considerable current debate. See, e.g., Refs. [1–3]. Advances in computer technology and the availability of large patent databases have recently made it possible to study aspects of the patent system quantitatively. To date the empirical analysis of the patent system has been undertaken primarily by economists, sociologists, and legal scholars. See, e.g., Refs. [4–8]. Because patents and the citations between

*Corresponding author.

E-mail address: csardi@rmki.kfki.hu (G. Csárdi).

1 them can be conceptualized as a growing network techniques from statistical physics that have been used in
2 the study of complex networks can be usefully applied to the patent citation network [9,10]. In this paper we
3 present what we believe to be the first results of such a network theory approach to the patent system. We
4 explore the kinetics of patent citation network growth and discuss some possible implications for
5 understanding the patent system.

6 The paper is organized as follows: In Section 2 we provide background on the United States patent system
7 and describe the citation data that is used in this study. In Section 3 we describe a general framework for
8 modeling the kinetics of citation networks, define an “attractiveness function” for the evolving network and
9 introduce a method for extracting this function from the history of the network. In Section 4 we use this
10 approach to analyze the US patent citation network and explore the changes in the kinetics from 1976 to 2000.
11 In Section 5 we discuss some possible implications of our results, and mention directions for future research.

13 2. Patentological background

14 While a similar approach could be applied to many patent systems, including the very important European
15 and Japanese patent systems, we begin our analysis with the United States patent system for which an
16 extensive database of citations has been made available through the work of economists Hall et al. [11].

17 An application for a US Patent is filed in the US Patent and Trademark Office (USPTO). A patent examiner
18 at the USPTO determines whether to grant a patent based on a number of criteria, most important of which
19 for present purposes are the requirements of novelty and non-obviousness with respect to existing technology.
20 Once a patent is issued by the USPTO, it is assigned a unique patent identification number. These numbers are
21 sequential in the order in which the patents were granted.

22 Novelty and non-obviousness are evaluated by comparing the claimed invention to statutorily defined
23 categories of “prior art”, consisting in most cases primarily of prior patents. Patents are legally effective only
24 for a limited term (currently 20 years from the date of application), but remain effective as “prior art”
25 indefinitely. Inventors are required to provide citations to known references that are “material” to
26 patentability, but are not required to search for relevant references (though they or their patent attorneys often
27 do so). During consideration of the application, patent examiners search for additional relevant references.

28 Patent citations include potential prior art that was considered by the examiner. They thus reflect the
29 judgment of patentees, their attorneys, and the USPTO patent examiners as to the prior patents that are most
30 closely related to the invention claimed in an application. Patent citations thus provide, to some
31 approximation, a “map” of the technical relationships between patents in the US patent system. This
32 “map” can be represented by a directed network, where the nodes are the patents and the directed edges the
33 citations. Our research uses a statistical physics approach inspired by studies of other complex networks to
34 attempt to gain insight from this “map”.

35 The patent database we use for the analysis in this paper was created by Hall, Jaffe and Trajtenberg based
36 on data available from the US Patent Office [11]. It is available online at <http://www.nber.org/patents/>. The
37 database contains data from over 6 million patents granted between July 13, 1836 and December 31, 1999 but
38 only reflects the citations made by patents after January 1, 1975: more than 2 million patents and over 16
39 million citations. Citations made by earlier patents are also available from the Patent Office, but not in an
40 electronic format. The Hall, Jaffe and Trajtenberg database also contains additional data about the included
41 patents, which is described in detail in Ref. [11].

43 3. Modeling patent citation networks

45 3.1. Defining the model framework

46 In this section we define the mathematical model framework we will use for studying patent citations. The
47 raw citation data gives us a complete history of citations made and received by each patent. Our goal is to
48 determine whether the evolution of the patent network may be consistently described in terms of variables
49 commonly used in understanding the evolution of complex networks and then to extract the time dependence
50 of the network growth from the detailed history. We assume as an initial matter (an assumption which turns
51

1 out to be consistent with the data) that the evolution of the network may be approximated by a discrete time,
 2 discrete space stochastic dynamic system. Time is measured in patent number units, so that each “time step”
 3 represents the citations made by a single patent. (Though we often “bin” the data from a range of patent
 4 numbers to obtain sufficient statistics for the analysis.) In our model, each patent is described by two variables:

- 5
- 6 (1) k , the number of citations it has received up to the current time step and
 - 7 (2) l , the age of the patent, which is simply the difference between the current time step (as measured in patent
 8 numbers) and the patent number. Because a given patent may cite more than one other patent, several
 9 citations may be made in one time step.

11

12 These two variables define what we call the “attractiveness” of a patent, $A(k, l)$, which determines the
 13 likelihood that the patent will be cited when the next citation is made. In every time step the probability that
 14 an older patent will be cited is proportional to the older patent’s attractiveness multiplied by the number of
 15 citations made in that time step. We find that this simple model gives a very good approximation of the
 16 observed kinetics of the growth of the patent citation network.

17 More formally, the state of the system is described by $k_i(t)$ and $l_i(t)$, ($1 < i < N$), where N is the patent
 18 number of the last patent studied and $k_i(t)$ and $l_i(t)$ are the in-degree and age, respectively, of patent i at the
 19 beginning of time step t . The attractiveness of any node with in-degree k and age l is denoted by $A(k, l)$. $A(k, l)$
 20 is defined such that the probability that node i will be cited by a given citation e in time step t is given by

$$21 \quad P[e \text{ cites node } i] = \frac{A(k_i(t), l_i(t))}{S(t)}, \quad (1)$$

22 where $S(t)$ is the total attractiveness of the system at time step t

$$23 \quad S(t) = \sum_{j=1}^t A(k_j(t), l_j(t)). \quad (2)$$

24 The total probability that node i will be cited in time step t is thus $E(t)A(k_i(t), l_i(t))/S(t)$, where $E(t)$ is the
 25 number of citations made by patent t . $A(k, l)$ and $S(t)$ are defined up to an arbitrary normalization parameter.
 26 To normalize, we arbitrarily define $A(0, 1) = 1$. With this normalization, $S(t)$ is the inverse probability that a
 27 “new” node, with $k = 0$ and $l = 1$, will be cited by a given citation during the next time step.

28 The $A(k, l)$ function determines the evolution of the network. It describes the average citation preferences of
 29 the citing patents (the inventors and patent examiners in reality). In this study, we measure and analyze $A(k, l)$
 30 for the United States patent system during the time period covered by our data. We find first that the
 31 parameterization by k and l consistently describes the average kinetics of the patent citation network. Of
 32 course, underlying patent citations are patentee and patent examiner evaluations of the significance of the
 33 cited patent and the technological relationship between the citing and cited patents that our probabilistic
 34 approach cannot capture. The way in which these “microscopic dynamics” are translated into the average
 35 behavior that we observe remains an open question.

36 Note that this mathematical framework is an assumption about the evolution of the patent network but it is
 37 also a model in the sense that it is a simplified description of the system and that the sensible results we obtain
 38 from our numerical procedure confirm that it is a reasonable model. Similar degree and age based models have
 39 been studied by others [12,13], but the work presented here is different in two respects. First, the degree and
 40 age dependence is a general function, we do not assume any particular form; second, we determine the shape
 41 of $A(k, l)$ from the patent citation network data.

42 In the remaining part of this section we explain our method for measuring the $A(k, l)$ and $S(t)$ functions for
 43 a given network. We believe that this method may be usefully applied to other networks as long as the
 44 necessary data is available.

51

3.2. Estimating the attractiveness function

Let us assume that edges are added to the system one after another in a fixed order; if two edges are added in the same time step (i.e., by the same citing patent), their order is fixed arbitrarily for the measurement. Let e be an edge and let $c_e(k, l)$ be indicator random variables, one for each (e, k, l) triple, ($1 < e < E_{\text{tot}}, k \geq 0, l > 0$), where E_{tot} is the total number of edges in the system. $c_e(k, l)$ is one if and only if edge e cites a (k, l) node (i.e., a node having in-degree k and age l) and zero otherwise. The probability that edge e cites a (k, l) node, i.e., that $c_e(k, l)$ is one, is thus given by

$$P[c_e(k, l) = 1] = \frac{N(t(e), k, l)A(k, l)}{S(t(e))}, \quad (3)$$

where $t(e)$ is the time step during which edge e is added, $S(t(e))$ is the total attractiveness of the system right before adding edge e , and $N(t(e), k, l)$ is the number of (k, l) nodes in the network right before adding edge e . We thus have a formula for $A(k, l)$

$$A(k, l) = \frac{P[c_e(k, l) = 1]S(t(e))}{N(t(e), k, l)}. \quad (4)$$

In Eq. (4) it is easy to determine $N(t(e), k, l)$ for any (e, k, l) , but $S(t(e))$ is unknown. Moreover, we have only a single experiment for $c_e(k, l)$ which is not enough to approximate $P[c_e(k, l) = 1]$ properly. To proceed further, let us define a new set of random variables, each of which is a simple transformation of the corresponding $c_e(k, l)$ variable

$$A_e(k, l) = \frac{c_e(k, l)S(t(e))}{N(t(e), k, l)} \quad \text{if } N(t(e), k, l) > 0. \quad (5)$$

If $N(t(e), k, l) = 0$ then $A_e(k, l)$ is not defined. It is easy to see that the expected value of any $A_e(k, l)$ variable (if defined) is $A(k, l)$; thus we can approximate $A(k, l)$ by

$$\bar{A}(k, l) = \frac{1}{E(k, l)} \sum_{e=1}^{|E_{\text{tot}}|} \frac{\bar{c}_e(k, l)S(t(e))}{N(t(e), k, l)}. \quad (6)$$

Here, $E(k, l)$ is the number of edges for which $N(t(e), k, l) > 0$ for any $t(e)$, and $\bar{c}_e(k, l)$ is the realization of $c_e(k, l)$ in the network being studied.

To calculate this approximation for $A(k, l)$ we need to determine $S(t(e))$, which itself is defined in terms of $A(k, l)$. To determine $A(k, l)$ and $S(t(e))$ self-consistently, we use the following iterative approach:

- (1) First we assume that $S_0(t)$ is constant, and use Eq. (6) to compute $A_0(k, l)$, normalizing the values such that $A_0(0, 1) = 1$.
- (2) Then we calculate $S_1(t)$ for each t based on $A_0(k, l)$ and use this to determine $A_1(k, l)$.
- (3) We repeat this procedure until the difference between $S_n(t)$ and $S_{n+1}(t)$ is smaller than a given small ε for all t .

3.3. The error of the estimation

Similarly to the expected value, the standard deviation of the $A(k, l)$ values can be estimated by taking

$$D^2 \bar{A}(k, l) = \frac{1}{E(k, l) - 1} \sum_{e=1}^{|E_{\text{tot}}|} [\bar{A}_e(k, l) - \bar{A}(k, l)]^2. \quad (7)$$

Here $\bar{A}_e(k, l)$ is the realization of the corresponding $A_e(k, l)$ random variable in the network. Again, the sum contains only cases for which $N(t(e), k, l) > 0$.

1 Based on this estimate, confidence intervals can be calculated for the $\bar{A}(k, l)$ values. The confidence intervals
 for $\bar{A}(k, l)$ were determined by

$$3 \quad \bar{A}(k, l) \pm t_{s/2, N-1} D\bar{A}(k, l) / \sqrt{E(k, l)}, \quad (8)$$

5 where $D\bar{A}(k, l)$ is the standard deviation of $\bar{A}(k, l)$, s is the desired significance level and $E(k, l)$ is the sample
 size. In our case this is the number of edges for which $N_e(t(e,), k, l) > 0$. $t_{s/2, N-1}$ is the upper critical value of the
 7 t distribution with $N - 1$ degrees of freedom. For the patent network we used $s = 0.99$.

9 3.4. Validating the method and software

11 To check the measurement method, we have applied it to various well-known models of growing networks,
 13 such as the Barabási–Albert model [14]. In these tests the method yielded the correct functional form of $A(k, l)$,
 which, for the BA-model, for example, is $A(k, l) = k + 1$. (The role of the +1 here is to assign positive citation
 15 probability to nodes without citations [15].) While these tests gave very good agreement overall they also
 suggested that the method cannot accurately measure the attractiveness of young nodes (small l) with high in-
 17 degree (high k), as these occur very rarely in any finite sample network.

19 4. Results

21 4.1. The attractiveness function

23 The analysis method described in the previous section was applied to the patent citation network and the
 forms of $S(t)$ and $A(k, l)$ were determined. Figs. 1 and 2 show sections of the $A(k, l)$ function. For all the
 25 figures in this paper we have binned the age values into 300 bins, each containing 7172 patents. Ages and times
 are measured in patent number units. Figs. 1 and 2 suggest that, for the patent network, the effects of in-
 27 degree and age can be separated to a good approximation and that $A(k, l)$ can be written approximately in the
 form

$$29 \quad A(k, l) = A_k(k) \cdot A_l(l). \quad (9)$$

31 While this is a reasonable and useful approximation, it is also clear that it is only approximately true. e.g.,
 $A(0, l)$ decays faster than $A(30, l)$, see the second plot in Fig. 1.

33 The measured $A_l(l)$ function for the patent citation network has two major features—a peak at
 approximately 200,000 patent numbers and a slowly decaying tail. (The very large absolute values of $A_l(l)$ are
 35 a result of the normalization, $A(0, 1) = 1$, and are of no independent significance.) The peak at 200,000 patent
 numbers corresponds to a large number of what might be called “ordinary”, relatively short-term citations. In
 37 1998–1999, 200,000 patent numbers corresponded to about 15 months. The tail is best described by a power-
 law decay: $A_l(l) \sim l^{-\beta}$ with $\beta \approx 1.6$. The observation of this power law decay is an important result. It indicates
 39 that while typical citations are relatively short-term, there are a significant number of citations that occur after
 very long delays. Very old patents are cited, suggesting that the temporal reach of some innovations, which
 41 perhaps can be described roughly as “pioneer”, is very long indeed. Moreover, because $A_l(l)$ is approximately
 independent of k —i.e., approximately the same power law decay is observed even for small k —the power law
 43 tail of $A_l(l)$ demonstrates that there is a significant possibility that patents that have gone virtually un-cited for
 long periods of time will re-emerge to garner citations. This slow power law decay of $A_l(l)$ thus suggests the
 45 unpredictability of innovative progress.

The measured $A_k(k)$ function increases monotonically with k , as Fig. 2 suggests. Higher in-degree always
 47 means higher attractiveness. Because the citation probability is proportional to the attractiveness, this means
 that the well-known preferential attachment, or “rich get richer” effect is at work here—the more citations a
 49 patent has received, the more likely it is to receive another. The functional form of $A_k(k)$ is a power law over
 the entire range of k values. $A_k(k) \sim k^\alpha + a$, where $\alpha = 1.2014 \pm 0.0059$ and $a = 1.0235 \pm 0.0313$. We estimated
 51 these parameters using the smaller values of k , for which we have more data. We then checked the results by
 comparing with more extensive fits.

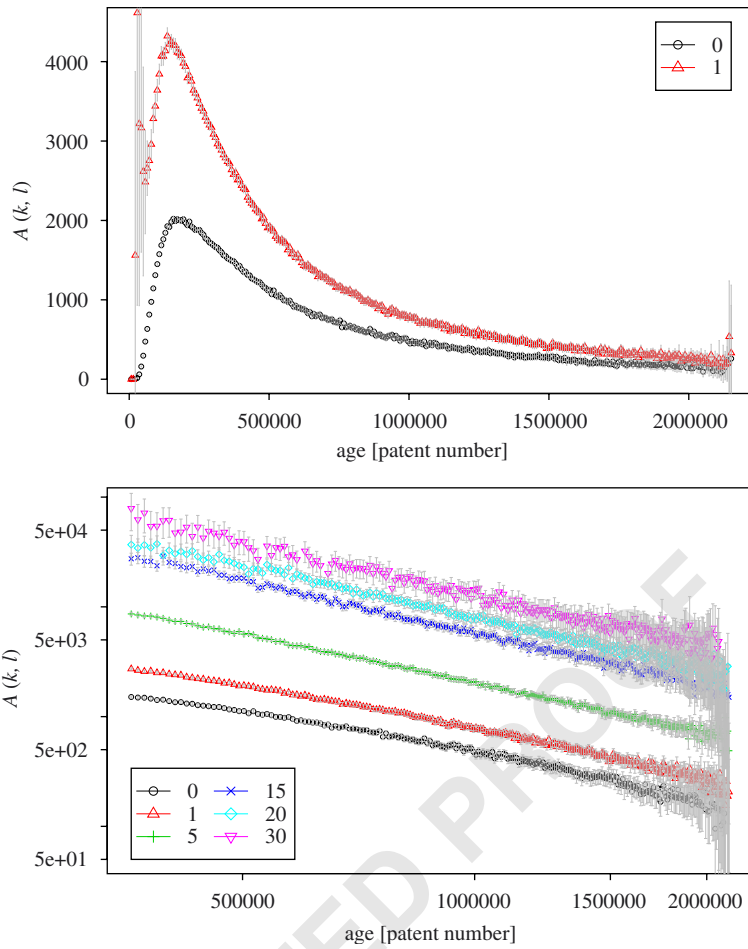


Fig. 1. The measured attractiveness $A(k, l)$ as a function of age l for various fixed values of in-degree, k . The bottom figure shows only the decreasing tail on log-log scales.

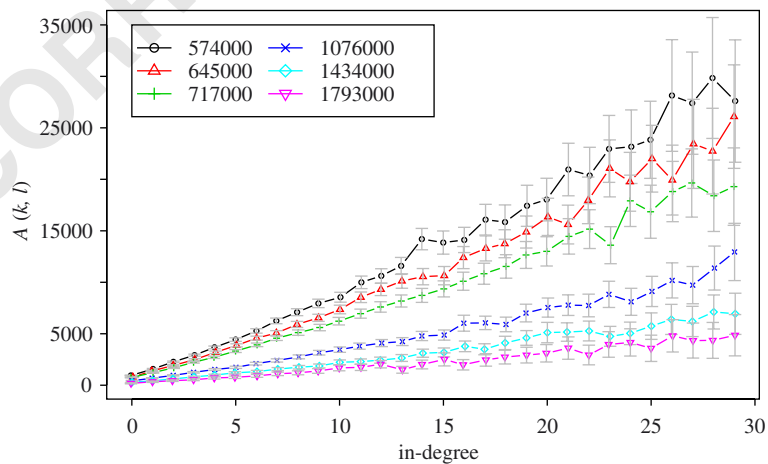


Fig. 2. The measured attractiveness $A(k, l)$ as a function of in-degree, k , for various fixed values of age, l .

1 Preferential attachment and its variations are well studied, see the reviews by Albert and Barabasi [9] and by
 Newman [10]. Linear preferential attachment ($\alpha = 1$) without aging has been shown to result in a degree
 3 distribution (frequency of nodes with degree k) with a power law tail [9,10]. Krapivsky et al. [16] have studied
 nonlinear preferential attachment. In the model they studied there was no aging, $A(k, l) = A_k(k) = k^\alpha + a$. For
 5 $\alpha > 1$, as is observed in the patent citation network, their calculations predict a condensation of node
 connectivity, in the sense that with high probability most of the edges are connected to only a small number of
 7 nodes. More specifically, in their model, if $(m + 1)/m < \alpha < m/(m - 1)$ the number of nodes with more than m
 incoming edges is finite, even in an infinite network. For the patent network $\frac{6}{5} < \alpha < \frac{5}{4}$ suggesting that, if there
 9 were no aging, the number of patents receiving more than five citations would be very small, though those
 patents would account for a large fraction of all of the citations. Aging complicates this picture, of course, and
 11 likely precludes a complete condensation onto a few nodes. However, the fact that the observed preferential
 attachment is super-linear does indicate a tendency toward what might loosely be called “stratification”—
 13 many nodes with very few citations and a few nodes with many citations.

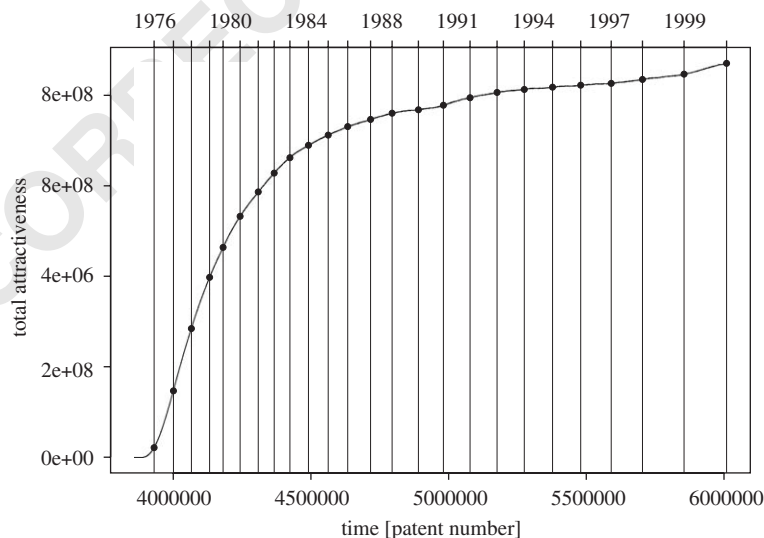
15 4.2. The total attractiveness

17 The total attractiveness function, $S(t)$, (see Fig. 3) of the US patent system increases with time. The initial
 step increase is only a finite size effect and comes from the fact that the citations made by pre-1975 patents are
 19 missing from our database. From about 1984 on, however, $S(t)$ displays a slow but steady increase. One way
 to interpret this increase is that the probability that a patent will be cited by a given citation (which is
 21 proportional to $1/S(t)$) is decreasing as the size of the network increases. The decrease is determined in part by
 the rate at which patents age, which determines the number of patents “available” for citation.

23 The probability that patent i will be cited in a given time step (in other words, by a particular patent rather
 than by a particular citation) is

$$25 \quad P[k_i(t+1) = k_i(t) + 1] = E(t) \frac{A(k_i(t), l_i(t))}{S(t)}. \quad (10)$$

27 The average number of citations made by each patent (and hence, since we measure time in units of patents,
 29 the number of citations made in each time step, $E(t)$), has increased approximately linearly with time in the
 real patent citation network, e.g., it was 4.69 in 1975 and 10.66 in 1999. See Fig. 4. The probability that a new
 31 patent ($k = 0, l = 1$) will be cited by the next patent is thus given by $E(t)/S(t)$, which is shown in Fig. 4. From
 this plot one can see that the increase in the number of citations being made slightly outweighs the increase in
 33



51 Fig. 3. The total attractiveness $S(t)$ of the patent network versus time in units of patent numbers. For ease of reference the time in years is indicated by filled circles and vertical lines.

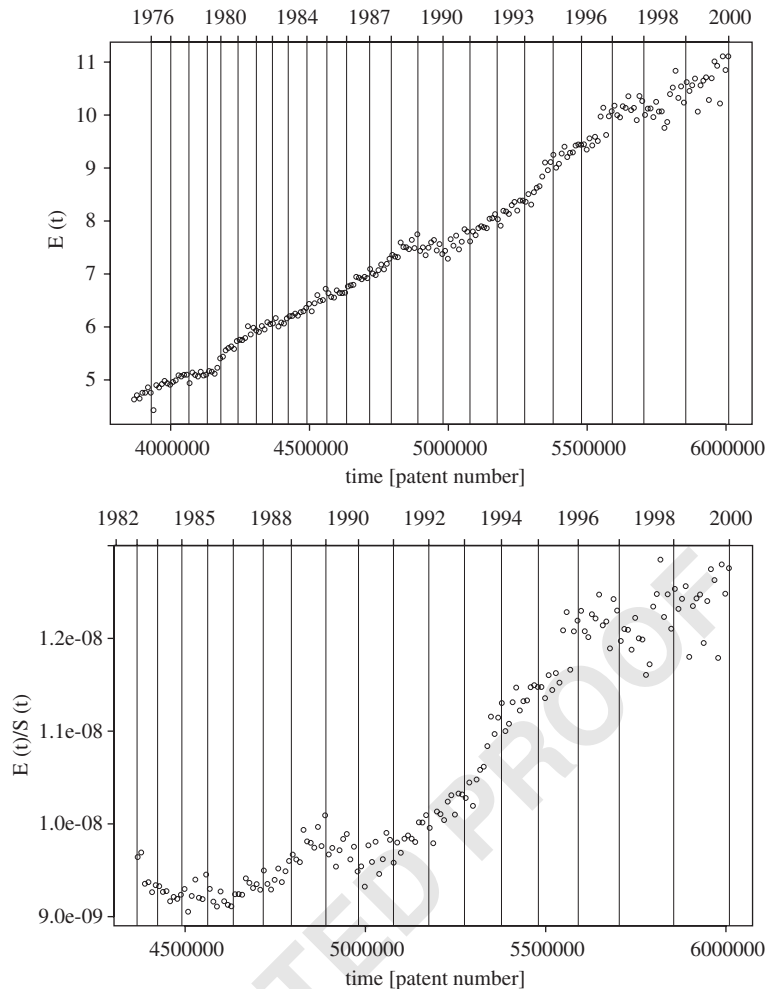


Fig. 4. The top figure shows the number of citations made per patent, $E(t)$, as a function of time in units of patent number. The bottom figure shows $E(t)/S(t)$, corresponding also to the probability that a new patent with $k = 0$, $l = 1$ will be cited, as a function of time in units of patent number.

$S(t)$, so that the probability that a new patent will be cited has increased over time, despite the increasing $S(t)$. Despite the persistent relevance of some old patents indicated by the power law tail in $A_l(l)$, new patents do not get “lost in the crowd” the way we might have predicted from simple models. Instead, patentees and patent examiners have on average increased the number of citations made by each patent to more than compensate for the increasing $S(t)$.

4.3. Change in the patent system dynamics

While it is well-known that there has been a significant increase in the number of US patents granted per year since 1984 [1,17], the underlying reason for this increase is widely disputed. Has there simply been an acceleration of technological development in the last twenty years or has there been a more fundamental change in the patent system, perhaps, as many have suggested, as a result of increased leniency in the legal standard for obtaining a patent [1]? A complete answer to this question is far beyond the scope of the present investigation. However, our kinetic model does permit us to ask whether there has been any deep change in the growth kinetics of the patent citation network. Because we measure time in units of patent number, a mere

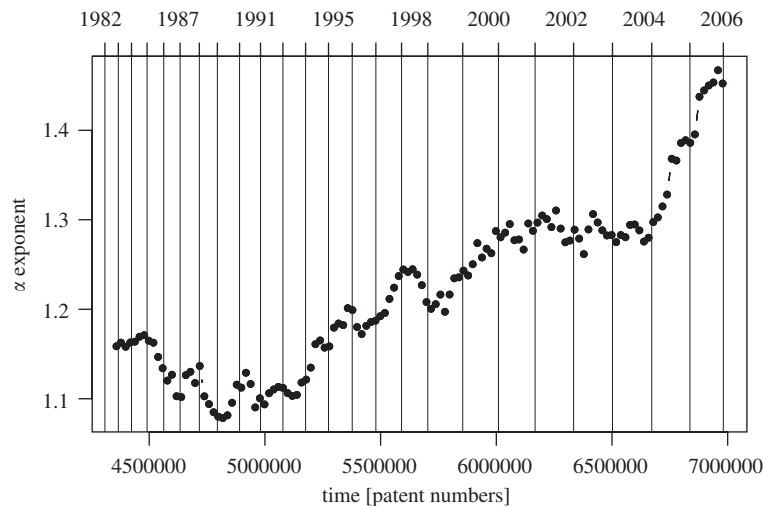


Fig. 5. The measured value of α as a function of time, measured as described in the text. The time in years is indicated by the grey vertical lines.

acceleration of technological progress should leave $A(k, l)$ unchanged in patent number “time”. A change in $A(k, l)$ indicates some other source of change.

Thus far, we have assumed a time-independent $A(k, l)$, which is reasonably consistent with our observations. In this section, we relax this assumption to ask the more subtle question of whether there has been a change in patent system kinetics over and above the acceleration that is already reflected in our choice of time units. Specifically, we allow α and β to vary with time and ask whether there has been a significant change in these parameters between 1980 and 2000.

To answer this question we measured the parameters of the system as functions of time. To perform the fits, we averaged over a 500,000-patent sliding time window and calculated the parameters after every 100,000 patents. The measured α parameters are plotted in Fig. 5. There is a significant variation over time. The time dependence of the important β parameter was also explored, but no significant time dependence was observed to within the statistical errors.

The plot for the α parameter shows that there are two regimes. In the first regime, prior to about 1991, α is decreasing slightly with time, while in the second, starting around 1993, there is a significant increase. As noted earlier, the α parameter has some very important consequences for the growth of the network: the higher α , the more “condensed” or “stratified” the network will be. The increasing α in the patent citation network indicates increasing stratification—a smaller and smaller fraction of the patents are receiving a larger and larger fraction of the citations. This change is not simply a result of accelerating numbers of patents being granted, but suggests a more fundamental change in the distribution of patents that are being issued.

5. Conclusions

We have presented a stochastic kinetic model for patent citation networks. Though a complex process underlies each decision by a patent applicant or examiner to cite a particular patent, the average citation behavior takes a surprisingly simple form. The citation probability can be approximated quite well by the ratio of an “attractiveness function”, $A(k, l)$, which depends on the in-degree, k , and age in patent numbers, l , of the cited patent, and a time-dependent normalization factor, $S(t)$, which is independent of k and l .

We introduced a method to extract the $A(k, l)$ and $S(t)$ functions of a growing network from a specification of the connection history. We applied this technique to the patent citation network and, though no assumptions were made as to the functional form of $A(k, l)$, the measured $A(k, l)$ function was well described by two approximately separable processes: preferential attachment as a function of in-degree, k , and power law age dependence. The interplay of these two processes, along with a growth in the number of citations

1 made by each patent, governs the emerging structure of the network. Particularly noteworthy are our finding
2 that the preferential attachment is super-linear, implying that patents are highly stratified in “citability”, and
3 our finding of a power law tail in the age dependence even for small k , indicating not only that some patents
4 remain important for very long times, but also that even “dormant” patents can re-emerge as important after
5 long delays.

6 We also used our technique to investigate the time dependence of the growth kinetics of the patent citation
7 network. Overall, we find that the increasing number of patents issued has been matched by increasing
8 citations made by each patent, so that the chance that a new patent will be cited in the next time period has
9 increased over time. This result suggests that on average patents are not becoming less “citable”. However, we
10 also find that there has been a change in the underlying growth kinetics since approximately 1993. Since that
11 time, preferential attachment in the patent system has become increasingly strong, indicating that patents are
12 more and more stratified, with fewer and fewer of the patents receiving more and more of the citations. A few
13 very important, perhaps “pioneer”, patents seem to dominate the citations. This trend may be consistent with
14 fears of an increasing patent “thicket”, in which more and more patents are issued on minor technical
15 advances in any given area. These technically dense patents must be cited by patents that build upon or
16 distinguish them directly, thus requiring that more citations be made, but few of them will be of sufficient
17 significance to merit citation by any but the most closely related patents. These observations are consistent
18 with recent suggestions that patent quality is decreasing as a result of insufficient standards of non-
19 obviousness. See, for discussion of these issues, e.g., Refs. [1–3] and references therein.

20 Note that while we used the in-degree and age of the patents to describe and measure the dynamics of the
21 patent network, the methodology presented here is general and can be applied to relate *any* node property to
22 the governing dynamics based on the *data driven* measurement.

23 In this paper we do not develop a microscopic theory of why the attractiveness function factorizes and why
24 it has the measured form. Future work will be done in this direction.

25 This work is only the beginning. In this paper we do not develop a microscopic theory of why the
26 attractiveness function factorizes and why it has the measured form. Future work will seek to address these
27 questions. Moreover, there are many further applications of network analysis to the patent citation network
28 that are likely to bear fruit. It will be possible, for example, to compare the structural and kinetic behavior of
29 the network for patents in different technological areas, to investigate the degree of relatedness between
30 patents in seemingly disparate technologies, and to explore more detailed structural indicators, such as
31 clustering coefficients and correlation functions. Also, it may be possible to compare the growth of patent
32 systems internationally, perhaps providing a means to distinguish between the effects of global technological
33 change and those of nation-specific legal changes. Finally, it will be interesting to compare the behavior of the
34 patent citation network with that of other networks (such as the scientific journal citations discussed in Ref.
35 [18]) to gain deeper insight into the behavior of complex networks in general.

37 Acknowledgments

38 This work was funded in part by the National Science Foundation and the Hungarian Academy of Sciences
39 under grant INT-0332075, the EU FP6 Programme under Grant numbers IST-4-027173-STP and IST-4-
40 027819-IP and by the Henry R. Luce Foundation. K. S.’s research is supported by the DePaul University
41 College of Law.

45 References

- 46 [1] A.B. Jaffe, J. Lerner, Innovation and its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and
47 What to Do About It, Princeton University Press, Englewood Cliffs, NJ, 2004.
48 [2] Federal Trade Commission, To promote innovation: the proper balance of competition and patent law and policy, Report (October
49 2003).
50 [3] S.A. Merrill, R.C. Levin, M.B. Myers (Eds.), A Patent System for the 21st Century, National Research Council of the National
51 Academies, National Academy Press, Washington, DC, 2004.

- 1 [4] A. Jaffe, M. Trajtenberg, *Patents, Citations and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA, 2003.
- 3 [5] J.R. Allison, M.A. Lemley, K.A. Moore, R.D. Trunkey, *Valuable patents*, *Georgetown Law J.* 92 (2004) 435.
- [6] K.A. Moore, *Worthless patents*, *Berkeley Technol. Law J.* 20.
- 5 [7] J.M. Podolny, T.E. Stuart, *A role-based ecology of technological change*, *Am. J. Sociol.* 100 (1995) 1224.
- [8] I. von Wartburg, T. Teichert, K. Rost, *Inventive progress measured by multi-stage patent citation analysis*, *Res. Policy* 34 (2005) 1591.
- 7 [9] R. Albert, A.-L. Barabási, *Statistical mechanics of complex networks*, *Rev. Mod. Phys.* 74 (2002) 47.
- [10] M.E.J. Newman, *The structure and function of complex networks*, *SIAM Rev.* 45 (2003) 167–256.
- 9 [11] B.H. Hall, A.B. Jaffe, M. Trajtenberg, *The NBER patent citation data file: lessons, insights and methodological tools*, in: A.B. Jaffe, M. Trajtenberg (Eds.), *Patents, Citations, and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA, 2003.
- 11 [12] S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of networks with aging of sites*, *Phys. Rev. E* 62 (2) (2000) 1842–1845.
- [13] H. Zhu, X. Wang, J.-Y. Zhu, *Effect of aging on network structure*, *Phys. Rev. E* 68 (2003) 056121.
- 13 [14] A.-L. Barabási, R. Albert, H. Jeong, *Mean-field theory for scale-free random networks*, *Physica A* 272 (1999) 173–187 cond-mat/9907068.
- [15] R. Albert, A.-L. Barabási, *Topology of evolving networks: local events and universality*, *Phys. Rev. Lett.* 85 (2000) 5234.
- 15 [16] P. Krapivsky, S. Redner, F. Leyvraz, *Connectivity of growing random networks*, *Phys. Rev. Lett.* 85 (2000) 4629–4632.
- [17] B.H. Hall, *Exploring the patent explosion*, *J. Technol. Transfer* 30 (2005) 35–48.
- 17 [18] S. Redner, *Citation statistics from 110 years of physical review*, *Phys. Today* 58 (2005) 49.

UNCORRECTED PROOF