Topology and evolution of technology innovation networks

Sergi Valverde,[1] Ricard V. Solé,[1,2] Mark A. Bedau,[3,4] and Norman Packard[2,4]
[1]*ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr. Aiguader 80, 08003 Barcelona, Spain*
[2]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*
[3]*Reed College, 3203 SE Woodstock Blvd., Portland, Oregon 97202, USA*
[4]*ProtoLife, Parco Vega, Via della Liberta 12, Marghera 30175, Venice, Italy*
(Received 2 December 2006; revised manuscript received 20 July 2007; published 28 November 2007)

The web of relations linking technological innovation can be fairly described in terms of patent citations. The resulting patent citation network provides a picture of the large-scale organization of innovations and its time evolution. Here we study the patterns of change of patents registered by the U.S. Patent and Trademark Office. We show that the scaling behavior exhibited by this network is consistent with a preferential attachment mechanism together with a Weibull-shaped aging term. Such an attachment kernel is shared by scientific citation networks, thus indicating a universal type of mechanism linking ideas and designs and their evolution. The implications for evolutionary theory of innovation are discussed.

DOI: 10.1103/PhysRevE.76.056118 PACS number(s): 89.75.Hc, 87.23.Kg

I. INTRODUCTION

Innovation takes place both in nature and technology [1]. Either through symbiosis [2], tinkering [3,4], or design [5] new functional structures and artifacts are obtained. Such new entities often result from the combination of predefined designs or building blocks, although a completely new solution can also emerge. This is the case, for example, in the replacement of vacuum tube technology by semiconductors. However, the majority of technological (and evolutionary) changes take place gradually [6]. Such steady transformation of designs is largely based on an extensive combination and refinement of existing inventions.

A surrogate of the ways in which innovations take place in time is provided by patent files. Patents are well-defined objects introducing a novel design, method, or solution for a given problem or set of problems and they can be analyzed in full detail [7]. Additionally, they indicate what previous novelties have been required to build the new one. In order to gain insight into the global organization of the patterns of innovation and their evolution in technology, here we study a very large database including all U.S. Patent and Trademark Office (USPTO) patents from 1975 to 2005.

As it occurs with the fossil record for biological evolution, the record of patents through time provides us with the opportunity of seeing how new inventions emerge and how they relate to previous ones. A given patent will typically require new solutions and also previously achieved results. Looking at how patents link to each other is the simplest way of having a large scale picture of the patterns and processes associated with the collective dynamics of innovation as it unfolds [8,9]. In this way patents are related to each other through a network of patent-patent associations. Such a network can be treated from a statistical physics perspective, looking for statistical patterns and their origins [10].

Many interesting questions can be formulated in relation to this large-scale picture of invention: What is the global organization of interactions among innovations? Is this a repeatable pattern? How are similar classes of innovations related among them? Do these patterns respond to history-dependent rules or are they instead describable by means of simple models? These questions are addressed here and it is shown that a standard statistical physics approach provides a good picture of how these webs emerge.

The paper is organized as follows. In Sec. II the data set analyzed is presented. In addition, we discuss the modular pattern exhibited by a small subgraph describing the patents in computed tomography. In Sec. III the topological trends exhibited by the full patent citation network are discussed under the light of a model of graph growth with aging (Sec. IV). In Sec. V our basic results are summarized and their evolutionary implications outlined.

II. PATENT CITATION NETWORKS

Previous studies have measured the value of an innovation by means of the analysis of patent citations, i.e., the rate of receiving new citations. However, innovation is an elusive notion that is difficult to measure properly and existing measures provide limited insight [8]. It is a difficult task to find useful indicators for the value of innovations. In this context, we introduce patent citation networks as an appropriate approach to the global analysis of the process of technological innovation. Recent work in complex networks provides several models that describe or reproduce structural features of natural and artificial evolving systems. Here, we will show how innovation can be described as a process of network growth following some specific rules. In particular, our model provides a rigorous statistical test to assess the balance between patent importance and patent age, i.e., Price's "immediacy factor" [8].

The set of patents and their citations describes a (so-called) patent citation network $G=(V,L)$. The patent network belongs to the general class of citation networks, which also includes scientific citation networks. Here, nodes $v_i \in G$ represent individual patents and the directed link (v_i, v_j) indicates that patent v_i is a descendent of patent v_j. In order to illustrate the power of the network approach, we have reanalyzed the evolution of patents in computed tomography (CT) [11]. G. N. Hounsfield, a senior research scientist at EMI Limited in England, developed CT in 1972. This technology, commonly referred to as CAT scan, uses x rays to produce high-resolution sectional images of the human body. Figure 1

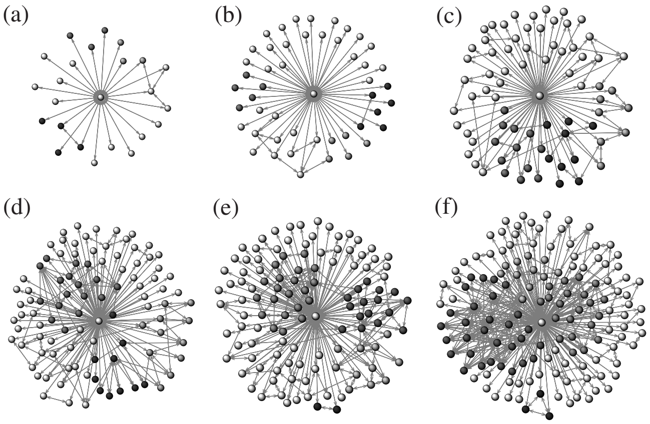

FIG. 1. From (a) to (f), evolution of the computed tomography (CT) patent network. This is a subgraph from the full USPTO patent network. The hub in the center corresponds to the precursor invention by Hounsfield [12]. We have detected $m=5$ different modules in the final CT network (f) using a specific algorithm of community detection (see [13]). Modules are indicated with different shades of gray (see Table I for a functional description). We can trace the evolution of modules by shading nodes according to the partition found in (f).

shows the time evolution for the computed tomography (CT) patent network. This data set comprises all the patents related to CT from 1973 to 2004. The CT network in2004 has $N=141$ nodes (patents) and $L=344$ directed links (citations).

This example illustrates some common features displayed by patent networks. Figure 1 indicates that some patents receive many more citations than others. In particular, the hub at the center corresponds to the very first CT patent. Interestingly, the network level analysis reveals some patterns that cannot be easily recovered by just looking at individual patents. Some groups of patents are more densely connected among them than with other patents. This pattern of modular structure seems to be a general trend in patent networks, thus indicating a nontrivial organization of relations between inventions.

For instance, in Fig. 1 we can appreciate the modular organization of the CT patents. The CT network is directed, but we will consider it as undirected when assessing its modularity. Here we have used Clauset *et al.*'s algorithm [13] to detect community structure in networks. The method considers a decomposition of the graph Ω into a set of μ subgraphs $\{C_r=(V_r,E_r)\}$ with $r=1,\ldots,\mu$. This defines a partition $\mathcal{C}$ such that $E=\cup R_r$ and $E_v \cap E_w=\varnothing$ for all pairs of subsets. Using the adjacency matrix $A=(a_{ij})$, the fraction of edges (for a given partition) that fall within subsets of $\mathcal{C}$ will be

$$f(\mathcal{C}) = \frac{\sum_{i,j} a_{ij}\delta(C_i,C_j)}{\sum_{i,j} a_{ij}}, \tag{1}$$

where $\delta(a,b)=1$ if $a=b$ and zero otherwise. Using $m=\Sigma_{ij}a_{ij}/2$ we can also write

TABLE I. Modular organization of the network shown in Fig. 1(f). The topological structure of the modular graph is well mapped with functionally meaningful classes of inventions. Here we list the five observed modules (as detected with the community detection algorithm), their size, and the characteristic keywords that appear in the patent description.

Module	Size	Common keywords
1	$N_1=20$	Transmission, 2D/3D image, tomography applications
2	$N_2=77$	Slice, CT scanner, radiography, compensating detector motion
3	$N_3=3$	Signal processing, radiation, radiology, reconstruction
4	$N_4=11$	Radiation, body, image, x-ray scanning
5	$N_5=30$	Tomographic, imaging system, diagnostic, examining

$$f(\mathcal{C}) = \frac{1}{2m}\sum_{i,j} a_{ij}\delta(C_i,C_j). \tag{2}$$

In order to define an appropriate modularity index, the previous measure needs to be compared with the expectation from a randomly wired graph with an identical number of nodes and links. Let us indicate as k_i the degree of v_i, which is obtained from the adjacency matrix as $k_i=\Sigma_j a_{ij}$.

The expected probability of having a link connecting two arbitrary nodes v_i and v_j will be simply $k_ik_j/2m$ and thus we can define modularity $\mathcal{Q}$ in terms of the average difference between the observed and the expected value of f, namely

$$\mathcal{Q} = \frac{1}{2m}\sum_{i,j}\left[a_{ij} - \frac{k_ik_j}{2m}\right]\delta(C_i,C_j), \tag{3}$$

which is properly normalized between zero (random network) and one (a single module is present).

The modularity of a network will be defined as the maximum $Q=\max\{\mathcal{Q}\}$ as evaluated by the search algorithm [13]. The size of the best partition μ defines the (potential) number of modules. For the CT patent network, we have found $m=5$ modules corresponding to the maximum modularity $Q=0.4136$ (see Fig. 1 where patents belonging to the same module have the same shade of gray). Although we have not explored this problem in detail, close inspection of the networks shown in Fig. 1 reveals that the modular structure seems to correlate well with shared functional traits (see Table I). In this context, this particular example seems to overcome the limitations of community detection methods that can jeopardize the relevance of modular partitions [14] but further analysis will be required to check the robustness of our finding.

Beyond specific patterns of patent evolution, here we aim to detect universal trends in the global evolution of the whole U.S. patent system. To this goal we have studied the full database of patent citations created by Hall, Jaffe, and Trajtenberg [8] (notice this database includes the CT network analyzed above). This database contains the citations made

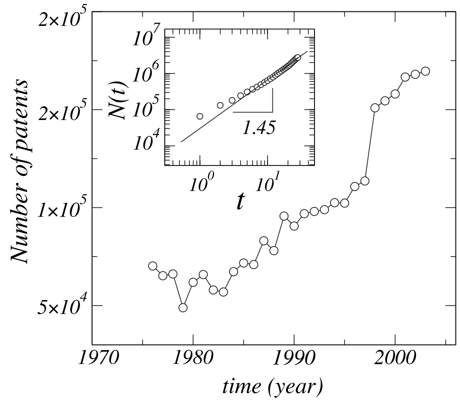

FIG. 2. Time evolution of the number of patents $N(t)$ in the USPTO data set from 1973 to 2004. Inset: Cumulative number of patents on a log-log scale, showing a scaling law $N(t) \sim t^{\theta}$ with $\theta \approx 1.45$.

by patents granted after January 1, 1975. The patent citation network (PCN) analyzed here has $N=2\,801\,167$ nodes and $L=18\,053\,661$ links. Its time evolution from 1976 to 2005 is shown in Fig. 2. The number of patents at a given time t scales as a power law:

$$N(t) \sim t^{\theta} \tag{4}$$

with an exponent $\theta=1.45\pm0.06$.

Some recent papers have explored the patent citation data sets at different levels, including a graph theoretical approach on a large scale [15] or involving a more specific case study, such as fuel cell research [16]. Here we will show that the statistical features of this network can be explained by using an appropriate attachment kernel describing how successful patents become more linked and how this preferential attachment decays with age. A related study [15] found that citation probability can be approximated by the ratio of an "attractiveness function" $A(k,l)$ that depends on both the in-degree k and patent age l but no specific functional form of $A(k,l)$ was provided. Here we estimate this functional form, which suggests there is a deep link between patent citation networks and scientific citation networks (see below).

III. GLOBAL TOPOLOGY OF THE USPTO PATENT NETWORK

Citations are often interpreted as indicators of innovation size or economic value [8]. The distribution of innovation size (defined as the number of citations to a patent) is skewed [9,17,18]. However, there is an ongoing discussion about the particular nature of this distribution. In particular, there is no general agreement on whether it follows a log-normal or Pareto distribution [18,19]. However, several indicators are clearly consistent with a power-law tail. Here we report similar features in the in-degree distribution studied here (see below).

The in-degree distribution $P_i(k)$ is equivalent to the so-called distribution of the number of patent citations. Figure 3(a) shows the in-degree distribution for the USPTO patent citation network in 2004. Notice that $P_i(k)$ is neither exponential nor a simple power law. Instead, we have found that an extended power-law form, described by a Zipf-Mandelbrot function, best describes the in-degree distribution:

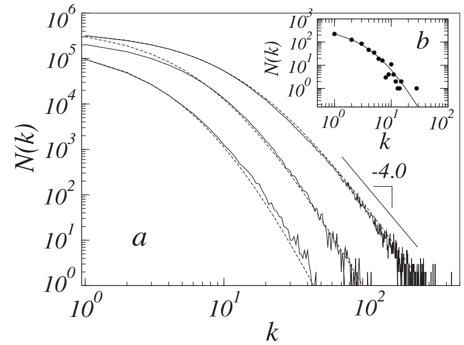

FIG. 3. (a) In-degree distribution for the patent citation network follows an extended power-law shape, $P_i(k) \sim (k+k_0)^{-\gamma}$. Three distributions are displayed for three different time windows, namely 1984 (leftmost), 1992 (center), and 2002 (rightmost). (b) The in-degree distribution (filled circles) for the subset of patents displayed in Fig. 1(g) associated to patents on microchip technology. The corresponding extended scaling fit is also shown (continuous line) using $k_0=6.80\pm0.16$ and $\gamma=4.73$.

$$P_i(k) \sim (k+k_0)^{-\gamma}, \tag{5}$$

where $k_0=19.46\pm0.22$ and $\gamma=4.55\pm0.04$. This extended power law typically deviates from a straight line in a log-log plot when k_0 is comparable with the largest degree. It reduces to a power law when $k \gg k_0$ and it fits an exponential form for $k \ll k_0$. The extended power-law distribution has been related to a mixed attachment mechanism [20,21]. However, here we will show that this explanation does not apply for the patent citation network. Instead, we propose that the extended power-law form for the in-degree distribution stems from a combination of both preferential attachment and aging [22].

A different measure can be made on patent networks that allows detecting correlations and hierarchical organization [23,24]. Since patent networks display both heterogeneity and modular organization, they can be an example of a hierarchical modular system. The measure is based on the use of the clustering coefficient and how it depends on degree. The clustering coefficient is defined as follows. Let us consider the set of links a_{ij} $(i,j=1,\ldots,N)$, where $a_{ij}=1$ if a link exists and zero otherwise and that the average number of links per node is $\langle k \rangle$. Let us indicate by $\Gamma_i=\{v_j | a_{ij}=1\}$ the set of nearest neighbors of a node v_i. The clustering coefficient for this node is defined as the number of connections between the components $v_j \in \Gamma_i$. By defining

$$\mathcal{L}_i = \sum_{j=1}^{N} a_{ij}\left[\sum_{k\in\Gamma_i; j<k} a_{jk}\right], \tag{6}$$

the clustering coefficient of v_i is $C_v(i)=\mathcal{L}_i/\binom{|\Gamma_i|}{2}$ and the average clustering coefficient is simply

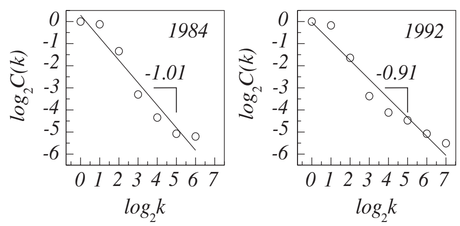

FIG. 4. Hierarchical organization of patent citation networks. Here two networks have been analyzed, corresponding to the 1984 and 1992 data sets. We can clearly appreciate that clustering decays roughly following a power-law behavior (a logarithmic binning has been used). The continuous lines correspond to the theoretical law $C(k) \sim k^{-\theta}$. The estimated exponents are indicated.

$$C = \frac{1}{N}\sum_{i=1}^{N} C_v(i) \tag{7}$$

and measures the average fraction of pairs of neighbors of a node that are also neighbors of each other. As it happens with many other complex networks, our network displays a high clustering. More interestingly, we can look at the relationship between clustering and degree. This function, $C(k)$, provides a measure of correlation between local structure and degree and has been shown to exhibit a scaling behavior

$$C(k) \sim k^{-1} \tag{8}$$

in some real networks (both natural and artificial, including language networks and social webs) and also in model graphs showing fractal rules of growth. Nonhierarchical graphs (such as the power grid) typically display a constant clustering (see [23,24] and references therein).

The analysis of patent networks shows that they also follow this scaling behavior. In Fig. 4 we plot $C(k)$ against degree for two citation networks. The estimated best fit power law gives in both cases a scaling exponent close to one. Specifically, we found $\theta = -1.01 \pm 0.06$ and $\theta = -0.91 \pm 0.08$ for the 1984 and 1992 data sets, respectively. This result supports the view that patent networks are hierarchical. Such a type of graph can emerge under different types of mechanisms [25–27], all of them dealing with multiplicative processes of some kind. In the next section we show that, using an appropriate attachment kernel (to be estimated from the USPTO data set), we can properly describe the underlying rules generating the observed topological patterns.

IV. NETWORK EVOLUTION

Let us assume that every patent has a unique identifier $0 < i < t$. Our model starts at time $t=0$ when there is only one patent in the network. From this initial network, we add a new patent i at every time step that references m previous patents. Two main forces drive the evolution of the patent citation network. First, it is natural to assume that the number of patent citations (i.e., incoming links) is a surrogate of its relevance [8]. Useful patents are more likely to receive further citations than marginal inventions. Thus the probability of receiving new citations should be proportional to the current number of citations. This rule parallels the preferential attachment mechanism of network growth [28]. Under this rule new elements entering the system connect with other nodes with a probability $\Pi(k)$ that is proportional to its degree, i.e.,

$$\Pi(k) \sim k. \tag{9}$$

However, old patents tend to be less relevant in the context of recent innovations: Attachment rates decay as the patent loses value. In particular, patents are released to the public domain after some finite period of exploitation.

The evolution of complex networks involving both preferential attachment and aging has been extensively studied. In particular, Dorogovtsev and Mendes (DM) determined analytically the scaling properties of the resulting networks [22]. In the DM model, the rule of attachment scales now as

$$\Pi(k,\tau) \sim k\tau^{-\alpha}, \tag{10}$$

where $\tau = t - i$ indicates the age of the ith node and the exponent α (which is positive) weighs how fast the aging is affecting the likelihood of attachment. Extensions of this attachment probability kernel include accelerated growth with $\Pi(k,\tau) \sim k^{\beta}\tau^{-\alpha}$ and exponential aging kernel $\Pi(k,\tau) \sim k\exp(-\tau^{\alpha})$ [21].

Finally, some models of scientific citation networks take into account the simultaneous evolution of author and paper networks [29]. In these models, the rule of attachment behaves as

$$\Pi(k_i,\tau) \sim k_i^{\beta}\tau^{\alpha-1}e^{-(\tau/\tau_0)^{\alpha}} \tag{11}$$

with a time-dependent component following a Weibull-like form. Here, τ_0 (the so-called scale parameter) gives the maximum of $\Pi(k_i,\tau)$ at a fixed k_i value and controls the rightward extension of the curve. As τ_0 increases, so does the probability of citing older papers. On the other hand, small values of τ_0 indicate strong aging that favors recently published patents [29]. Here we choose the simplest assumption (preferential attachment, $\beta=1$) and consider the aging function in Eq. (11). Consequently, the average connectivity of the ith patent at time t evolves according to the following equation:

$$\frac{\partial \bar{k}(i,t)}{\partial t} = \frac{m\bar{k}(i,t)f(t-i)}{\int_0^t \bar{k}(u,t)f(t-u)du}, \tag{12}$$

where m is the number of links introduced at each step ($m=1$ is the DM model). Now we address the following question: Is the above equation consistent with the patent network evolution? In the following, we will estimate the form of the attachment kernel (and the corresponding α, β, and τ_0 parameters) for the patent citation data.

First, we consider system size N as our time index instead of real time t. In this way we avoid any bias due to the pattern of nonlinear growth and thus follow the standard formulation of network models. Then, Eq. (12) becomes

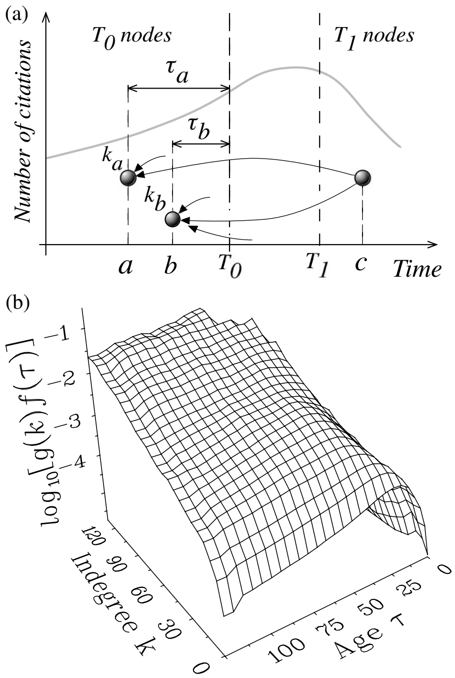

FIG. 5. (a) Illustration of the kernel estimation procedure. We compare two consecutive patent network snapshots indicated by all patents existing at time T_0 and the patents created in $[T_1, T_1+\Delta N]$. Here, $T_1 > T_0$ in order to avoid unwanted biases in the measurements [31]. The probability to cite the sth patent is a function of its age $\tau_s = T_0 - s$ and in-degree k_s, with $s < T_0$. Patent attachment probability $\Pi(k,\tau)$ is approximated by the fraction of T_0 citations in the T_1 network (see text for details). (b) The empirical attachment kernel when T_0 corresponds to the year 2002 and T_1 corresponds to the year 2003.

$$\frac{\partial \bar{k}(i,N)}{\partial N} = \frac{m\bar{k}(i,N)f(N-i)}{\int_0^t \bar{k}(u,t)f(N-u)du}. \tag{13}$$

Using $\partial\bar{k}/\partial N = (\partial\bar{k}/\partial t)(\partial t/\partial N)$ and the time-dependent scaling $N(t) = At^{\theta}$, we have

$$\frac{\partial \bar{k}(i,N)}{\partial N} = \frac{1}{A\theta}t^{1-\theta}\frac{\partial \bar{k}(i,t)}{\partial t}. \tag{14}$$

We assume the rate of attachment of new links is the product of a preferential attachment function $g(k)$ and an aging function $f(\tau)$:

$$\Pi(k,\tau) \sim g(k)f(\tau). \tag{15}$$

In order to minimize the impact of noisy fluctuations we partition the whole time interval N into $N/\Delta N$ time slots (bins) comprising the same number $\Delta N \ll N$ of patents (see [30]). Here, $N \approx 2.8$ million patents corresponding to the time interval 1976–2005. The sth time slot has the same number of new $\Delta N = 10^5$ patents. The large number of nodes N in the system ensures that we will gather sufficient samples. Following [31], we study the citation process by comparing two consecutive patent network snapshots T_0 and

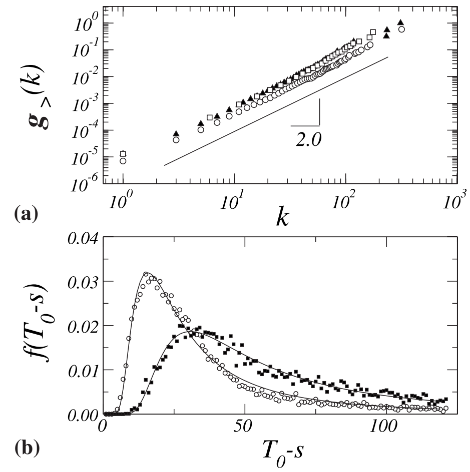

FIG. 6. Estimation of the attachment rule for the patent citation network at $T_1 = 2002$. (a) The preferential attachment function fits a scaling law $g_>(k) \sim k^{\beta+1}$ with $\beta \approx 1$. Each curve corresponds to nodes having the same age ($\tau = 120$ for the open circles, $\tau = 75$ for the filled triangles, and $\tau = 50$ for the open squares). (b) Fitting for the aging function $f(\tau)$ predicts the Weibull distribution described in the text with $\alpha \approx 1.45$. Each curve corresponds to nodes having the same in-degree ($k = 1$ for open circles and $k = 5$ for filled circles). For every curve we have used $T_1 = T_0 + 1$. These curves are cross sections of the attachment kernel [see Fig. 5(b)].

T_1. The T_0 network is defined by all the patents existing at time T_0 and the T_1 network comprises all the patents added between $[T_1, T_1+\Delta N]$, where $\Delta N \ll T_1$ and $T_1 > T_0$ [31]. To measure the attachment rule $\Pi(k_i,\tau)$ we will compute the fraction of links acquired by T_0 nodes with exactly in-degree k_i and age τ:

$$\Pi[k,\tau;T_0,T_1] = \frac{\sum_{i\in T_1, j\in T_0} a_{ij}\theta(k-k_j)\theta(\tau-\tau_j)}{\sum_{j\in T_0}\theta(k-k_j)\theta(\tau-\tau_j)}, \tag{16}$$

where a_{ij} are the adjacency matrix, $\theta(z)=1$ if $z=0$ and $\theta(z)=0$ otherwise, and k_j and $\tau_j = T_0 - j$ is the in-degree and the age of the jth node ($j < T_0$), respectively [see Fig. 5(a)]. The attachment kernel can be further decomposed in the preferential attachment $g(k)$ and aging $f(\tau)$ components [see Eq. (15)]. We can estimate these components by taking cross sections from the empirical attachment kernel [see Fig. 5(b)]. Using our data set, we have estimated that $g(k) \sim k^{\beta}$ and found $\beta = 1$, which further validates our assumption of preferential attachment [see Fig. 6(a)]. Notice that in our fittings we have used the cumulative function

$$g_>(k) = \int_0^k g(k)dk \tag{17}$$

to reduce the noise level. On the other hand, Fig. 6(b) shows the $f(\tau)$ sections together with the approximated Weibull distributions, which fit very well the aging function $f(\tau)$:

$$f(\tau) = \frac{\alpha}{\tau_0}\left(\frac{\tau}{\tau_0}\right)^{\alpha-1} \exp\left(-\frac{\tau}{\tau_0}\right)^{\alpha} \quad (18)$$

with an exponent $\alpha \approx 1.45$ and $\tau_0 \approx 40$. An obvious advantage of using the Weibull form is that it naturally includes as limit cases both exponential and Gaussian distributions. Using this distribution, we can estimate the average τ, given from

$$\langle \tau \rangle = \tau_0 \Gamma\left(1 - \frac{1}{\alpha - 1}\right) \quad (19)$$

with $\Gamma(z)$ being the gamma function. From our estimated parameters, we obtain $\langle \tau \rangle \approx 27.58$ which defines the average time required for a patent to start getting cited (using our patent-based time scale). Such a value thus represents an estimate of the time required for a patent to get known and developed in order to start being used by other inventions.

The common structure of the aging term found here and in the network of paper citations [29] suggests that common patterns of organization and evolution might be shared. The paper citation graph, obtained by looking at the list of references included in each paper, is in fact close to the basic rules defining the patent citation graph. In both cases, cross links are associated to some underlying set of features which is shared by both patents and papers. As it occurs with the patent case, new papers are based on previous ones providing the background required to build a new idea. On the other hand, as new ideas and concepts develop into well-defined areas, they will tend to attach less to more generic or older works. Additionally, the modular organization might also contribute to deviation from the simple power-law attachment assumed in previous theoretical studies. What seems clear is that there might be some universal trends canalizing the growth of innovation networks, whether scientific or technologic.

V. DISCUSSION

The patterns of innovation emerging in our society are the outcome of an extensive exchange of shared information linked with the capacity of inventors to combine and improve previous designs. Even very original inventions are not isolated from previous achievements. A patent can be identified as an object which needs a minimum amount of originality to be considered as truly different from previous patents. Moreover, to be obtained, it must properly refer to related patents in a fair way. Such constraints make this system especially interesting since we can conjecture that it represents the expansion of real designs through some underlying technology landscape. Such a concept [32] is consistent with a number of commonalities shared by both organisms and artifacts [33]. Technological evolution, as biological evolution, displays radiations, stasis, extinctions, and novelty. That technological change might actually occur on fitness landscapes is illustrated by the so-called learning curves [32,34], where universal patterns of improvement can be explained in terms of adaptive walks on so-called NK rugged landscapes [35–37]. Our analysis provides a different quantitative approach to this evolving structure by using the approach of complex networks.

In this paper, we have analyzed the architecture of large patent citation networks obtained from the USPTO database. The networks have been shown to be heterogeneous, with most patents having just a few citations whereas a small set of patents have many of them. The distribution was shown to follow a Zipf-Mandelbrot function. Together with some examples of the modular character of the network (using a community detection algorithm) we used the $C(k)$ curve as a straighforward way of identifying the presence of hierarchical organization. As it happens with other complex networks, modularity is not defined in terms of coexisting, independent groups of elements, but instead as a nested hierarchy of interconnected clusters of related inventions.

We have shown that the underlying rules of network change for our system reveal a mixture of preferential attachment favoring a rich-gets-richer mechanism together with an aging term weighting the likelihood of citing old patents. As the network grows, recent patents will tend to cite recent designs (since innovation is likely to involve redefining recent inventions) and less likely to link to old patents. The consequence of this, as predicted by previous mean field models, is that the expected scaling law in the degree distribution associated to preferential attachment kernels will be modified in significant ways. Here we have shown that the network of patents, defined by using the in degree as a surrogate of patent relevance, scales as $P(k) \sim (k+k_0)^{-\gamma}$ with $\gamma > 4$. This is not far from previous predicted scaling laws (DM) associated with preferential attachment and power-law aging [i.e., $f(t) \sim t^{-\alpha}$] which predict $P(k) \sim k^{-\gamma(\alpha)}$ (with $\gamma \sim 4$ for $\alpha \sim 0.5$). However, the humped shape of our aging term (as described by the Weibull distribution) makes it necessary to modify these approximations.

In our study of the patent network, we have tentatively characterized network modularity using standard methods of community detection, which are known to have several limitations [14]. Our partial exploration of this feature suggests that there is a good correlation between topological modules and the properties that define the sets of patents within them, but further work is required in order to prove the generality of these observations. The emergence of correlations is a natural consequence of the specialized features shared by related patents. But it might also reveal the structure of the innovation landscape itself: New patents related to previous ones can also be understood as improved solutions that explore the neighborhood of previous inventions. This view would provide a quantitative picture of the topology of technology landscapes [32,34]. Such an evolutionary interpretation in terms of fitness functions will be explored elsewhere.

ACKNOWLEDGMENTS

We thank Vincent Anton, Marti Rosas-Casals, and Stuart A. Kaufmann for useful discussions. This work has been supported by Grant No. FIS2004-05422, by the EU within the 6th Framework Program under Contract No. 001907 (DELIS), by the James McDonnell Foundation, and by the Santa Fe Institute.

[1] D. H. Erwin and D. Krakauer, Science **304**, 1117 (2004).

[2] *Symbiosis as a Source of Evolutionary Innovation*, edited by L. Margulis and R. Fester (MIT Press, Cambridge, MA, 1991).

[3] F. Jacob, Science **196**, 1161 (1977).

[4] R. V. Solé, R. Ferrer-Cancho, J. M. Montoya, and S. Valverde, Complexity **8**, 20 (2002).

[5] J. H. Lienhard, *How Invention Begins* (Oxford University Press, New York, 2006).

[6] G. Basalla, *The Evolution of Technology* (Cambridge University Press, Cambridge, UK, 1990).

[7] The collection of US patents is fully searchable and accessible from this web page (http://www.google.com/patents).

[8] *Patents, Citations and Innovations: A Window on the Knowledge Economy*, edited by A. B. Jaffe and M. Trajtenberg (MIT Press, Cambridge, MA, 2003).

[9] S. Kuznets, in *The Rate and Direction of Incentive Activity: Economic and Social Factors*, edited by R. R. Nelson (Princeton University Press, Princeton, NJ, 1962).

[10] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, New York, 2003); S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006); M. E. J. Newman, SIAM Rev. **45**, 167 (2003); R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[11] M. Trajtenberg, Rand J. Econ. **21**, 172 (1990).

[12] G. Hounsfield, U.S. Patent No. 3,778,614.

[13] A. Clauset, M. E. J. Newman, and C. Moore, Phys. Rev. E **70**, 066111 (2004).

[14] S. Fortunato and M. Barthelemy, Proc. Natl. Acad. Sci. U.S.A. **104**, 36 (2007).

[15] G. Csárdi, K. J. Strandburg, L. Zalányi, J. Tobochnik, and P. Érdi, Physica A **374**, 783 (2007).

[16] B. Verspagen, Eindhoven Center for Innovation Studies, Working Paper 05.11, 2005.

[17] F. M. Scherer, Am. Econ. Rev. **55**, 1097 (1965).

[18] F. M. Scherer, Ann. Econ. Stat. **49/50**, 495 (1998).

[19] D. Harhoff, F. M. Scherer, and K. Vopel, in *The Economics of the Patent System*, edited by O. Grandstrand (Routledge, London, 2003).

[20] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, New York, 2003).

[21] Z.-G. Shao, X.-W. Zou, Z.-J. Tan, and Z.-Z. Jin, J. Phys. A **39**, 2035 (2006).

[22] S. N. Dorogovtsev and J. F. F. Mendes, Phys. Rev. E **62**, 1842 (2000).

[23] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).

[24] E. Ravasz and A. L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[25] K. Klemm and V. M. Eguíluz, Phys. Rev. E **65**, 057102 (2002); G. Szabo, M. Alava, and J. Kertesz, *ibid.* **67**, 056102 (2003).

[26] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).

[27] G. Szabó, M. Alava, and J. Kertész, Phys. Rev. E **67**, 056102 (2003).

[28] A.-L. Barabasi and R. Albert, Science **286**, 509 (1999).

[29] K. Borner, J. T. Maru, and R. L. Goldstone, Proc. Natl. Acad. Sci. U.S.A. **101**, 5266 (2004).

[30] H. Zhu, X. Wang, and J. Y. Zhu, Phys. Rev. E **68**, 056121 (2003).

[31] H. Jeong, Z. Néda, and A. L. Barabási, Europhys. Lett. **61**, 567 (2003).

[32] S. A. Kauffman, *Investigations* (Oxford University Press, New York, 2000).

[33] S. J. Gould, *The Structure of Evolutionary Theory* (Harvard University Press, New York, 2002).

[34] S. A. Kauffman, *At Home in the Universe* (Oxford University Press, New York, 1995).

[35] S. Kauffman and S. J. Levin, J. Theor. Biol. **128**, 11 (1987).

[36] E. D. Weinberger, Phys. Rev. A **44**, 6399 (1991).

[37] S. A. Kauffman, J. Lobo, and W. G. J. Macready, J. Econ. Behav. Organ. **43**, 141 (2000).