Information Sciences 677 (2024) 120937

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Knowledge complexity based on coupled equations within the bipartite network

Shenshen Sergio Zhang

School of Economics and Finance, Shanghai International Studies University, Shanghai 200080, China

ARTICLE INFO

Keywords:
Bipartite network
Knowledge complexity
The fitness and complexity algorithm
The matrix-estimation exercise

ABSTRACT

In the era of competing for core competence, the increasing inputs of knowledge factors have brought the issue of "efficiency-enhancing and quality-improving" into focus; and concerns about the "quality" perspective of knowledge require mining information related to complex knowledge hidden in the economic system. In order to quantify knowledge complexity at both the national (or regional) and technological levels, this article combines the Fitness and Complexity algorithm with matrix-estimation exercises based on the framework of the bipartite network. On the basis of these measurements, this article analyzes and discusses the economic implications and evolutionary features while considering the "expiration" of patents; additionally, community detection is conducted to discuss the evolution of the "location" of complex knowledge. The results show that knowledge complexity depends on the structural similarity and specialization of patents; furthermore, the timeliness of patents may affect knowledge complexity conspicuously; moreover, the significance of the "location" of complex knowledge in the past has been downplayed over the past few decades.

1. Introduction

As globalization advances, it has been increasingly significant to invest knowledge into global competition [11,13]. Many scholars believe that differences in the output and dissemination of knowledge are crucial factors of economic growth and unbalanced development of countries [10]. Sorenson and Fleming [42] believed that the difference between the output and dissemination of knowledge is related to the complexity of knowledge. If it is more difficult to grasp some complex knowledge, going into the corresponding fields will be more costly, which usually impedes the dissemination of such knowledge among economic entities [27,43]. The higher the complexity of knowledge, the stronger its "exclusivity"; and it will be more challenging to produce such knowledge in areas other than the original area and more difficult to transfer between different areas [7], which makes it has higher economic value [37,45].

Under the background of globalization, while referring to a certain economic entity, the "core competence" is closely associated with the output of high-value, tacit, and complex knowledge [11,13]. In the economic system, such complex knowledge is highly valuable information hidden deep within, urgently awaiting extraction through efficient approaches. Therefore, studies on economic geography hope to find a practical approach to measure complex knowledge, in order to explore how the complex knowledge is related to regional economic growth, and try to clarify the relationships between the complex knowledge and the mode of economic

E-mail address: s.sergio.zhang@gmail.com.

https://doi.org/10.1016/j.ins.2024.120937
Received 20 May 2023; Received in revised form 16 May 2024; Accepted 4 June 2024
Available online 10 June 2024

operation or the mechanism of economic development [1,3,6,22,36]. In previous studies, in order to measure the complexity of knowledge, many scholars have conducted attempts [15,17,34].

However, the key problem lies with most of these measurements of knowledge complexity being only based on simple calculations of the input and output of knowledge, e.g., the number of patent applications, the times of breakthroughs in crucial domains, or the trade volume of new products [14,46,47]. Nevertheless, such approaches also imply that only a small portion of the superficial information about "complex knowledge" has been extracted, while a considerable amount of deeper information remains concealed. In particular, these approaches generally involve two issues. First, they ignore the evaluation of the "heterogeneity" of knowledge itself, assuming that knowledge from different domains only differs in quantity. Second, they overlook the "interconnectedness" between different types of knowledge, assuming that knowledge from various domains is mutually isolated, which finally leads to such measurements being unconvincing [7].

In fact, the "quality" of complex knowledge, rather than its "quantity", is the more valuable information that deserves our attention [20,33], which means that it requires us to pay more attention to the "heterogeneity" and "interconnectedness" of complex knowledge. First, the "heterogeneity" of knowledge suggests that knowledge from different domains contributes to the complexity of an economic entity differently, i.e., it is necessary to consider the differences in the complicacy attributes of knowledge in their respective domains. Meanwhile, we also need to consider the "interconnectedness" of knowledge, meaning that changes in the complicacy of knowledge domains will affect economic entities that are deeply learning or have fully mastered them, which, in turn, indirectly affects the complicacy of other knowledge domains. Such repercussions will propagate continuously through the transmissivity of complex systems.

The method based on the NK model proposed by Fleming and Sorenson [15] is a meritorious attempt, whose applicability and feasibility have been confirmed by several studies. However, the measurement of knowledge complexity at the regional level still needs to be included [42,43]. Despite that, their utilization of binary variables coincides with the idea proposed by Hidalgo and Hausmann [23], which is based on the bipartite network. In order to study the economic complexity, Hidalgo and Hausmann pioneered a linear iterative operation called the "Method of Reflection (MR)", which provides a feasible idea for quantifying the knowledge complexity.

As mentioned above, similar to some traditional measurements, economic complexity is also based on trade volume or value added of production, which also quantifies complex knowledge to a certain extent. However, the problem lies with some kind of "self-realization" of complexity, especially since value-added itself measures the value chain position of commodities to some extent, while complex knowledge is associated with the creation of economic value [7,37] which also involves the value chain. The relationship between the two may be circular reasoning. Moreover, although the MR overcomes the shortcomings of the former methods and involves the complexity of regional knowledge, it has inherent defects of information omission [25,30], which will lead to biased conclusions that the interaction between the complexity of knowledge domains and that of economic entities cannot be described thoroughly and accurately; and in the calculation of the Eigenvector Centrality, ad hoc assumptions cannot be avoided, which may lead to circular argumentation fallacy [39].

Therefore, for the purpose of finding a more practical approach to measuring knowledge complexity, this article first aims to place analyses within the framework of the bipartite network. As mentioned above, complex knowledge implies the creation of advantages, and its geographic stickiness means that such advantages are prone to be self-sustaining [2,8]. Continuous, extensive, or deep involvement in technology domains containing more complex knowledge may suggest the establishment of such advantages. Therefore, more specifically, this article aims to identify such an "apparent advantage" by calculating countries' "outperformance" in terms of patent outputs relative to other countries. Based on this, a country-technology network is constructed to serve as the analytical foundation for uncovering deeply hidden information. And then, on the basis of the calculation results, this article gives a reasonable explanation on the theoretical and algorithmic level combined with current statistical data, and conducts analyses and discussions on its evolution. The results indicate that there is not a strong correlation between knowledge complexity and the volume of patents; knowledge complexity depends on the structural similarity and specialization of patents. The "expirations" of patents may have a considerable impact on knowledge complexity, but it varies depending on the development and R&D capabilities of countries. Additionally, this article also conducts community detection to inspect whether and how the "location" heterogeneity of complex knowledge changes, and the results suggest that such locational heterogeneity has been fading.

The rest of this article is structured as follows: Section 2 elaborates on the method of indicator construction in detail and explains the topological significance of the indicators; Section 3 describes the data pre-processing and the details of constructing the bipartite network; Section 4 presents the computational results in two parts based on the patent "expirations" and conducts discussions; finally, this article performs community detection and analyzes the "location" of knowledge in Section 5, and concludes with a discussion on policy implications.

2. Method

2.1. Framework based on bipartite network

In research on international trade, the bipartite network can map the trade information of different countries on different products, doing well in clarifying the underlying relationships among countries or among products, which cannot be realized by traditional complex networks [23,40,44]. This article constructs a country-technology bipartite network (the Adjacent Matrix corresponding is $\boldsymbol{M} = \{M_{ct}\}$) by mapping information of patent outputs of different countries on different technology domains, based on defining the two different types of vertexes: countries and products, and edges between two different types of vertexes.

Within the framework of the bipartite network, issues on complexity can be divided into two dimensions, namely, discussing the complexity of two objects corresponding to the two types of vertexes. For the framework constructed in this article, issues of knowledge complexity can be described by determining two properties — X_c and Y_t, which represent the complexity properties of counties and products, respectively. The property X_c can be expressed as a function composed of property Y_1, Y_2,…, Y_t and the Adjacency Matrix M_{ct} of the bipartite network; at the same time, property Y_t can be also expressed as a function composed of property X_1, X_2,…, X_c and M_{ct}. And then, studying economic problems is just equal to solving the coupled equations below:

$$\begin{cases} X_c = f(Y_1, Y_2, \dots, Y_t, M_{ct}) \\ Y_t = g(X_1, X_2, \dots, X_c, M_{ct}) \end{cases} \tag{1}$$

The two functions in Eq. (1) allows one to recast the determination of X_c and Y_t as the solutions of an eigen-problem of a suitable transformation matrix $\boldsymbol{W}$.

$$\begin{cases} X_c = \frac{1}{\sqrt{\lambda}} \sum_t W_{ct} Y_t \\ Y_t = \frac{1}{\sqrt{\lambda}} \sum_c W_{ct} X_c \end{cases} \Leftrightarrow \begin{cases} X_c = \frac{1}{\lambda} \sum_t \sum_{c'} W_{ct} W_{c't} X_{c'} = \frac{1}{\lambda} \sum_{c'} N_{cc'} X_{c'} \\ Y_t = \frac{1}{\lambda} \sum_c \sum_{t'} W_{ct} W_{ct'} Y_{t'} = \frac{1}{\lambda} \sum_{t'} G_{tt'} Y_{t'} \end{cases} \tag{2}$$

where, W_{ct} can be calculated from M_{ct} by some kind of algorithm; λ is eigenvalue. $N_{cc'} = (\boldsymbol{W}\boldsymbol{W}^T)_{cc'}$ and $G_{tt'} = (\boldsymbol{W}^T\boldsymbol{W})_{tt'}$ are symmetrical square matrices; thereinto, $\boldsymbol{N}$ can be explained as the proximity matrix of the country, describing the similarity of the structure of countries' participating in technology domains; $\boldsymbol{G}$ can be correspondingly explained as the proximity matrix of the technology, describing the similarity of the structure of technology domains' being participated by countries.

2.2. The fitness and complexity algorithm

Actually, the research objects of this article — represented by X_c and Y_t — denote just some kind of "property" of nodes in the network, which is referred to as "complicacy" or so-called "centrality" (but distinct from "complexity") in the context of this article. One of the tasks of complex network analysis is to quantify such "complicacy" by constructing algorithms, with the essence of these algorithms lying in mathematically describing the relationships between nodes through iterative equations. By constructing algorithms, W_{ct} can be obtained, which further facilitates solving the eigenvalue problem represented by Eq. (2). Initially, Hidalgo and Hausmann [23] proposed the "Method of Reflection" and systematically studied the coupled equations in Eq. (2). The MR can be numerically expressed as:

$$\begin{cases} \widetilde{k_c}^{(n)} = \sum_t M_{ct} {k_t}^{(n-1)}, \quad \forall c \\ \widetilde{k_t}^{(n)} = \sum_c M_{ct} {k_c}^{(n-1)}, \quad \forall t \end{cases} \Leftrightarrow \begin{cases} {k_c}^{(n)} = \frac{\widetilde{k_c}^{(n)}}{\sum_t M_{ct}} \\ {k_t}^{(n)} = \frac{\widetilde{k_t}^{(n)}}{\sum_c M_{ct}} \end{cases} \tag{3}$$

where ${k_c}^{(n)}$ and ${k_t}^{(n)}$ denote the normalized quantifications of centrality, referring to X_c and Y_t respectively; and the initial conditions are: $\begin{cases} \widetilde{k_c}^{(0)} = \sum_t M_{ct}, \quad \forall c \\ \widetilde{k_t}^{(0)} = \sum_c M_{ct}, \quad \forall t \end{cases}$.

The Method of "Reflection" also derives its name from the linear form of its iteration; but precisely because of this, as mentioned above, the flaw of information omission is essentially rooted in the MR, which means that the MR may conduct incorrect evaluations on the diversification of countries and ubiquity of products [21,30]; what's more, the non-convergency in the iterations may also have impacts on calculation results significantly [35]. That implies that, while measuring the complexity based on the MR within the bipartite network, the complexity property, of one type of vertex, may not be sensitive enough to changes on the same property of the other type; namely, the MR cannot reflect fine structure of the pattern on interactions between two types of vertexes.

Actually, there should be a difference when discussing the complexity of these two types of vertexes, which can be reflected in the difference in their own formulas. While referring to a certain country, certeris paribus, the wider the technology domains it participates in, the higher the knowledge complexity it will be; in addition, when the centrality or importance of a technology the country participating in, and more deeply the country participating in such technology, the higher the knowledge complexity it will be. As for a certain technology, the situation is different. When the participating countries of the technology are more important, the knowledge complexity of the technology is higher; but, if it is more widely participated by countries, namely the number of participating countries is larger, the knowledge complexity should be lowered.

Tacchella et al. [44] proposed the Fitness and Complexity (FC) algorithm, which expresses the interrelationships between two types of vertexes as non-linear functions. This algorithm can not only acquire almost all information calculated from the MR [28,40], but also avoid the shortcomings of the MR mentioned above [25,30]. Therefore, this article will adopt the FC algorithm, and recast the mathematical expression of the complexity issues from Eq. (2), by determining two complexities: F_c — fitness for countries, and Q_t — complexity for technologies:

$$\begin{cases} \widetilde{F_c}^{(n)} = \sum_t M_{ct} {Q_t}^{(n-1)}, \quad \forall c \\ \widetilde{Q_t}^{(n)} = \frac{1}{\sum_c M_{ct} \frac{1}{{F_c}^{(n-1)}}}, \quad \forall t \end{cases} \Leftrightarrow \begin{cases} {F_c}^{(n)} = \frac{\widetilde{F_c}^{(n)}}{\sum_c \widetilde{F_c}^{(n)}/C} \\ {Q_t}^{(n)} = \frac{\widetilde{Q_t}^{(n)}}{\sum_t \widetilde{Q_t}^{(n)}/T} \end{cases} \tag{4}$$

and the initial conditions are: $\begin{cases} \widetilde{F_c}^{(0)} = 1, & \forall c \\ \widetilde{Q_t}^{(0)} = 1, & \forall t \end{cases}$.

In Eq. (4), the intermediations of fitness and complexity, $\widetilde{F_c}^{(n)}$ and $\widetilde{Q_t}^{(n)}$, are all standardized in each step, which make $\widetilde{F_c}^{(n)}$ and $\widetilde{Q_t}^{(n)}$ tend to be converged. Therefore, the equations in iterative form can be expressed as non-iterative form, by introducing rescaling factors $c_F = {}^{C}/_{\sum_t Q_t s_t}$ and $c_Q = {}^{\sum_t Q_t s_t}/_{T}$:

$$\begin{cases} F_c = c_F \sum_t M_{ct} Q_t \\ Q_t = c_Q \frac{1}{\sum_c M_{ct} \frac{1}{F_c}} \end{cases} \tag{5}$$

where s_t is the weight of technology vertex, namely $s_t = \sum_c M_{ct}$. In Eq. (5), Q_t is expressed as a non-linear function $f_{non-linear}(F_1, F_2, \dots, F_c, M_{ct})$. Based on the Taylor Expansion [40], Eq. (5) can be approximatively expressed in a linear form:

$$\begin{cases} F_c \backsimeq c_F \sum_t M_{ct} Q_t \\ Q_t \backsimeq \frac{c_Q}{(s_t')^2} \sum_c \frac{M_{ct} F_c}{{s_c}^2} \end{cases} \tag{6}$$

where s_c is the weight of country vertex, namely $s_c = \sum_t M_{ct}$ and $s_t' = \sum_c {}^{M_{ct}}/_{s_c}$.

2.3. Matrix-estimation exercise

Due to superior properties compared to MR, the FC algorithm can provide a more feasible solution for quantifying centrality. Therefore, in subsequent sections, this article will investigate the country-technology bipartite network and measure the knowledge complexity based on the FC algorithm. Notwithstanding, there is still an issue to address at this point, namely, it is inevitable for either the MR or the FC algorithm to generate ad hoc assumptions for their purpose of calculating the Eigenvector Centrality to some extent, which will make the results lack of effectiveness [39].

Consider adopting the methods of matrix-estimation exercise. In terms of an adjacency matrix $\boldsymbol{A}$, the matrix estimator $\hat{A}_{ij}$ will depend on the centrality of vertex i and j, namely:

$$\hat{A}_{ij} = f(x_i, x_j) \tag{7}$$

The same as OLS, $\hat{A}_{ij}$ can be estimated by minimizing its Squared Error (SE). Actually, SE be partitioned into SE_0 and SE_k; thereinto SE_0 is independent of x_k, and SE_k depending on x_k:

$$SE = SE_0 + SE_k = \sum_i \sum_j (A_{ij} - \hat{A}_{ij})^2 = \sum_i \sum_j (A_{ij} - f(x_i, x_j))^2 \tag{8}$$

Because of SE_0 being independent of x_k, minimizing SE just equal to minimizing SE_k, which means that the derivate of SE_k should be equal to 0, namely:

$$\begin{aligned} \frac{\partial SE}{\partial x_k} &= \frac{\partial SE_k}{\partial x_k} = 0 \\ &= \partial\Big(\sum_{i\neq k}(A_{ik} - f(x_i, x_k))^2 + \sum_{k\neq j}(A_{kj} - f(x_k, x_j))^2 + (A_{kk} - f(x_k, x_k))^2\Big)\Big/\partial x_k \\ &= 4\sum_{i\neq k}\left(A_{ik} - f\left(x_i, x_k\right)\right)\frac{\partial f\left(x_i, x_k\right)}{\partial x_k} + 2\left(A_{kk} - f\left(x_k, x_k\right)\right)\frac{\partial f\left(x_k, x_k\right)}{\partial x_k} \end{aligned} \tag{9}$$

By introducing the bound variable z_m, Eq. (9) can be expressed as:

$$\begin{aligned} \frac{\partial SE_k}{\partial x_k} &= 4\sum_{i\neq k}\left(A_{ik} - f\left(x_i, x_k\right)\right)\frac{\partial f(x_i, z_m)}{\partial z_m}\Bigg|_{z_m = x_k} + 2\left(A_{kk} - f\left(x_k, x_k\right)\right)\cdot 2\frac{\partial f(x_k, z_m)}{\partial z_m}\Bigg|_{z_m = x_k} \\ &= 4\sum_i (A_{ik} - f(x_i, x_k))\frac{\partial f(x_i, z_m)}{\partial z_m}\Bigg|_{z_m = x_k} \end{aligned} \tag{10}$$

And based on the concept of unique contribution of commonality analysis, the unique contribution of the vertex k can be expressed as:

$$
\begin{aligned}
UC_k &= R_N^2 - R_{N\backslash k}^2 = \frac{SE_{N\backslash k} - SE_N}{TSS} = \frac{SE_{N\backslash k} - SE_N}{\sum_i \sum_j (A_{ij} - \bar{A}_{ij})^2} \\
&= \frac{SE_{N\backslash k} - SE_N}{\sum_i \sum_j (A_{ij} - \frac{\sum_i \sum_j A_{ij}}{N^2})^2} \\
&= \frac{SE_{N\backslash k} - SE_N}{\sum_i \sum_j {A_{ij}}^2 - 2\frac{\sum_i \sum_j A_{ij}}{N^2} \sum_i \sum_j A_{ij} + \frac{(\sum_i \sum_j A_{ij})^2}{N^2}} \\
&= \frac{SE_{N\backslash k} - SE_N}{\sum_i \sum_j A_{ij}(1 - \frac{\sum_i \sum_j A_{ij}}{N^2})}
\end{aligned}
\tag{11}
$$

where R_N^2 denotes the goodness of the estimation $\hat{A}_{ij}$ by considering all centrality of vertexes, and $R_{N\backslash k}^2$ denotes that by excluding vertex k. Since invariance of TSS whether excluding vertex k or not, UC_k just depends on $\Delta SE = SE_{N\backslash k} - SE_N$:

$$
\begin{aligned}
\Delta SE &= 2\sum_{i\neq k} \left(\left(A_{ik} - f\left(x_i, 0\right)\right)^2 - \left(A_{ik} - f\left(x_i, x_k\right)\right)^2 \right) + (A_{kk} - f(x_0, x_0))^2 - (A_{kk} - f(x_k, x_k))^2 \\
&= 2\sum_{i\neq k} (f(x_i, 0) - f(x_i, x_k))(f(x_i, 0) + f(x_i, x_k) - 2A_{ik}) + (f(0,0) - f(x_k, x_k))(f(0,0) + f(x_k, x_k) - 2A_{kk})
\end{aligned}
\tag{12}
$$

Since coupling Equations of Eq. (2) refers to eigen problem, given multi-component eigenvector centrality, estimation $\hat{A}_{ij}$ can be re-written as:

$$
\hat{A}_{ij}(s) = f(x_i, x_j) = \sum_{t=1}^{s} \gamma_t x_{i,t} x_{j,t} = \gamma_1 x_{i,1} x_{j,1} + \gamma_2 x_{i,2} x_{j,2} + \dots + \gamma_s x_{i,s} x_{j,s}
\tag{13}
$$

Then, δSE can be expressed as:

$$
\begin{aligned}
\Delta SE(s) &= 2\sum_{i\neq k} \left(-\sum_{t=1}^{s} \gamma_t x_{i,t} x_{k,t} \right) \left(\sum_{t=1}^{s} \gamma_t x_{i,t} x_{k,t} - 2A_{ik} \right) - \left(\sum_{t=1}^{s} \gamma_t x_{k,t}^2 \right) \left(\sum_{t=1}^{s} \gamma_t x_{k,t}^2 - 2A_{kk} \right) \\
&= -2\sum_{t=1}^{s} \gamma_t^2 x_{k,t}^2 \sum_i x_{i,t}^2 + 4\sum_{t=1}^{s} \gamma_t x_{k,t} \sum_i A_{ik} x_{i,t} + \left(\sum_{t=1}^{s} \gamma_t x_{k,t}^2 \right)^2
\end{aligned}
\tag{14}
$$

The unique contribution UC_k of vertex k can be expressed as:

$$
UC_k(s) = 2\sum_{t=1}^{s} \gamma_t^2 x_{k,t}^2 + \left(\sum_{t=1}^{s} \gamma_t x_{k,t}^2 \right)^2
\tag{15}
$$

Undoubtedly, unique contribution UC_k represents the importance of vertex k; the more important the vertex k is, the more significant the difference between $SE_{N\backslash k}$ and SE_N will be, which means that the larger the unique contribution UC_k will be. Worthy mentioning, the parameter s may be crucial, because it will determine the number of eigenvectors being introduced into the matrix-estimation. Every eigenvector of N/G may include effective information of topological or clustering fine-structure; therefore, introducing more eigenvectors in the calculation may do favor to mapping deeper information of complexity, even results on dramatically different conclusions [41]. But as eigenvalues descend, introducing excessive eigenvectors may generate overmuch noise [18]. So this article will just employ 2 eigenvectors as Sciarra et al. [39] did, namely set $s = 2$.

Thus, based on combining the FC algorithm and the matrix-estimation exercise, this article constructs indicators for measuring the knowledge complexity of countries and technologies by solving the eigen-problem described by Eq. (2), which are GPYC (Generalized comPlexitY of Country) and GPYT (Generalized comPlexitY of Technology) respectively.

Eq. (5) is a linear form of Eq. (1) based on the FC algorithm, let $\begin{cases} X_c = {}^{F_c}/_{s_c} \\ Y_t = {}^{Q_t}/_{s_{t'}} \end{cases}$, and $\boldsymbol{N/G}$ can be expressed as:

$$
\begin{cases}
N_{cc'} = \frac{\sum_{c'} M_{ct} M_{c't}}{(s_t')^2 s_{c'} s_c} \\
G_{tt'} = \frac{\sum_{t'} M_{ct} M_{ct'}}{(s_t)^2 s_{t'}' s_t'}
\end{cases}
\tag{16}
$$

On the basis of Eq. (16), this article constructs GPYC and GPYT as follows:

$$
\begin{cases}
GPYC_c = \left(\sum_{i=1}^{2} \lambda_i^N \left(v_{c,i}^N \right)^2 \right)^2 + 2\sum_{i=1}^{2} \left(\lambda_i^N \right)^2 \left(v_{c,i}^N \right)^2 \\
GPYT_t = \left(\sum_{i=1}^{2} \lambda_i^G \left(v_{t,i}^G \right)^2 \right)^2 + 2\sum_{i=1}^{2} \left(\lambda_i^G \right)^2 \left(v_{t,i}^G \right)^2
\end{cases}
\tag{17}
$$

Thereinto, λ_1^N/λ_2^N and λ_1^G/λ_2^G are the largest 2 eigenvalues of the proximity matrices $\boldsymbol{N}$ and $\boldsymbol{G}$, $\upsilon_{c,1}^N/\upsilon_{c,2}^N$ and $\upsilon_{t,1}^G/\upsilon_{t,2}^G$ are the eigenvectors corresponding to λ_1^N/λ_2^N and λ_1^G/λ_2^G respectively. Worthing mentioning, $\boldsymbol{N}$ and $\boldsymbol{G}$ here are the modified matrix, by setting all diagonal elements to an arbitrary constant value (such as 0), namely $\begin{cases} N_{cc'} = \frac{\sum_{c'} M_{ct} M_{c't}}{(s_t')^2 s_{c'} s_c}, & \forall c' \neq c \\ N_{cc'} = 0, & c' = c \end{cases}$ and $\begin{cases} G_{tt'} = \frac{\sum_{t'} M_{ct} M_{ct'}}{(s_t)^2 s_{t'}' s_t'}, & \forall t' \neq t \\ G_{tt'} = 0, & t' = t \end{cases}$,
and redundant self-proximity can be deleted, making the modified $\boldsymbol{N}$ and $\boldsymbol{G}$ just reflect the inter-proximity information of vertex.

Intuitively, the higher the GPYC or GPYT, the greater the importance of the vertex in the network, representing higher knowledge complexity, whether in terms of countries or technology domains in this article. According to the steps outlined above, it is apparent that the construction rationale of GPYC/GPYT is based on traditional Eigenvector Centrality within the one-mode network, meaning that their topological significance is similar. Traditional Eigenvector Centrality describes such a fact: the centrality or importance of a certain vertex depends not only on its property, but also on its neighbors' properties. Correspondingly, GPYC/GPYT indicators indicate that the knowledge complexity of a country vertex depends on both the knowledge complexity of other country vertexes (which are unlinked to each other) and the knowledge complexity of technology vertices (which are linked directly); and vice versa for a certain technology vertex. However, unlike the Eigenvector Centrality, based on specific algorithms, GPYC/GPYT will be applicable to bipartite networks; moreover, the latter undergoes the rectifications of matrix-estimation, avoiding the usual shortcomings of the former while retaining as much valid information from the network as possible.

3. Data pre-processing

The concept of "technology vertex" and "country vertex" mentioned in this article are stemmed from the country-technology bipartite network. Intuitively, one feasible approach to construct a country-technology bipartite network is creating a network based on the number of patent outputs in different technology domains, which are categorized according to a certain classification method. In fact, data on patent outputs have been a staple of a considerable amount of macro-level literature on knowledge production [19,24,26,31]. Empirically, a patent office grants a "patent" in recognition of an innovation in knowledge or ideas. It can be expected that the number of patents recognized by a patent office contains information about the wealth of knowledge that they provide [26]. Here, an underlying assumption is that all patents within the same technology field contain roughly the same complex knowledge.

Therefore, this article considers using the volume of patent outputs as raw data to construct a country-technology domain bipartite network [7,48]. The data is collected from the item "Science, Technology and Patents" in the OECD statistics dataset, which encompasses the number of patent grants of 100 countries (or regions) in detailed 35 sub-domains classified by WIPO, from 1985 to 2019. Additionally, there are also several merits of using WIPO fields as the classification standard, e.g., such 35 fields are neither too large nor too small, covering all IPC classifications with a surjective mapping, and avoiding the duplication of patent counts [38].

Worthy mentioning, rather than directly based on initial data of patent number, the country-technology bipartite network constructed in this article will be based on the RCA (Revealed Comparative Advantage) proposed by [4,5]. The advantages of using RCAs include: first, patent data can be non-dimensionalize by doing this; and second, the use of the "outperformance" implication inherent in RCAs to represent a country's relative performance in patent outputs compared to other countries, enhancing comparability across technology domains and countries. Employing RCAs means that the element M_{ct} of the Adjacency Matrix $\boldsymbol{M}$ corresponding to the bipartite network is actually the RCA_{ct}. The RCA_{ct} can be calculated from the data mentioned above through the following formula:

$$M_{ct} = RCA_{ct} = \frac{P_{ct} / \sum_t P_{ct}}{\sum_c P_{ct} / \sum_c \sum_t P_{ct}} \tag{18}$$

However, it is worth noticing that, given the property of knowledge included by the patent, since a patent completes to be "output", such a patent will be valid for a long term; or in other words, it will not expire within a certain spell, contributing to the knowledge complexity of the technology domain corresponding. Therefore, while calculating the RCA of a certain year, this article counts the number of patents within a certain period before this year, not just the number of patents of this year.

On account of that, this article will conduct two parts of analyses based on the "expiration" of patents. One is based on the hypothesis of patents being valid in the long term, which means that the RCA will be calculated by using the number of patents from 1985 to a certain year, namely $P_{ct,y} = \sum_{i=1985}^{y} N_{ct,i}$. The other is based on the hypothesis of patents being expired after certain years, meaning that the RCA will be calculated by: $P_{ct,y} = \sum_{i=y-exp}^{y} N_{ct,i}$, where exp denotes "expiration".

4. Results

4.1. Long-term "validity"

Table 1 shows technology domains with their GPYTs, ranking 1st to 5th and 31st to 35th in 2019; Table 2 shows countries with their GPYCs, ranking 1st to 10th and 91st to 100th in 2019; and Fig. 1 shows changes of GPYC and GPYT rankings from 1990-2019.

From the view of GPYTs, results show that the top-ranked technology domains are generally considered high-tech with knowledge or capital intensity. These fields have fewer participating countries, and the participating countries generally have higher knowledge

Table 1
GPYT results[1] of technology domains in 2019[2].

GPYT Ranking	Technology Domain	GPYT[3]	Patents	Patents Ranking
1	Semiconductors	1.0000	122140	11
2	Audio-visual technology	0.9802	146219	9
3	Digital communication	0.9593	265260	3
4	Basic communication processes	0.9487	32440	34
5	Optics	0.9410	118609	12
31	Pharmaceuticals	0.5655	212699	5
32	Other consumer goods	0.5532	81267	23
33	Medical technology	0.5079	269729	2
34	IT methods for management	0.4796	68747	26
35	Environmental technology	0.0000	51273	31

1 The full results of GPYT in 2019 have been provided in the Appendix Table A1.
2 The complete results of GPYT from 1990-2019 have been provided in the Supplementary Material.
3 NOTE: GPYTs have been Min-Max Normalized.

Table 2
GPYC results[1] of countries or regions in 2019[2].

GPYC Ranking	Country or Region	GPYC[3]	Patents	Patents Ranking
1	China	1.0000	376819	4
2	Japan	0.9598	698133	2
3	Chinese Taipei	0.9035	9881	26
4	Singapore	0.8544	13271	23
5	South Korea	0.8410	190977	5
6	Malaysia	0.8326	4377	34
7	Netherlands	0.8112	80280	8
8	Finland	0.8036	37752	15
9	United States	0.7560	1204578	1
10	Israel	0.7305	42187	14
91	Bermuda	0.3072	131	92
92	El Salvador	0.2797	81	96
93	Andorra	0.2770	27	88
94	Zimbabwe	0.2668	36	95
95	Mongolia	0.2638	18	99
96	Nigeria	0.2568	45	93
97	Guatemala	0.2493	45	94
98	Ecuador	0.2439	137	80
99	Seychelles	0.2147	26	97
100	Djibouti	0.0000	0	100

1 The full results of GPYC have been provided in the Appendix Table A2.
2 The complete results of GPYC from 1990-2019 have been provided in the Supplementary Material.
3 NOTE: GPYCs have been Min-Max Normalized.

complexity. Additionally, these participating countries have substantial comparative advantages in these fields. In contrast, the bottom-ranked technology domains differ from the aforementioned situation. Most of these fields correspond to resource-intensive or labor-intensive industries. These fields are widely participated in by countries worldwide, resulting in lower knowledge complexity. Worthy mentioning, pharmaceuticals and medical technology were previously considered advanced technologies, but their GPYTs rank last in 2019. Besides, there is no apparent correlation between GPYT and patent outputting.

From the view of GPYCs, China is ranked 1st and Chinese Taipei, as one of the “Four Asian Tigers”, is ranked 3rd; Japan, Singapore, and South Korea are ranked 2nd, 4th and 5th respectively, which are considered to be high-tech countries; the countries or regions that rank lower are mainly developing countries in Africa or Latin America. It is noteworthy that some developing countries that were not previously considered to have high levels of technological advancement have a high GPYC, e.g., Malaysia and Pakistan (ranked 6th and 12th respectively); countries that were seen to be technologically developed do not have high GPYCs, e.g., United States and Germany (ranked 9th and 14th respectively). In addition, there is also no explicit correlation between GPYC and the volume of patent outputting.

In terms of GPYT ranking changes, it can be observed that the top 5 technology domains have been keeping in high ranks, while the lower-ranked domains tend to maintain a lower level of knowledge complexity; and it can also be easily observed that the rankings of IT management and Pharmaceuticals just show a rapid decline and eventually stabilize at a lower level.

From the view of GPYC ranking changes, South Korea, Singapore, and Japan have consistently ranked high, while countries in Africa or Latin America have consistently ranked lower. The rankings of China, Chinese Taipei, and developing countries such as

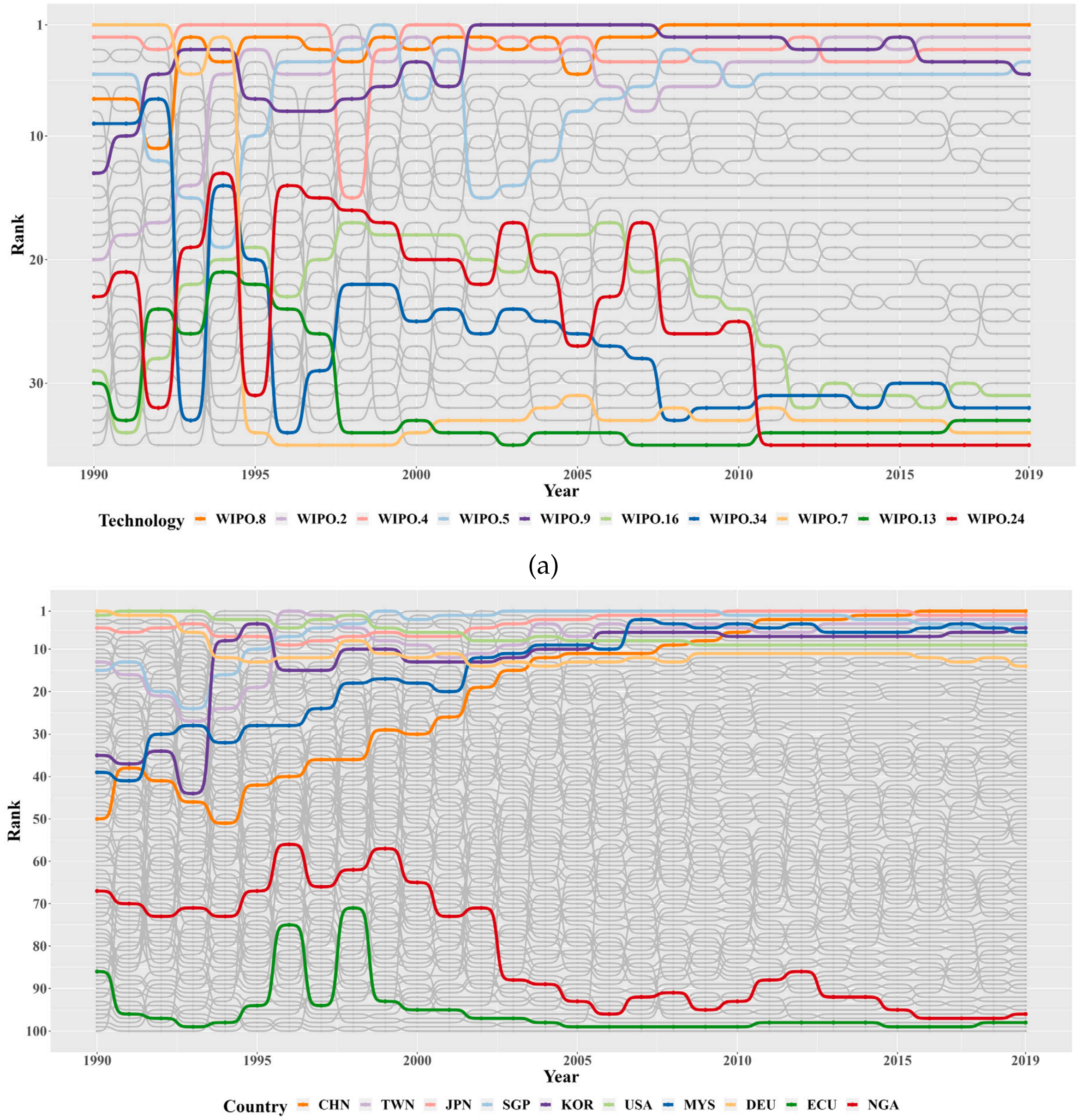


Fig. 1. Ranking changes of GPYT and GPYC in 1990-2019. (a) shows the GPYT ranking changes (10 specific domains have been marked); (b) shows the GPYC ranking changes (10 specific countries have been marked).

Malaysia have risen rapidly in the past 30 years; while developed countries like the United States have shown a slow decline in their GPYC rankings.

To further analyze and explain the calculation results obtained above, this article maps the two largest eigenvectors $\boldsymbol{v}_1^{\boldsymbol{G}}/\boldsymbol{v}_2^{\boldsymbol{G}}$ and $\boldsymbol{v}_1^{\boldsymbol{N}}/\boldsymbol{v}_2^{\boldsymbol{N}}$ corresponding to the proximity matrix $\boldsymbol{G}$ and $\boldsymbol{N}$ onto the x- and y-coordinate, respectively. Results of 35 technology domains and 100 countries (or regions) are shown in Fig. 2.

Actually, $\boldsymbol{v}_1^{\boldsymbol{G}}$ and $\boldsymbol{v}_1^{\boldsymbol{N}}$ denotes the Eigenvector Centrality of the weighted undirected network corresponding to the $\boldsymbol{G}$ and $\boldsymbol{N}$. A higher value of $v_{t,1}^{G}$ implies that the corresponding technology domain tends to have more similarity in patent output to those with higher GPYT, and a higher value of $v_{c,1}^{N}$ implies that the corresponding country tends to have more similarity in patent contribution to countries with higher GPYC. $\boldsymbol{v}_2^{\boldsymbol{G}}$ and $\boldsymbol{v}_2^{\boldsymbol{N}}$ denote the clustering structure of the network derived from $\boldsymbol{G}$ and $\boldsymbol{N}$ [29,40]. In the framework of this article, $\boldsymbol{v}_2^{\boldsymbol{G}}$ describes the concentration of technology domains participated by different countries; i.e., a higher absolute value of $v_{t,2}^{G}$ of a certain technology domain indicates that it is concentratedly participated by fewer countries. In addition, if $v_{t,2}^{G}$ is positive, it indicates that this technology domain tends to be concentratedly participated by countries with

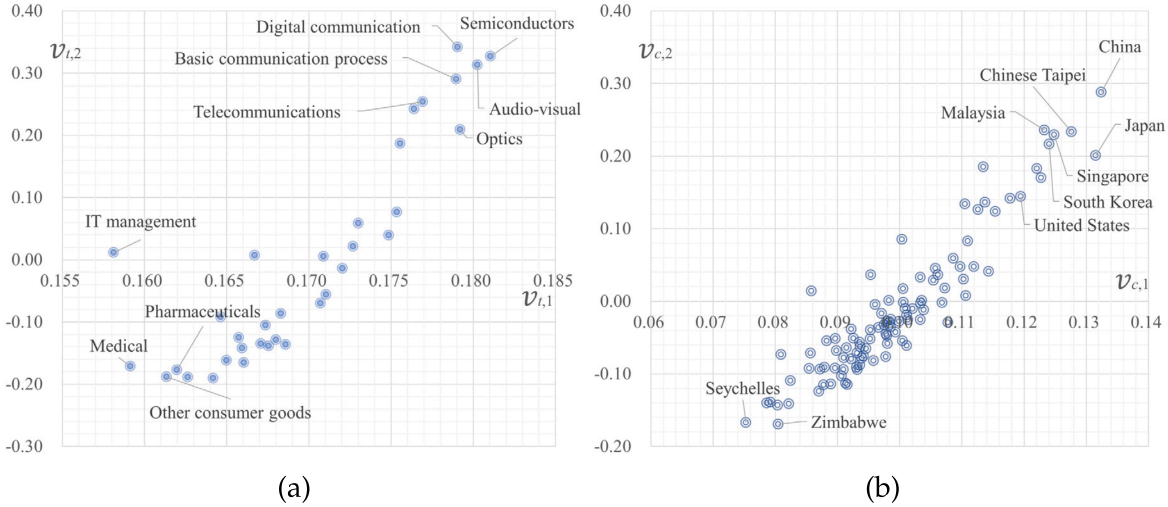


Fig. 2. Scatter plots on $\{v_1^G, v_2^G\}$ and $\{v_1^N, v_2^N\}$ plane in 2019. (a) shows the scatter points of 35 technology domains; (b) shows the scatter points of 100 countries.

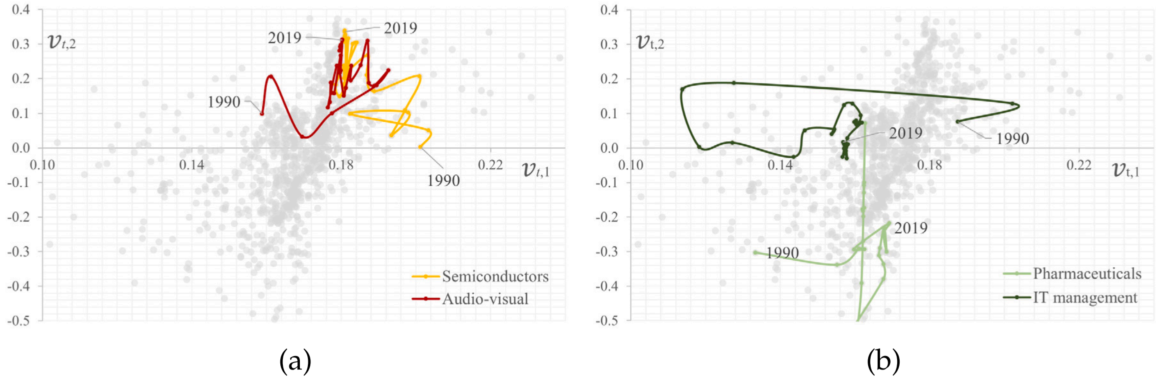


Fig. 3. Trajectories on $\{v_1^G, v_2^G\}$ plane in 1990-2019[1]. (a) shows the trajectories of semiconductors and audio-visual technology; (b) shows the trajectories of pharmaceuticals and IT management. The complete calculation results of v_1^G/v_2^G and v_1^N/v_2^N have been provided in the Supplementary Material.

higher GPYC, and vice versa. v_2^N describes the specialization degree of countries in various technology domains; that is, a higher absolute value of $v_{c,2}^N$ of a certain country indicates a higher degree of specialization in a certain technology domain. In addition, if $v_{c,2}^N$ for a certain country is positive, it tends to have higher specialization in technology domains with higher GPYT, and vice versa.

Fig. 2 shows that technologies such as semiconductors, audio-visual technology and digital communications are scattered in the upper right part of $\{v_1^N, v_2^N\}$ plane, namely v_1^N and v_2^N are relatively higher. It means that the situation of patent outputting of such technology domains is similar to those with high GPYT, and countries with high GPYC participate in these technologies more intensively. Technologies like pharmaceuticals, IT management, and medical technology are located in the lower left corner.

China and Chinese Taipei are located in the upper right corner of the $\{v_1^N, v_2^N\}$ plane, namely values of v_1^N and v_2^N are relatively higher; it means that the structure of patents contribution of China and Chinese Taipei is relatively similar to that of countries with high GPYC (e.g., Korea and Japan), and China and Chinese Taipei focus more on technology domains with high GPYT (e.g., semiconductors and computer technology). The position in which Japan, South Korea, and Singapore are located is similar to China and Chinese Taipei, and a majority of countries in Africa and Latin America are located in the lower left corner, with relatively low values of v_1^N and v_2^N corresponding. Worth mentioning, the US is located in the central part of $\{v_1^N, v_2^N\}$ plane with v_2^N almost equal to 0, which results on low-ranking of GPYC.

In order to explain ranking changes of GPYT and GPYC, this article calculates all the scatters of $\{v_1^N, v_2^N\}$ and $\{v_1^N, v_2^N\}$ plane from 1990 to 2019, and results are shown in Fig. 3 and Fig. 4. For convenience, this article only labeled a few objects, and the links between the scatter plots represent the trajectories of ranking changes from 1990 to 2019.

The scatters of technologies that are considered to be high-tech are mostly located in the upper right part of $\{v_1^G, v_2^G\}$ plane, such as semiconductors and audio-visual technology, especially in recent years. It indicates that these technol-

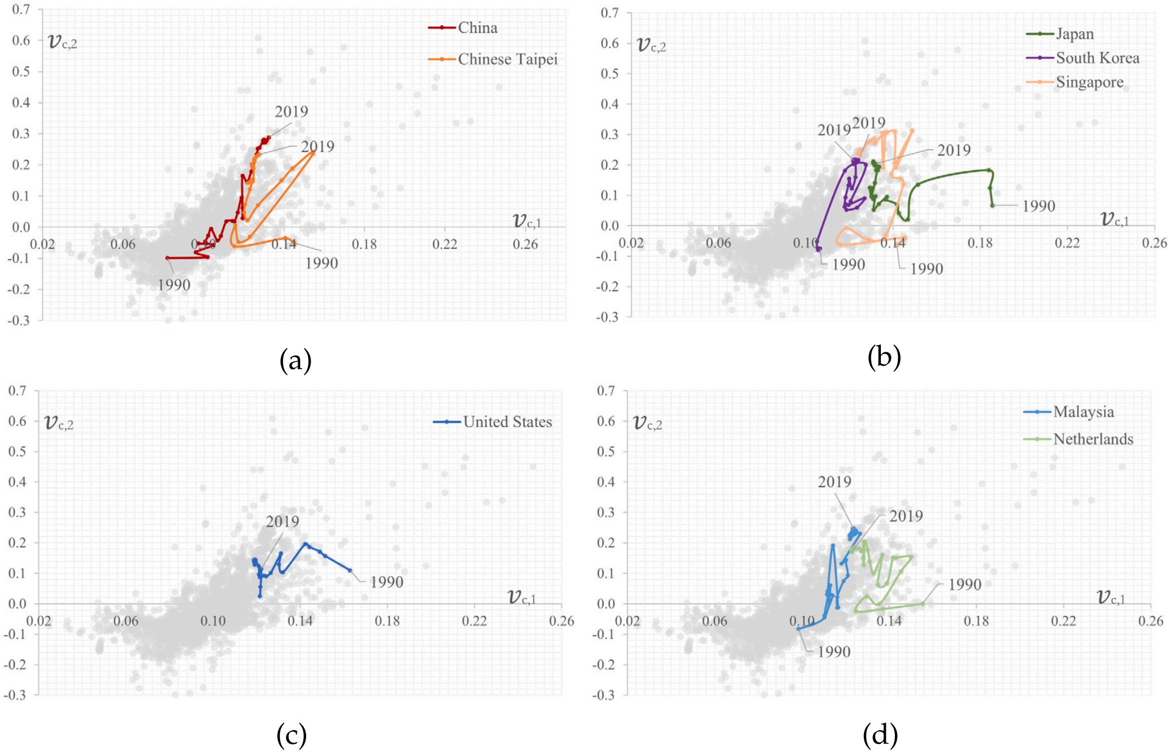


Fig. 4. Trajectories on $\{\boldsymbol{v}_1^N, \boldsymbol{v}_2^N\}$ plane in 1990-2019. (a) shows the trajectories of China and Chinese Taipei; (b) shows the trajectories of Japan, South Korea and Singapore; (c) shows the trajectories of the US; (d) shows the trajectories of Malaysia and the Netherlands. The complete calculation results of $\boldsymbol{v}_1^G/\boldsymbol{v}_2^G$ and $\boldsymbol{v}_1^N/\boldsymbol{v}_2^N$ have been provided in the Supplementary Material.

ogy domains have been participated in by countries with high GPYC steadily and deeply over the past decades. Such domains contain complex knowledge that will be difficult for countries to acquire and master, which means that there will be a rare spillover of complex knowledge taking place in those countries having mastered such technologies; in addition, complex knowledge will be generated, utilized, and renewed step by step, making it hard for other countries to benefit from such complex knowledge. Thus, these technology domains can maintain a high level of GPYT for an extended period.

Technologies like pharmaceuticals and IT management are highly ranked in the early stage, but their GPYT rapidly declined as time passed. For instance, the $v_{t,2}^G$ of IT management has remained relatively stable for a long period, but its $v_{t,1}^G$ value has undergone significant changes —its $v_{t,1}^G$ exceeded that of most technology domains in 1990 but reached a low level in 2019. While pharmaceuticals have a relatively low $v_{t,1}^G$ value and a high negative $v_{t,2}^G$ value, indicating that it has been continuously dominated by countries with low GPYC for a long time.

The scatters corresponding to China and Chinese Taipei move gradually closer to the upper part of $\{\boldsymbol{v}_1^N, \boldsymbol{v}_2^N\}$ plane and tend to be stable eventually, which explains their rapid rise in GPYC rankings. In terms of $\boldsymbol{v}_2^N$, they have shifted their focus from domains with low GPYT to those with higher GPYT, especially after 2000. In fact, as one of the “Four Asian Tigers”, Chinese Taipei has long made significant progress in the fields of semiconductors and computer technology, and has established forceful industrial advantages. Benefiting from the “Reform and Opening-up”, China has gradually established a globally unparalleled full-industry-chain and high industrial barriers in the high-tech fields (e.g., electrical apparatus and computer technology), making its capabilities on scientific and technological innovation improved rapidly. It is also consistent with the patent statistics data, as Chinese Taipei and China began to produce a large number of patents in semiconductors and computer technology around 1995 and after 2000, respectively.

The scatters of Japan, South Korea and Singapore have been located in the right part of $\{\boldsymbol{v}_1^N, \boldsymbol{v}_2^N\}$ plane for a long time, and relatively high $v_{c,1}^N$ values make their GPYCs rank at the top. It indicates that they have consistently maintained a high level of focus on fields such as semiconductors and audio-visual technology over the past 30 years. Due to the high costs of entering these fields, substantial industrial barriers have been formed, allowing these countries to maintain their level of GPYC.

The US moves from the upper right quarter of $\{\boldsymbol{v}_1^N, \boldsymbol{v}_2^N\}$ plane to the central part, namely the values of $v_{c,1}^N$ and $v_{c,2}^N$ are both decline. Based on the patent data, no concentration can be found in specific technology domains; on the contrary, the US contributes

Table 3
GPYT (10-year) results[1] of technology domains in 2019[2].

GPYT Ranking (10-year)	Technology Domain	GPYT (10-year)[3]	Patents	Patents Ranking
1	Semiconductors	1.0000	1.000	1
2	Telecommunications	0.9695	0.896	6
3	Audio-visual technology	0.9388	0.980	2
4	Digital communication	0.9324	0.959	3
5	Basic communication processes	0.9225	0.949	4
31	Other special machines	0.2388	0.648	24
32	IT methods for management	0.1567	0.480	34
33	Food chemistry	0.1459	0.579	30
34	Pharmaceuticals	0.1373	0.565	31
35	Other consumer goods	0.0000	0.553	32

1 The full results of GPYT (10-year) have been provided in the Appendix Table A3.
2 The complete results of GPYT (10-year) from 1990-2019 have been provided in the Supplementary Material.
3 NOTE: GPYTs have been Min-Max Normalized.

a large number of patents in all 35 domains, which disperses its comparative advantages on technologies with high GPYT and leaves the industrial barriers that do benefits unestablished.

The situation in Malaysia is similar to that in China, which has continued to rise in GPYC ranking in the past 30 years. In the past decade, Malaysia has benefited from low labor costs and the transference of industrial chains from developed countries, meaning Malaysia can acquire and master advanced technologies through continuous technology spillovers and achieve technological innovation and replacement. In fact, in the fields of semiconductor and electrical apparatus, the patent output showed a significant increase around 2000. Notably, patents of such two fields account for a large proportion of all fields, indicating a high degree of specialization, which corresponds to a high $v^N_{c,2}$ value.

The situation in the Netherlands is similar to those of Japan, South Korea, and Singapore. As an old developed country in Europe, the advantage of the Netherlands in terms of knowledge complexity is due to its well-known photolithography technology. In the past 30 years, the Netherlands has paid high attention to and developed its "monopolistic" technology, which is also evidenced by its relatively high $v^N_{c,1}$ value and the upward trend of $v^N_{c,2}$ value.

4.2. Different "expirations"

However, considering that technological innovation usually has its life cycle, which means that countries around the world will generally master the innovation after a certain period of its release, and then it will be seen to no longer contribute to the knowledge complexity of the field. Therefore, this article will change the statistical method by setting different years as "expiration" to re-calculate GPYTs/GPYCs.

Above all, this article will take a 10-year "expiration" as an example and conduct an analysis. In detail, this article will re-construct the bipartite network of each year based on Eq. (18) by defining $P_{ct,y} = \sum_{i=y-exp}^{y} N_{ct,i}$, and then re-calculate GPYTs/GPYCs. Table 3 and Table 4 displays some of the calculation results.

Actually, for most technology domains, there is not much difference between the GPYT ranking obtained by considering "expiration" and the results obtained by assuming long-term "validity". For example, the fields of semiconductors, audiovisual technology, and digital communications still rank in the top 5, while the fields of IT management, pharmaceuticals, and other consumer goods still rank last; only a few fields have experienced fluctuations in their rankings.

It will be slightly different while referring to countries or regions. In fact, only a few countries or regions with high levels of technological development and high knowledge complexity, such as China, Japan, South Korea, and the US, have remained relatively the same in their GPYC rankings. The rankings of most countries do not match the results obtained in the previous calculations, especially for countries initially ranked lower, where the GPYC rankings show significant differences.

This article still constructs $\{\boldsymbol{v}^{\boldsymbol{G}}_1, \boldsymbol{v}^{\boldsymbol{G}}_2\}$ and $\{\boldsymbol{v}^{\boldsymbol{N}}_1, \boldsymbol{v}^{\boldsymbol{N}}_2\}$ plane and displays the trajectories of countries (or regions) in the $\{\boldsymbol{v}^{\boldsymbol{G}}_1, \boldsymbol{v}^{\boldsymbol{G}}_2\}$ plane from 1990 to 2019, as shown in Fig. 5 and Fig. 6.

The trajectory of China and Chinese Taipei on $\{\boldsymbol{v}^{\boldsymbol{N}}_1, \boldsymbol{v}^{\boldsymbol{N}}_2\}$ plane is consistent with the results obtained in the previous text. It means that China and Chinese Taipei have a sustained and stable output of patents at the forefront of high-techs, which helps them maintain high rankings. The situation of Malaysia is generally similar to that of China and Chinese Taipei.

The trajectories of Japan and South Korea do not differ essentially from results in the previous text, but that of Singapore shows a more significant retreat from the upper right area of $\{\boldsymbol{v}^{\boldsymbol{N}}_1, \boldsymbol{v}^{\boldsymbol{N}}_2\}$ plane to the center. It indicates Singapore's specialization in technologies with high GPYT has decreased significantly in recent years. The statistical data of patents also show that industrial advantages of Singapore in high-techs are mainly derived from early large-scale outputting and rapid accumulation. However, in recent years, the output in such fields has slowed down, and the development of other industries has gradually dispersed its original comparative advantages in high-techs.

The trajectory of the US is almost identical to the results obtained in the previous text, which also indicates that the US has had a relatively stable outputting of patents in various technology domains for a long time, and the structure of

Table 4
GPYC (10-year) results[1] of countries or regions in 2019[2].

GPYC Ranking (10-year)	Country or Region	GPYC (10-year)[3]	Patents	Patents Ranking
1	China	1.0000	1.0000	1
2	Japan	0.8618	0.9598	2
3	Chinese Taipei	0.8550	0.9035	3
4	Malaysia	0.8500	0.8326	6
5	Korea	0.8258	0.8410	5
6	Finland	0.7571	0.8036	8
7	Sweden	0.7474	0.6929	11
8	Singapore	0.7195	0.8544	4
9	United States	0.7165	0.7560	9
10	Pakistan	0.6761	0.6736	12
91	Jordan	0.3027	0.4344	56
92	Uzbekistan	0.2998	0.3279	87
93	Zimbabwe	0.2955	0.2668	94
94	Nigeria	0.2827	0.2568	96
95	Bermuda	0.2630	0.3078	90
96	Armenia	0.2493	0.3781	74
97	El Salvador	0.2236	0.2770	93
98	Seychelles	0.2145	0.2147	99
99	Jamaica	0.1871	0.3844	71
100	Djibouti	0.0000	0.0000	100

1 The full results of GPYC (10-year) have been provided in the Appendix Table A4.
2 The complete results of GPYC (10-year) from 1990-2019 have been provided in the Supplementary Material.
3 NOTE: GPYCs have been Min-Max Normalized.

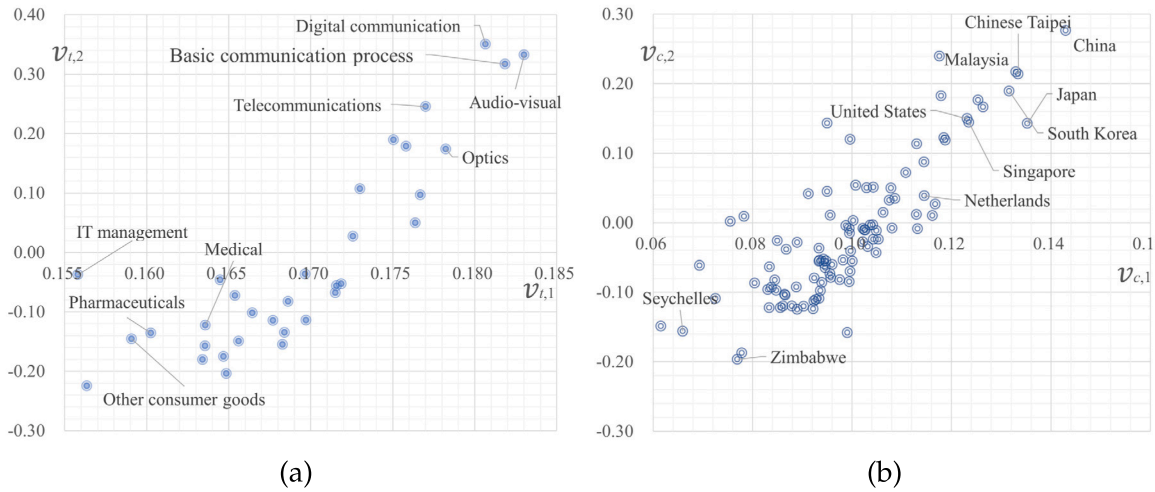


Fig. 5. Scatter plots (10-year) on $\{v_1^G, v_2^G\}$ and $\{v_1^N, v_2^N\}$ plane in 1990-2019. (a) shows the scatter points of 35 technology domains; (b) shows the scatter points of 100 countries.

patents has hardly changed significantly. However, the situation of the Netherlands is worse than that of Singapore, continuously retreating from the right side of $\{v_1^N, v_2^N\}$ plane to the central area. From the patent statistics, the Netherlands had a considerable volume of patent outputting in audio-visual technology and optics before 2005; but in recent years, such outputting has rapidly declined, leading to a rapid dissipation of its comparative advantages in technologies with high GPYT.

Luxembourg and Lebanon show a more obvious movement from the lower left corner to the upper right corner on $\{v_1^N, v_2^N\}$ plane. It means that both of them had insufficient output in domains with high GPYT in the early stage; but in recent years, they have made significant breakthroughs and formed a certain degree of specialization in specific domains.

Conducting analyses just based on setting a 10-year "expiration" as a case is not enough; thus, this article calculates GPYCs for different years under 10 different "expirations" to compare its impact on the knowledge complexity of different countries. Fig. 7 shows the changes in GPYC rankings of some countries (or regions) under different "expirations" from 2000 to 2019.

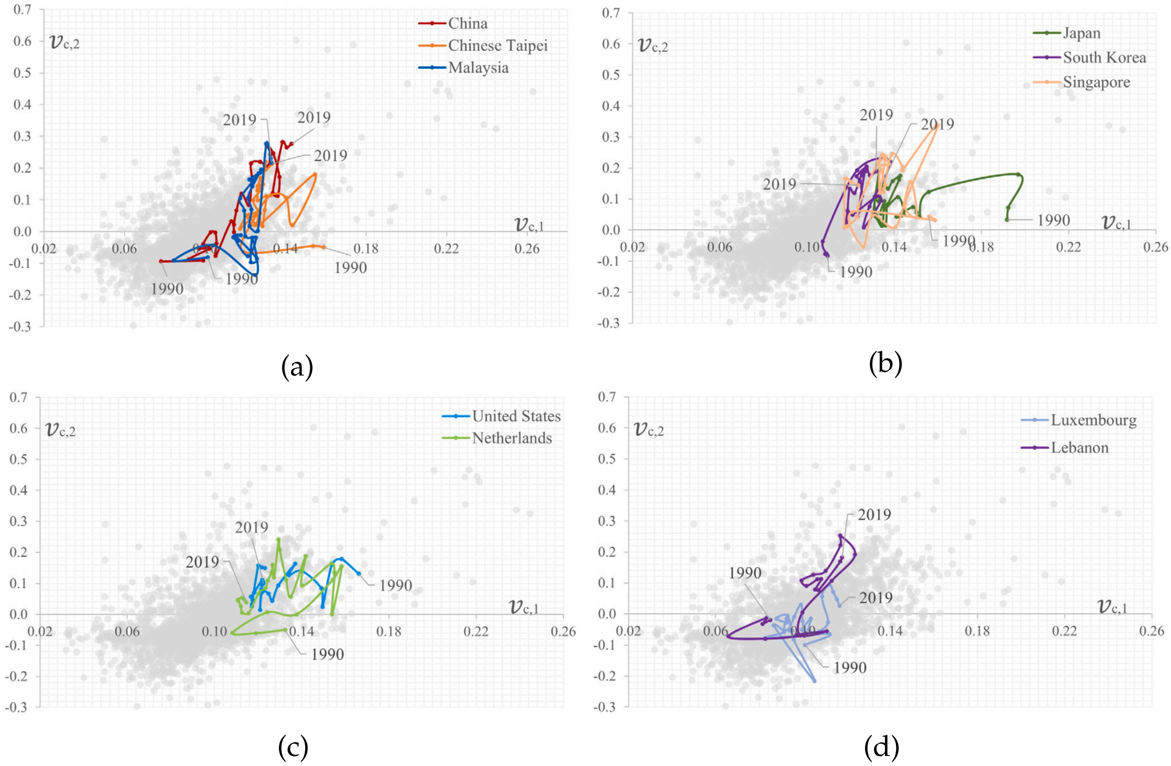


Fig. 6. Trajectories (10-year) on $\{\boldsymbol{v}_1^N, \boldsymbol{v}_2^N\}$ plane in 1990-2019. (a) shows the trajectories of China, Chinese Taipei and Malaysia; (b) shows the trajectories of Japan, South Korea and Singapore; (c) shows the trajectories of the US and the Netherlands; (d) shows the trajectories of Luxembourg and Lebanon. The complete calculation results of $\boldsymbol{v}_1^G/\boldsymbol{v}_2^G$ and $\boldsymbol{v}_1^N/\boldsymbol{v}_2^N$ have been provided in the Supplementary Material.

From the perspective of China and Chinese Taipei, the common feature is that their GPYC rankings steadily increase by assuming long-term “validity”, and those obtained based on different “expirations” exhibit relatively synchronized changes. It confirms what was mentioned in the previous text, that China and Chinese Taipei have stable and continuous high-tech patent outputting.

The situations in Japan and the US are similar. Their feature in common is that regardless of whether giving long-term “validity” or different “expirations”, their GPYC rankings almost do not fluctuate. It means that they have stable outputs in their specific domains, respectively, and the structure of patents has not changed over the long term.

Singapore and the Netherlands are similar. Their GPYC rankings obtained by assuming long-term “validity” show a slow downward trend; but those based on different “expirations” show a more apparent declining trend. It indicates that their knowledge complexity is more dependent on the large-scale outputting in frontier technology domains in the early stage. However, as time goes by, their capabilities to re-generate complex knowledge in such domains decline gradually, changing the patent structure. Moreover, at the same time, the early accumulated complex knowledge no longer contributes to their knowledge complexity due to the spillover effect, resulting in significant differences in the calculation results under different “expirations”.

The situation in Malaysia is similar to that of Pakistan. Although both countries rank relatively high in GPYC rankings while assuming long-term “validity”, their GPYC rankings based on different “expirations” fluctuate significantly in different years. It implies that their knowledge complexity depends more on the transference of industrial chains and technology spillovers from developed countries. Also, due to the relatively low total patent outputting in all domains, the concentration in specific fields in a certain year can affect their patent structure significantly, leading to significant changes in GPYC. In this regard, most countries in Africa and Latin America are generally the same, with their GPYCs and rankings often experiencing “boom and bust” cycles, which also confirms the previous discussion about the inconsistent results for African and Latin American countries under different calculations.

5. Community detection

Previous studies have shown that complex knowledge exhibits “spatial stickiness” [7,43], which hinders its diffusion among regions, leading to certain locational differences, and forming certain geographical patterns [16,27]. However, with the progress of globalization, the spillover effect of complex knowledge may be strengthened, leading to a reduction in its spatial heterogeneity [9,32].

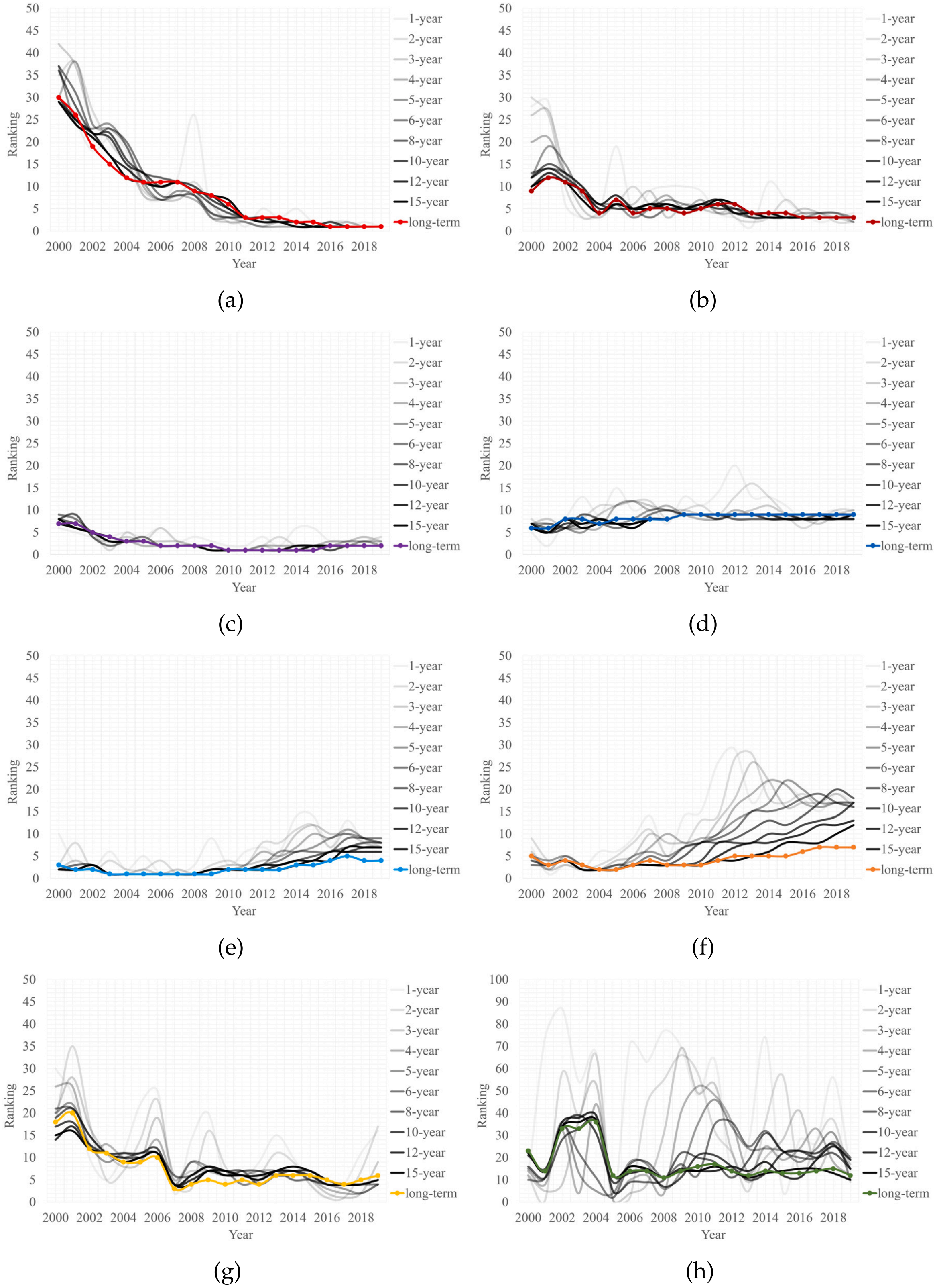


Fig. 7. GPYCs under different “expiration” from 2000-2019. (a) and (b) show the results of China and Chinese Taipei; (c) and (d) show the results of Japan and the US; (e) and (f) show the results of Singapore and the Netherlands; (g) and (h) show the results of Malaysia and Pakistan. The complete calculation results of GPYCs under different “expiration” from 2000-2019 have been provided in the Supplementary Material.

Therefore, this paper will analyze the location of complex knowledge and its pattern or structure through community detection. Specifically, this article will conduct community detection by combining the Weighted Extremal Optimization (WEO) algorithm and Co-occurrence Matrix based on the weighted undirected network corresponding to the proximity matrix $\boldsymbol{N}$.

As mentioned above, the matrix $\boldsymbol{N}$ can be seen as an adjacency matrix corresponding to a weighted network that describes some kind of similarity in the patent structure. Such a structure refers to how countries participate in technology domains, and it can be calculated by: $N_{cc'} = \sum_{c'} M_{ct} M_{c't} / (s_t')^2 s_{c'} s_c$.

For the WEO algorithm, community detection is achieved by maximizing the weighted modularity Q^w [12]:

$$\begin{aligned} \mathbb{r} &= arg\min_{\mathbb{r}} Q^w_{\mathbb{r}} = arg\min_{\mathbb{r}} \sum_{\mathbb{r}} (e^w_{\mathbb{r}\mathbb{r}} - (a^w_{\mathbb{r}})^2) \\ &= arg\min_{\mathbb{r}} \sum_{\mathbb{r}} (\sum_{i,j\in\mathbb{r},i\neq j} s_{ij} / \sum_{i,j} s_{ij} - (\sum_{i\in\mathbb{r},i\neq j} s_{ij} / \sum_{i,j} s_{ij})^2) \end{aligned} \tag{19}$$

where, $\mathbb{r}$ denotes communities; s_{ij} denotes the weighted edge between vertex i and j; k^w_i denotes the weighted degree of vertex i. To be more specific, community detection can be completed by steps following:

•**Step 1.** Split the vertex of the network derived from $\boldsymbol{N}$ into two random partitions having the same scale, and calculate the fitness λ^w_i of each vertex, and the initial modularity Q^w_0. The fitness of the vertex can be defined by:

$$\lambda^w_i = \frac{\kappa^w_{\mathbb{r}}(i)}{k^w_i} - a^w_{\mathbb{r}}(i) = \sum_{j\in\mathbb{r},i\neq j} s_{ij} / k^w_i - \sum_{i\neq j} s_{ij} / \sum_{i,j} s_{ij} \tag{20}$$

•**Step 2.** Move the vertex with the lowest fitness from one partition to the other, and re-calculate the fitness λ^w_i of each vertex and the modularity Q^w_1.

•**Step 3.** If $Q^w_0 < Q^w_1$, repeat Step 2; while if $Q^w_0 > Q^w_1$, repeat Step 1.

The process of Step 1 to Step 3 will be repeated, and the community detection will finish when the modularity Q^w cannot be further improved.

Actually, it is inevitable that the community detection result based on the WEO algorithm is stochastic, because partitions of the vertex are split randomly, as mentioned in Step 1. For the purpose of diminishing such stochasticity, this article constructs the Co-occurrence Matrix and re-conducts the community detection. Specifically, this process can be completed by:

•**Step 4.** Conduct partitions 100 times based on proximity matrix $\boldsymbol{N}$ and construct the initial Co-occurrence Matrix $\boldsymbol{C}^{(1)} = \{c^{(1)}_{ij}\}$. Thereinto, $c^{(1)}_{ij}$ denotes the probability of vertex i and j being classified into the same partition.

•**Step 5.** If all the elements of $\boldsymbol{C}^{(1)}$ are equal to 0 or 100, it implies that the community detection result based on proximity matrix **N** is distinct, and no more detection needs to be conducted. Otherwise, re-conduct partitions 100 times based on $\boldsymbol{C}^{(1)}$ and re-construct the Co-occurrence Matrix.

•**Step 6.** Repeat Step 5, and $\boldsymbol{C}^{(k)}$ can be obtained, where k means the times of Step 5 being conducted.

Similar to Step 5, Step 6 will be finished when all the elements of $\boldsymbol{C}^{(k)}$ are equal to 0 or 100. Then, the final and distinct result of community detection based on $\boldsymbol{N}$ will be the same as the result of detecting $\boldsymbol{C}^{(k)}$ one more time.

Assuming long-term "validity", this article conducts community detection based on the proximity matrix $\boldsymbol{N}$ from 1990 to 2019 based on Step 1 to Step 6. Results are displayed in Fig. 8.

The results of 1990 showed that countries or regions could be definitely divided into two principal communities — the North American and European communities. The production of complex knowledge was mainly concentrated in the representative regions of these two communities (e.g., the US and Germany), which indicated that the "location" of complex knowledge is explicit.

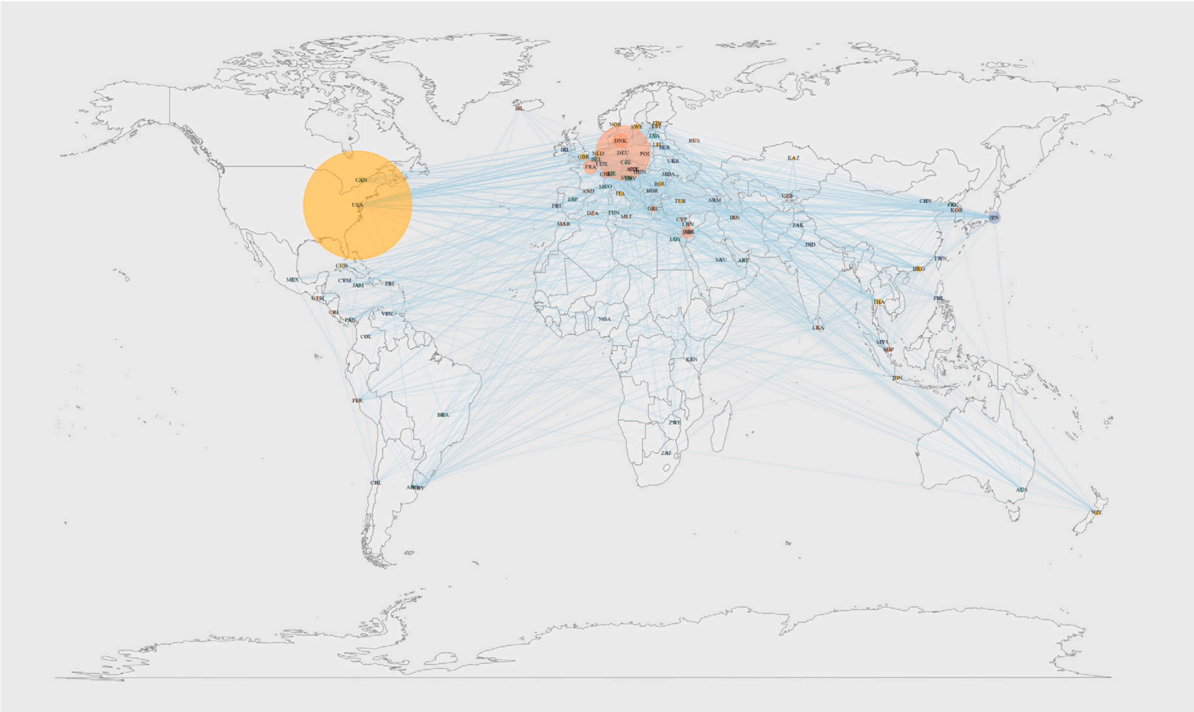
In 2000, in addition to the North American and European communities, the Asian region was also separately divided into two communities — the Southeast Asia community, mainly led by Malaysia and Singapore, and the East Asia community, mainly led by China and Japan. The community structure at this time still suggested "location" to some extent, but was slightly weaker than 1990. In particular, the Middle East region, mainly led by Lebanon and Israel, was also classified into the same community as Europe, indicating a relatively consistent patent outputting situation between these two regions.

The results of 2010 showed that the production of complex knowledge in the East Asia and Southeast Asia communities continued to grow, and the heterogeneity of location further weakened. For example, some European countries were classified into such two communities, and the community of North America could not be detected definitely.

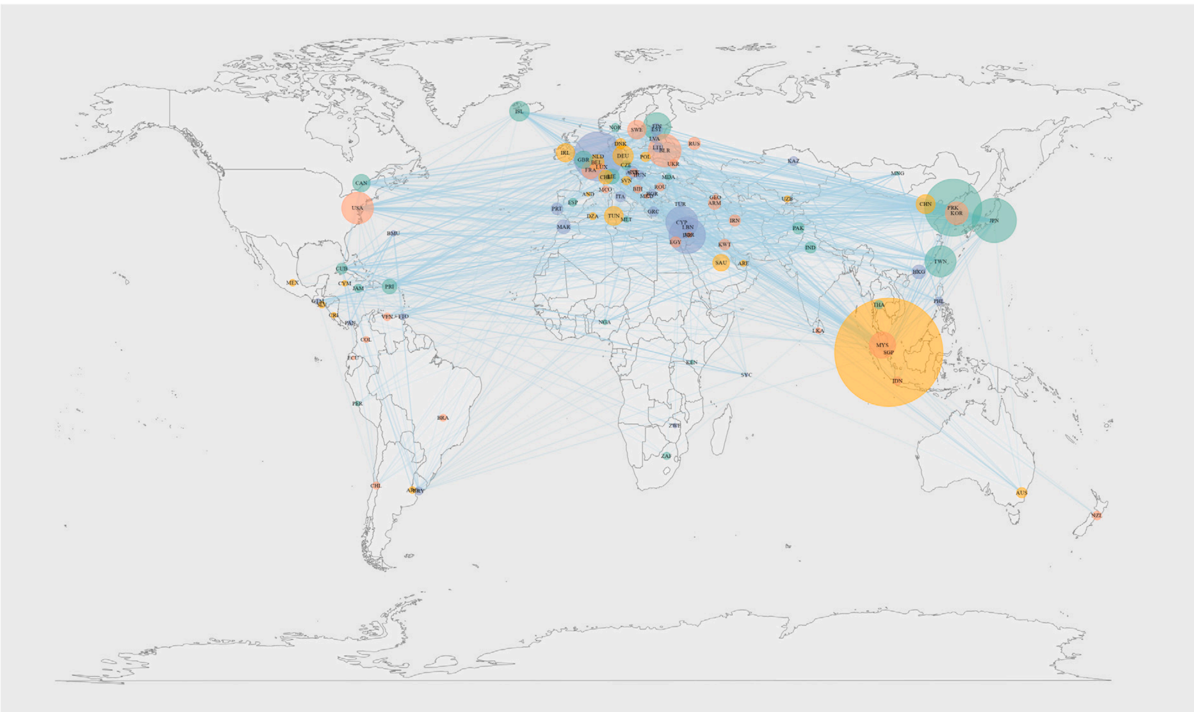
Meanwhile, in 2019, the locational heterogeneity of complex knowledge was scarcely significant. For instance, China and South Korea, as well as Chinese Taipei and Japan, which were initially in two communities, were divided into four communities, respectively, and no prominent community could be detected definitely in Europe.

6. Conclusion

This article quantifies the knowledge complexity of technologies and countries (or regions) by combining the FC algorithm matrix-estimation exercise, re-calculates knowledge complexity given patents' "expiration", and finally conducts commu-

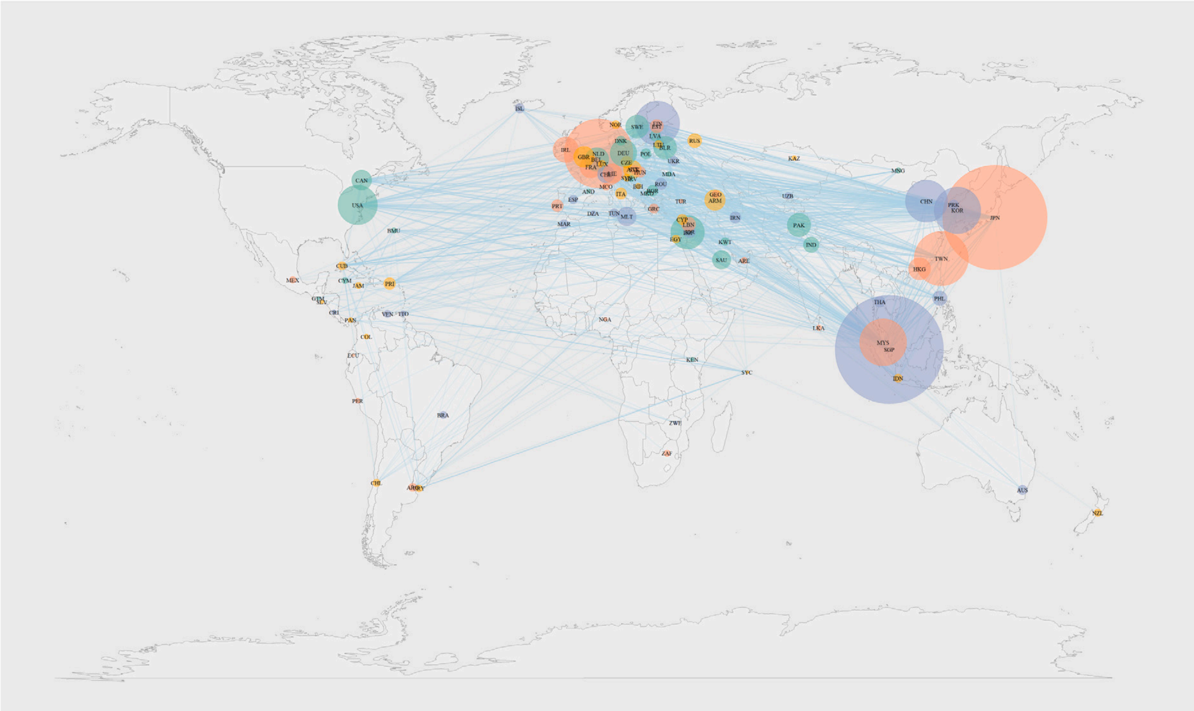


(a)

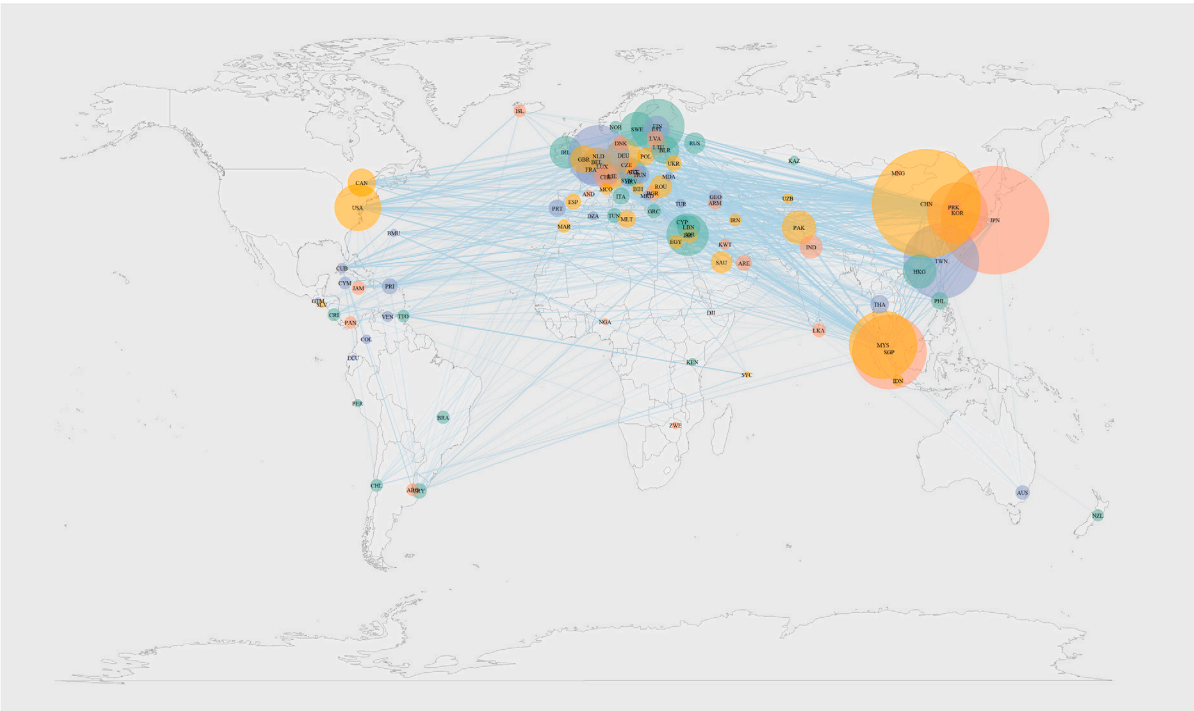


(b)

Fig. 8. Results of community detection. (a), (b), (c) and (d) show the results of community detection in 1990, 2000, 2010 and 2019 respectively.



(c)



(d)

Fig. 8. (*continued*)

nity detection based on the WEO algorithm and Co-occurrence Matrix. In fact, the generalized knowledge complexity (GPYC and GPYT) is the "density" of complex knowledge rather than the total amount. A higher GPYC/GPYT of a country/technology indicates that such a country/technology contains more complex knowledge within the same patent output and vice versa. Alternatively, to put it more vividly, "knowledge complexity" means that a country can produce more and denser complex knowledge compared to other countries. It stems from the country's "outperformance" in patent outputs within specific fields relative to other countries, thereby bringing the country "industrial advantages" for creating more economic value.

The results of this article show that the GPYC of a country just depends on the similarity of patent structure among countries, as well as the specialization of the country in technology domains. In addition, if the "expiration" of patents is considered, the GPYC of countries with stable outputting of high-tech patents does not change much. However, if a country overly relies on the advantages formed by patent outputting and technological accumulation in high-tech fields in the early stage, its knowledge complexity will be customarily weakened gradually. For countries that have not formed a complete industrial chain, their knowledge complexity will fluctuate significantly in different years. In addition, the results of community detection imply the decrease of the "location" property of complex knowledge.

Based on the above conclusions, the following insights can be obtained: First, the knowledge complexity of a country will be jointly determined by various countries worldwide. Overly concentrated participation in a certain domain will cause the complexity of such domain to change considerably, thereby affecting the expected effects of these countries participating in that. Second, improving the knowledge complexity of a country can be achieved by participating in domains with higher complexity, or by achieving higher specialization in a certain domain. Therefore, achieving division of labor in various domains is the best way to improve the global knowledge complexity. As for different countries, they should choose which domains to participate in or achieve specialization based on their actual development status and resource endowments in order to improve their knowledge complexity effectively. In addition, what a country can do to maintain knowledge complexity is to pay more attention to realizing continuous and stable outputs of cutting-edge technologies and achieving updating and iteration of technology based on existing complex knowledge.

CRediT authorship contribution statement

Shenshen Sergio Zhang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The author declared no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Data availability

The data that support the findings of this article are available from the author upon request.

Code availability

The codes for the computations in this article are available from the author upon request.

Acknowledgement

The author expresses the most sincere gratitude to the editor and reviewers for their valuable and professional comments and advice, and to Prof. Dr. Mingqian Zhang and Siwu (Paul) Zhang for their excellent research assistance.

Appendix A. Long-term "validity"

Due to the computation of GPYT and GPYC involving 35 technical domains and 100 countries (or regions), displaying all the results will generate excessively lengthy tables. Therefore, this article only displays partial results of important technology domains and countries (or regions) in the main text, while presenting the full computation results here. Under the assumption of Long-term "validity" for patents, the GPYT for 35 technology domains and the GPYC for 100 countries (or regions) in 2019 are shown in Table A1 and Table A2 respectively. The complete results from 1990 to 2019 will be provided in the Supplementary Material.

Appendix B. 10-year "expiration" results

Under the assumption of 10-year "expiration" for patents, the GPYT (10-year) for 35 technology domains and the GPYC (10-year) for 100 countries (or regions) in 2019 are shown in Table A3 and Table A4 respectively. The complete results from 1990 to 2019 will be provided in the Supplementary Material.

Table A1
Full GPYT results of 35 technology domains in 2019.

GPYT ranking	Technology Domain	GPYT	Patents	Patents Ranking
1	Semiconductors	1.0000	122140	11
2	Audio-visual technology	0.9802	146219	9
3	Digital communication	0.9593	265260	3
4	Basic communication processes	0.9487	32440	34
5	Optics	0.9410	118609	12
6	Telecommunications	0.8965	116120	13
7	Computer technology	0.8810	285306	1
8	Micro-structural and nano-technology	0.8534	6197	35
9	Macromolecular chemistry, polymers	0.8428	86223	21
10	Measurement	0.8308	178011	6
11	Textile and paper machines	0.7927	60170	29
12	Electrical machinery, apparatus, energy	0.7834	251200	4
13	Surface technology, coating	0.7700	66337	28
14	Machine tools	0.7482	77891	24
15	Control	0.7449	67855	27
16	Chemical engineering	0.7394	97631	20
17	Mechanical elements	0.6991	105329	17
18	Thermal processes and apparatus	0.6901	58351	30
19	Transport	0.6865	158388	7
20	Handling	0.6769	102590	19
21	Organic fine chemistry	0.6689	139063	10
22	Materials, metallurgy	0.6665	75283	25
23	Analysis of biological materials	0.6613	40520	32
24	Other special machines	0.6479	111897	15
25	Basic materials chemistry	0.6428	111162	16
26	Furniture, games	0.6383	81831	22
27	Engines, pumps, turbines	0.6260	103654	18
28	Biotechnology	0.6161	146587	8
29	Civil engineering	0.6111	112649	14
30	Food chemistry	0.5786	40150	33
31	Pharmaceuticals	0.5655	212699	5
32	Other consumer goods	0.5532	81267	23
33	Medical technology	0.5079	269729	2
34	IT methods for management	0.4796	68747	26
35	Environmental technology	0.0000	51273	31

Table A2
Full GPYC results of 100 countries in 2019.

GPYC Ranking	Country or Region	GPYC	Patents	Patents Ranking
1	China	1.0000	376819	4
2	Japan	0.9598	698133	2
3	Chinese Taipei	0.9035	9881	26
4	Singapore	0.8544	13271	23
5	Korea	0.8410	190977	5
6	Malaysia	0.8326	4377	34
7	Netherlands	0.8112	80280	8
8	Finland	0.8036	37752	15
9	United States	0.7560	1204578	1
10	Israel	0.7305	42187	14
11	Sweden	0.6929	79591	9
12	Pakistan	0.6736	102	85
13	Hong Kong, PRC	0.6699	8010	30
14	Germany	0.6694	415829	3
15	Ireland	0.6520	8391	28
16	France	0.6355	162710	6
17	Lebanon	0.6254	168	75
18	Canada	0.6241	67431	11
19	Belgium	0.6159	26902	20
20	Austria	0.6118	30246	19
21	United Kingdom	0.6048	162059	7
22	Belarus	0.5900	483	60
23	Liechtenstein	0.5771	561	56
24	India	0.5693	33808	16
25	Switzerland	0.5632	55159	12
26	Estonia	0.5564	676	55
27	Romania	0.5516	1171	45

(*continued on next page*)

Table A2 (*continued*)

GPYC Ranking	Country or Region	GPYC	Patents	Patents Ranking
28	Russia	0.5442	22389	21
29	Luxembourg	0.5254	1426	43
30	Saudi Arabia	0.5188	4149	35
31	Malta	0.5172	182	71
32	Hungary	0.5172	6071	31
33	Thailand	0.5171	1524	42
34	Lithuania	0.5017	492	59
35	Poland	0.4916	5736	32
36	Czechia	0.4907	4060	36
37	Portugal	0.4893	2813	37
38	Puerto Rico	0.4884	139	79
39	Sri Lanka	0.4826	274	66
40	Philippines	0.4818	678	54
41	Italy	0.4816	71275	10
42	United Arab Emirates	0.4814	962	49
43	Spain	0.4729	32882	17
44	Latvia	0.4656	502	58
45	Uruguay	0.4611	188	70
46	Iran	0.4544	699	53
47	Bulgaria	0.4530	939	50
48	Cyprus	0.4523	252	68
49	Slovakia	0.4514	1096	48
50	Denmark	0.4500	30691	18
51	Korea, DPR	0.4489	91	86
52	Greece	0.4484	2431	39
53	Australia	0.4482	49830	13
54	Ukraine	0.4440	2749	38
55	Egypt	0.4389	924	51
56	Jordan	0.4344	130	82
57	Georgia	0.4337	169	74
58	Brazil	0.4250	11776	25
59	Tunisia	0.4170	174	72
60	Cayman Islands	0.4158	110	84
61	Morocco	0.4157	538	57
62	Norway	0.4102	17433	22
63	Chile	0.4052	1970	41
64	Slovenia	0.4027	2357	40
65	Iceland	0.4009	822	52
66	Costa Rica	0.3989	212	69
67	Bosnia and Herzegovina	0.3977	143	78
68	Argentina	0.3942	1160	47
69	Trinidad and Tobago	0.3939	58	91
70	New Zealand	0.3937	8129	29
71	Jamaica	0.3844	24	98
72	Mexico	0.3825	4721	33
73	Monaco	0.3785	423	61
74	Armenia	0.3781	169	73
75	Indonesia	0.3743	370	63
76	Venezuela	0.3730	148	77
77	Peru	0.3702	305	64
78	South Africa	0.3689	8401	27
79	Croatia	0.3674	1164	46
80	Turkey	0.3547	12552	24
81	Algeria	0.3539	152	76
82	Cuba	0.3499	256	67
83	Kuwait	0.3452	62	89
84	North Macedonia	0.3368	87	87
85	Colombia	0.3353	1366	44
86	Kazakhstan	0.3339	375	62
87	Uzbekistan	0.3279	62	90
88	Moldova	0.3259	114	83
89	Panama	0.3080	280	65
90	Bermuda	0.3078	51	92
91	Kenya	0.3072	131	81
92	Andorra	0.2797	81	88
93	El Salvador	0.2770	27	96
94	Zimbabwe	0.2668	36	95
95	Mongolia	0.2638	18	99

Table A2 (*continued*)

GPYC Ranking	Country or Region	GPYC	Patents	Patents Ranking
96	Nigeria	0.2568	45	93
97	Guatemala	0.2493	45	94
98	Ecuador	0.2439	137	80
99	Seychelles	0.2147	26	97
100	Djibouti	0.0000	0	100

Table A3
Full GPYT (10-year) results of 35 technology domains in 2019.

GPYT ranking (10-year)	Technology Domain	GPYT (10-year)	Patents	Patents Ranking
1	Semiconductors	1.0000	122140	11
2	Audio-visual technology	0.9802	146219	9
3	Digital communication	0.9593	265260	3
4	Basic communication processes	0.9487	32440	34
5	Optics	0.9410	118609	12
6	Telecommunications	0.8965	116120	13
7	Computer technology	0.8810	285306	1
8	Micro-structural and nano-technology	0.8534	6197	35
9	Macromolecular chemistry, polymers	0.8428	86223	21
10	Measurement	0.8308	178011	6
11	Textile and paper machines	0.7927	60170	29
12	Electrical machinery, apparatus, energy	0.7834	251200	4
13	Surface technology, coating	0.7700	66337	28
14	Machine tools	0.7482	77891	24
15	Control	0.7449	67855	27
16	Chemical engineering	0.7394	97631	20
17	Mechanical elements	0.6991	105329	17
18	Thermal processes and apparatus	0.6901	58351	30
19	Transport	0.6865	158388	7
20	Handling	0.6769	102590	19
21	Organic fine chemistry	0.6689	139063	10
22	Materials, metallurgy	0.6665	75283	25
23	Analysis of biological materials	0.6613	40520	32
24	Other special machines	0.6479	111897	15
25	Basic materials chemistry	0.6428	111162	16
26	Furniture, games	0.6383	81831	22
27	Engines, pumps, turbines	0.6260	103654	18
28	Biotechnology	0.6161	146587	8
29	Civil engineering	0.6111	112649	14
30	Food chemistry	0.5786	40150	33
31	Pharmaceuticals	0.5655	212699	5
32	Other consumer goods	0.5532	81267	23
33	Medical technology	0.5079	269729	2
34	IT methods for management	0.4796	68747	26
35	Environmental technology	0.0000	51273	31

Table A4
Full GPYC (10-year) results of 100 countries in 2019.

GPYC Ranking (10-year)	Country or Region	GPYC (10-year)	Patents	Patents Ranking
1	China	1.0000	1.0000	1
2	Japan	0.8618	0.9598	2
3	Chinese Taipei	0.8550	0.9035	3
4	Malaysia	0.8500	0.8326	6
5	Korea	0.8258	0.8410	5
6	Finland	0.7571	0.8036	8
7	Sweden	0.7474	0.6929	11
8	Singapore	0.7195	0.8544	4
9	United States	0.7165	0.7560	9
10	Pakistan	0.6761	0.6736	12
11	Lebanon	0.6643	0.6254	17
12	Israel	0.6620	0.7305	10
13	Hong Kong, PRC	0.6590	0.6699	13
14	Luxembourg	0.6310	0.5254	29
15	Germany	0.6250	0.6694	14
16	Canada	0.6114	0.6241	18

(*continued on next page*)

Table A4 (*continued*)

GPYC Ranking	Country or Region	GPYC	Patents	Patents Ranking
17	Netherlands	0.6082	0.8112	7
18	Ireland	0.5986	0.6520	15
19	Austria	0.5929	0.6118	20
20	France	0.5902	0.6355	16
21	Sri Lanka	0.5713	0.4826	39
22	Romania	0.5466	0.5516	27
23	Belgium	0.5401	0.6159	19
24	India	0.5394	0.5693	24
25	United Kingdom	0.5352	0.6048	21
26	Russia	0.5220	0.5442	28
27	Lithuania	0.5127	0.5017	34
28	Thailand	0.5093	0.5171	33
29	Poland	0.5089	0.4916	35
30	Puerto Rico	0.5041	0.4884	38
31	Switzerland	0.5032	0.5632	25
32	Hungary	0.5018	0.5172	32
33	United Arab Emirates	0.4970	0.4814	42
34	Malta	0.4929	0.5172	31
35	Cyprus	0.4914	0.4523	48
36	Portugal	0.4870	0.4893	37
37	Saudi Arabia	0.4854	0.5188	30
38	Spain	0.4827	0.4729	43
39	Egypt	0.4826	0.4389	55
40	Costa Rica	0.4705	0.3989	66
41	Liechtenstein	0.4685	0.5771	23
42	Kenya	0.4670	0.3072	91
43	Slovakia	0.4644	0.4514	49
44	Uruguay	0.4643	0.4611	45
45	Italy	0.4615	0.4816	41
46	Czechia	0.4615	0.4907	36
47	Iran	0.4606	0.4544	46
48	Belarus	0.4575	0.5900	22
49	Estonia	0.4571	0.5564	26
50	Australia	0.4510	0.4482	53
51	Morocco	0.4472	0.4157	61
52	Greece	0.4433	0.4484	52
53	Cayman Islands	0.4291	0.4158	60
54	Denmark	0.4278	0.4500	50
55	Chile	0.4270	0.4052	63
56	Brazil	0.4249	0.4250	58
57	Tunisia	0.4223	0.4170	59
58	Philippines	0.4182	0.4818	40
59	Turkey	0.4166	0.3547	80
60	Indonesia	0.4156	0.3743	75
61	Mexico	0.4141	0.3825	72
62	Norway	0.4121	0.4102	62
63	Slovenia	0.4116	0.4027	64
64	Bulgaria	0.4108	0.4530	47
65	Peru	0.4094	0.3702	77
66	Ukraine	0.4053	0.4440	54
67	New Zealand	0.4043	0.3937	70
68	Korea, DPR	0.4040	0.4489	51
69	South Africa	0.4033	0.3689	78
70	Kazakhstan	0.4014	0.3339	86
71	Algeria	0.4012	0.3539	81
72	Latvia	0.3980	0.4656	44
73	Guatemala	0.3852	0.2493	97
74	Bosnia and Herzegovina	0.3848	0.3977	67
75	Trinidad and Tobago	0.3744	0.3939	69
76	Colombia	0.3692	0.3353	85
77	Moldova	0.3658	0.3259	88
78	Croatia	0.3649	0.3674	79
79	Iceland	0.3519	0.4009	65
80	Argentina	0.3509	0.3942	68
81	Mongolia	0.3506	0.2638	95
82	Venezuela	0.3485	0.3730	76
83	Georgia	0.3470	0.4337	57
84	Andorra	0.3368	0.2797	92
85	Cuba	0.3335	0.3499	82
86	Ecuador	0.3330	0.2439	98

Table A4 (*continued*)

GPYC Ranking	Country or Region	GPYC	Patents	Patents Ranking
87	Panama	0.3294	0.3080	89
88	North Macedonia	0.3292	0.3368	84
89	Kuwait	0.3238	0.3452	83
90	Monaco	0.3232	0.3785	73
91	Jordan	0.3027	0.4344	56
92	Uzbekistan	0.2998	0.3279	87
93	Zimbabwe	0.2955	0.2668	94
94	Nigeria	0.2827	0.2568	96
95	Bermuda	0.2630	0.3078	90
96	Armenia	0.2493	0.3781	74
97	El Salvador	0.2236	0.2770	93
98	Seychelles	0.2145	0.2147	99
99	Jamaica	0.1871	0.3844	71
100	Djibouti	0.0000	0.0000	100

References

[1] W.B. Arthur, Complexity and the economy, in: Handbook of Research on Complexity, Edward Elgar Publishing, 2009.
[2] B.T. Asheim, M.S. Gertler, The geography of innovation: regional innovation systems, 2006.
[3] D.B. Audretsch, M. Belitski, Knowledge complexity and firm performance: evidence from the European smes, J. Knowl. Manag. 25 (4) (2021) 693–713.
[4] B. Balassa, Trade liberalisation and "revealed" comparative advantage 1, Manch. Sch. 33 (2) (1965) 99–123.
[5] B. Balassa, The Theory of Economic Integration (Routledge Revivals), Routledge, 2013.
[6] P.-A. Balland, R. Boschma, J. Crespo, D.L. Rigby, Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification, Reg. Stud. 53 (9) (2019) 1252–1268.
[7] P.-A. Balland, D. Rigby, The geography of complex knowledge, Econ. Geogr. 93 (1) (2017) 1–23.
[8] A.D. Chandler Jr, Regional Advantage: Culture and Competition in Silicon Valley and Route 128, 1995.
[9] C.-L. Chang, S.-P. Chen, M. McAleer, Globalization and knowledge spillover: international direct investment, exports and patents, Econ. Innov. New Technol. 22 (4) (2013) 329–352.
[10] C.A. Corrado, C.R. Hulten, How do you measure a "technological revolution"?, Am. Econ. Rev. 100 (2) (2010) 99–104.
[11] P. Dicken, Global Shift: Mapping the Changing Contours of the World Economy, SAGE Publications Ltd., 2007.
[12] J. Duch, A. Arenas, Community detection in complex networks using extremal optimization, Phys. Rev. E 72 (2) (2005) 027104.
[13] J.H. Dunning, Regions, Globalization, and the Knowledge-Based Economy, OUP, Oxford, 2000.
[14] O. Ejermo, Regional innovation measured by patent data—does quality matter?, Research paper, Ind. Innov. 16 (2) (2009) 141–165.
[15] L. Fleming, O. Sorenson, Technology as a complex adaptive system: evidence from patent data, Res. Policy 30 (7) (2001) 1019–1039.
[16] K. Frenken, R.A. Boschma, A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process, J. Econ. Geogr. 7 (5) (2007) 635–649.
[17] O. Goldreich, E. Petrank, Quantifying knowledge complexity, Comput. Complex. 8 (1999) 50–98.
[18] G.H. Golub, C.F. Van Loan, Matrix Computations, JHU Press, 2013.
[19] B.H. Hall, A.B. Jaffe, M. Trajtenberg, The nber patent citation data file: Lessons, insights and methodological tools, 2001.
[20] U.-E. Haner, Innovation quality—a conceptual framework, Int. J. Prod. Econ. 80 (1) (2002) 31–37.
[21] R. Hausmann, C.A. Hidalgo, S. Bustos, M. Coscia, A. Simoes, The Atlas of Economic Complexity: Mapping Paths to Prosperity, Mit Press, 2014.
[22] D. Helbing, A. Kirman, Rethinking economics using complexity theory, Real-world Econ. Rev. 64 (2013).
[23] C.A. Hidalgo, R. Hausmann, The building blocks of economic complexity, Proc. Natl. Acad. Sci. 106 (26) (2009) 10570–10575.
[24] A.B. Jaffe, M. Trajtenberg, Patents, Citations, and Innovations: A Window on the Knowledge Economy, MIT Press, 2002.
[25] E. Kemp-Benedict, An interpretation and critique of the method of reflections, 2014.
[26] D.F. Kogler, D.L. Rigby, I. Tucker, Mapping knowledge space and technological relatedness in us cities, in: Global and Regional Dynamics in Knowledge Flows and Innovation, Routledge, 2015, pp. 58–75.
[27] C. Lawson, E. Lorenz, Collective learning, tacit knowledge and regional innovative capacity, Reg. Stud. 33 (4) (1999) 305–317.
[28] M.S. Mariani, A. Vidmer, M. Medo, Y.-C. Zhang, Measuring economic complexity of countries and products: which metric to use?, Eur. Phys. J. B 88 (2015) 1–9.
[29] P. Mealy, J.D. Farmer, A. Teytelboym, Interpreting economic complexity, Sci. Adv. 5 (1) (2019) eaau1705.
[30] G. Morrison, S.V. Buldyrev, M. Imbruno, O.A. Doria Arrieta, A. Rungi, M. Riccaboni, F. Pammolli, On economic complexity and the fitness of nations, Sci. Rep. 7 (1) (2017) 15332.
[31] B. Ó Huallacháin, D.-S. Lee, Technological specialization and variety in urban invention, Reg. Stud. 45 (1) (2011) 67–88.
[32] S. Ozsoy, B. Fazlioglu, S. Esen, do fdi and patents drive sophistication of exports? A panel data approach, 2021.
[33] S. Pekovic, F. Galia, From quality to innovation: evidence from two French employer surveys, Technovation 29 (12) (2009) 829–842.
[34] M. Pintea, P. Thompson, Technological complexity and economic growth, Rev. Econ. Dyn. 10 (2) (2007) 276–293.
[35] E. Pugliese, A. Zaccaria, L. Pietronero, On the convergence of the fitness-complexity algorithm, Eur. Phys. J. Spec. Top. 225 (2016) 1893–1911.
[36] D.L. Rigby, C. Roesler, D. Kogler, R. Boschma, P.-A. Balland, et al., Do eu regions benefit from smart specialization, Utrecht University Papers in Evolutionary Economic Geography, 2019, p. 19.
[37] J.W. Rivkin, Imitation of complex strategies, Manag. Sci. 46 (6) (2000) 824–844.
[38] U. Schmoch, Concept of a technology classification for country comparisons, in: Final Report to the World Intellectual Property Organisation (Wipo), WIPO, 2008, report to the world intellectual property organisation (wipo), WIPO.
[39] C. Sciarra, G. Chiarotti, F. Laio, L. Ridolfi, A change of perspective in network centrality, Sci. Rep. 8 (1) (2018) 15269.
[40] C. Sciarra, G. Chiarotti, L. Ridolfi, F. Laio, Reconciling contrasting views on economic complexity, Nat. Commun. 11 (1) (2020) 3352.
[41] D. Skillicorn, Understanding Complex Datasets: Data Mining with Matrix Decompositions, CRC Press, 2007.
[42] O. Sorenson, L. Fleming, Science and the diffusion of knowledge, Res. Policy 33 (10) (2004) 1615–1634.
[43] O. Sorenson, J.W. Rivkin, L. Fleming, Complexity, networks and knowledge flows, in: The Handbook of Evolutionary Economic Geography, Edward Elgar Publishing, 2010.
[44] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, L. Pietronero, A new metrics for countries' fitness and products' complexity, Sci. Rep. 2 (1) (2012) 723.
[45] S. Tamer Cavusgil, R.J. Calantone, Y. Zhao, Tacit knowledge transfer and firm innovation capability, J. Bus. Ind. Mark. 18 (1) (2003) 6–21.

[46] M. Trajtenberg, A penny for your quotes: patent citations and the value of innovations, Rand J. Econ. (1990) 172–187.
[47] B. Verspagen, Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research, Adv. Complex Syst. 10 (01) (2007) 93–115.
[48] S. Zhang, P. Zhang, A new approach on measuring the knowledge complexity in the view of the bipartite network, Appl. Econ. Lett. 31 (2) (2024) 146–151.