Applied Economics Letters

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rael20

A new approach on measuring the knowledge complexity in the view of the bipartite network

Shenshen Zhang & Paul Zhang

To cite this article: Shenshen Zhang & Paul Zhang (2024) A new approach on measuring the knowledge complexity in the view of the bipartite network, Applied Economics Letters, 31:2, 146-151, DOI: 10.1080/13504851.2022.2128290

To link to this article: https://doi.org/10.1080/13504851.2022.2128290

Published online: 26 Sep 2022.

Submit your article to this journal

Article views: 161

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rael20

APPLIED ECONOMICS LETTERS
2024, VOL. 31, NO. 2, 146–151
https://doi.org/10.1080/13504851.2022.2128290

ARTICLE

A new approach on measuring the knowledge complexity in the view of the bipartite network

Shenshen Zhang and Paul Zhang

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

ABSTRACT

Developed countries are guaranteed to have the ownership of higher knowledge complexity, which is naturally accepted and universally acknowledged. However, the reality may tell a different story. In this article, a brand-new approach is utilized to quantify the knowledge complexity. Within the bipartite network model, based on the Fitness and Complexity algorithm and the matrix-estimation exercise, a couple of indicators are constructed to measure generalized knowledge complexities of countries and technologies. Results illuminate that admittedly knowledge complexities of countries and those of technologies are interrelated, and those established and developed countries, compared with developing ones, do not necessarily own higher knowledge complexity. To be more specific, an increasing number of facts demonstrate that some less-developed countries with less-advanced technologies are progressing in leaps and bounds, for they attach great importance to investing in high-knowledge-complexity technologies.

KEYWORDS

Knowledge complexity; bipartite network; The Fitness and Complexity algorithm; matrix-estimation exercise; OLS

JEL CLASSIFICATION

C45; C60; P49

I. Introduction and literature

Complex knowledge is of high economic value (Pugliese, Zaccaria, and Pietronero 2016), for its dissemination, among economic agents (Sciarra et al. 2020), and is usually hindered by the considerable cost of its acquisition (Olav, Rivkin, and Fleming 2002). To stand out in globalization, countries have to accumulate tacit, complex, high-value and non-ubiquitous knowledge (Asheim and Gertler 2006).

Previous studies emphasize on the summation of knowledge input and output. Therefore, more than often, their ignorance of characteristics pertaining to knowledge causes the invalidity (Balland and Rigby 2017). However, the method proposed by Balland and Rigby (2017) has strong feasibility and applicability (Broekel 2019). Within the bipartite network model, they adopt the 'Method of Reflection (MR)' advanced by Hidalgo and Hausmann (2009). Admittedly, MR can be defective as it lacks attention in terms of the country's diversification and product's ubiquity (Sciarra et al. 2018). Tacchella et al. (2012) put forward the Fitness and Complexity (FC) algorithm, retaining most of the information computed by MR (Morrison et al. 2017) while avoiding defects mentioned above. Frankly speaking, the FC approach calculates some kind of 'eigenvector centrality', which typically descends from ad hoc assumptions (Rivkin 2000). In order to avoid that, Sciarra et al. (2018, 2020) harness the matrix-estimation exercise to construct two new indicators, on the basis of the FC algorithm, giving birth to a new idea on measuring 'complexity'.

This article is a step forward on Balland and Rigby's method. Two indicators are constructed: GPYC and GPYT, both of which are based on Sciarra's method. Knowledge complexity of countries and technologies is quantified via calculating GPYCs and GPYTs.

II. Method

A general framework

Within the bipartite network, economic information can be collected by defining exporters and products as two sets of vertexes, with the trade data of them as edges. In order to quantify the complexity, we define two properties X_c and Y_p in describing complexity of countries and technologies, respectively. Thereinto, X_c can be determined

CONTACT Shenshen Zhang mgc1232@163.com Shenshen Zhang School of Economics and Finance, Shanghai International Studies University No.550, Dalian Rd. (W), Hongkou District, Shanghai, 200080, China

© 2022 Informa UK Limited, trading as Taylor & Francis Group

by $Y_1, Y_2, \ldots, Y_t$ based on M_{ct}, and Y_t just the same. That is, the problem of measuring complexity is equal to acquiring the solutions of the following coupling linear equations:

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

Equation 1 can be demonstrated into an eigenproblem, by means of using a transformational matrix W:

$$\begin{cases} X_c = \frac{1}{\sqrt{\lambda}} \sum\limits_t W_{ct} Y_t \\ Y_t = \frac{1}{\sqrt{\lambda}} \sum\limits_c W_{ct} X_c \end{cases} \Leftrightarrow \begin{cases} X_c = \frac{1}{\lambda} \sum\limits_t \sum\limits_{c'} W_{ct} W_{c't} X_{c'} = \frac{1}{\lambda} \sum\limits_{i'} N_{cc'} X_{c'} \\ Y_t = \frac{1}{\lambda} \sum\limits_c \sum\limits_{t'} W_{ct} W_{ct'} Y_{t'} = \frac{1}{\lambda} \sum\limits_{t'} G_{tt'} Y_{t'} \end{cases} \tag{2}$$

where W_{ct} can be calculated via M_{ct}; λ is the eigenvalue; and $N_{cc'} = \left(\boldsymbol{W}\boldsymbol{W}^{\boldsymbol{T}}\right)_{cc'}$ and $G_{tt'} = \left(\boldsymbol{G}^{\boldsymbol{T}}\boldsymbol{G}\right)_{tt'}$ are both symmetric square matrixes. N/G can be interpreted as the proximity matrixes, describing the similarity of any two countries/technologies in terms of their 'participating constructure'.

The Fitness and Complexity algorithm

As to countries, covering a wider range of technology domains will spur the knowledge complexity to a higher position; in addition to that, developing high-end sophisticated technologies in a more involved way will also result in higher knowledge complexities.

As for technologies, supported by countries with a higher knowledge complexity, the complexity of technologies is higher, but with the increasing number of participant countries, the complexity of technologies will be cut down.

Therefore, the FC algorithm is embraced to solve eigenproblem mentioned above, and two complexities are defined, namely the Fitness F_c for countries and the Complexity Q_t for technologies, abiding by the following iterations:

$$\begin{cases} \widetilde{F_c}^{(n)} = \sum\limits_t M_{ct} Q_t^{(n-1)} \\ \widetilde{Q_t}^{(n)} = \frac{1}{\sum_c M_{ct} \frac{1}{F_c^{(n-1)}}} \end{cases} \Leftrightarrow \begin{cases} F_c^{(n)} = \frac{\widetilde{F_c}^{(n)}}{\sum_c \widetilde{F_c}^{(n)} / c} \\ Q_t^{(n)} = \frac{\widetilde{Q_t}^{(n)}}{\sum_t \widetilde{Q_t}^{(n)} / t} \end{cases} \tag{3}$$

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

Because of $\widetilde{F_c}^{(n)}/\widetilde{Q_t}^{(n)}$, $F_c^{(n)}/Q_t^{(n)}$ tend to be convergent (Pugliese, Zaccaria, and Pietronero 2016), parameters are introduced as follows: $c_F = \frac{C}{\sum_t Q_t s_t}$ and $c_Q = \frac{\sum_t Q_t s_t}{T}$. Equation 4 can thus be transformed into a non-iterative form: $\begin{cases} F_c = c_F \sum\limits_t M_{ct} Q_t \\ Q_t = c_Q \frac{1}{\sum_c M_{ct} \frac{1}{F_c}} \end{cases}$, where $s_t = \sum\limits_t M_{ct}$; and further into a linear expression, via the Taylor Expansion (Sciarra et al. 2018):

$$\begin{cases} F_c \simeq c_F \sum\limits_t M_{ct} Q_t \\ Q_t \simeq \frac{c_Q}{\left(s_t'\right)^2} \sum\limits_c \frac{M_{ct} F_c}{s_c^2} \end{cases} \tag{4}$$

where $s_c = \sum_c B_{ct}$ and $k_t' = \sum_c M_{ct}/s_c$.

Equation 4 can be mapped in our general framework via using $\begin{cases} X_c = F_c / s_c \\ Y_t = Q_t s_t' \end{cases}$, and the proximity matrixes $\boldsymbol{N}$ and $\boldsymbol{G}$ can be represented as:

$$\begin{cases} N_{cc'} = \frac{\sum_{c'} M_{ct} M_{c't}}{\left(s_t'\right)^2 s_{c'} s_c} \\ G_{tt'} = \frac{\sum_{t'} M_{ct} M_{ct'}}{\left(s_c\right)^2 s_{t'}' s_t'} \end{cases} \tag{5}$$

Generalized complexity indicators

To quantify the generalized knowledge complexity, the method advanced by Sciarra et al. is employed (Sciarra et al. 2018). Indicators are constructed as GPYC (Generalized comPlexitY of Country) and GPYT (Generalized comPlexitY of Technology):

$$\begin{cases} GPYC_c = \left(\sum\limits_{i=1}^{2} \lambda_i^N \left(v_{t,i}^N\right)^2\right)^2 + 2 \sum\limits_{i=1}^{2} \left(\lambda_i^N\right)^2 \left(v_{t,i}^N\right)^2, c = 1, 2, \ldots \\ GPYT_t = \left(\sum\limits_{i=1}^{2} \lambda_i^G \left(v_{t,i}^G\right)^2\right)^2 + 2 \sum\limits_{i=1}^{2} \left(\lambda_i^G\right)^2 \left(v_{t,i}^G\right)^2, c = 1, 2, \ldots \end{cases} \tag{6}$$

λ_1^N/λ_2^N and λ_1^G/λ_2^G are the two largest eigenvalues of the proximity matrix $\boldsymbol{N}$ and $\boldsymbol{G}$. It is worth noting that the proximity matrix $\boldsymbol{N}$ and $\boldsymbol{G}$ here are modified ones:

$$\begin{cases} N_{cc'} = \dfrac{\sum_{c'} M_{ct} M_{c't}}{\left(s_t'\right)^2 s_{c'} s_c}, & c' \neq c \\ N_{cc'} = 0, c' = c \end{cases}$$

and

$$\begin{cases} G_{tt'} = \dfrac{\sum_{t'} M_{ct} M_{ct'}}{\left(s_c\right)^2 s_{t'}' s_t'}, & t' \neq t \\ G_{tt'} = 0, t' = t \end{cases}$$

, which omit the redundant information of the self-proximity of vertexes.

Based on the eigenvectors of the proximity matrix N and G, GPYC and GPYT are both the derivative indicators of eigenvector centrality, with similar topological implications: the importance of a vertex in network depends not only on its own properties, usually referring to the 'degree', but also on properties of its neighbours. A larger GPYC/GPYT generally stands for the higher complexity at the national/technological level.

III. Data

The data in this article is quoted from the OECD Stat, the 'Patents Statistics' subitem beneath the 'Science, Technology and Patents' item. In this article, the country–technology bipartite network is constructed based on the Revealed Comparative Advantage (RCA):

$$M_{ct} = RCA_{ct,year} = \frac{\sum_{y=1989}^{year} year \mathrm{P}_{ct,y} \Big/ \sum_{y=1989}^{year} year\left(\sum_t \mathrm{P}_{ct,y}\right)}{\sum_{y=1989}^{year} year\left(\sum_c \mathrm{P}_{ct,y}\right) \Big/ \sum_{y=1989}^{year} year\left(\sum_c \sum_t \mathrm{P}_{ct,y}\right)} \tag{7}$$

in which P_{ct} is the patent number of country c in technology t.

To calculate RCA, this article quotes the 'patent number' data of 100 countries in 35 technology domains classified by WIPO, from 1989 to 2018. While calculating RCA_{ct} in a given year, the patent number refers to the total amount from 1989 to that year, namely $\mathrm{P}_{ct,y} = \sum_{i=1989}^{y} \mathrm{N}_{ct,i}$. The general overview on 'patent number' can be seen in Figure 1.

Results

GPYC

Figure 2 reveals that all of the 100 countries' trajectories of GPYC rank change, and some of them are highly developed and established. Results show that Japan, Korea and Singapore have high placings in GPYC, respectively, ranking 4th, 2nd and 3rd in 2018, which is consistent with the common sense. That is because they have comparative edges in the domains of semiconductors, computer technology, electrical machinery, etc., all of which possess elevated knowledge complexity, according to the calculation following in the article.

China and Chinese Taipei rank 5th and 1st in 2018, respectively. As one of 'the Four Asian Dragons', Chinese Taipei has contributed a good deal of patents on three technology domains mentioned above for a long period, and it underlines the importance of building the strengths of high techs. Thanks to the

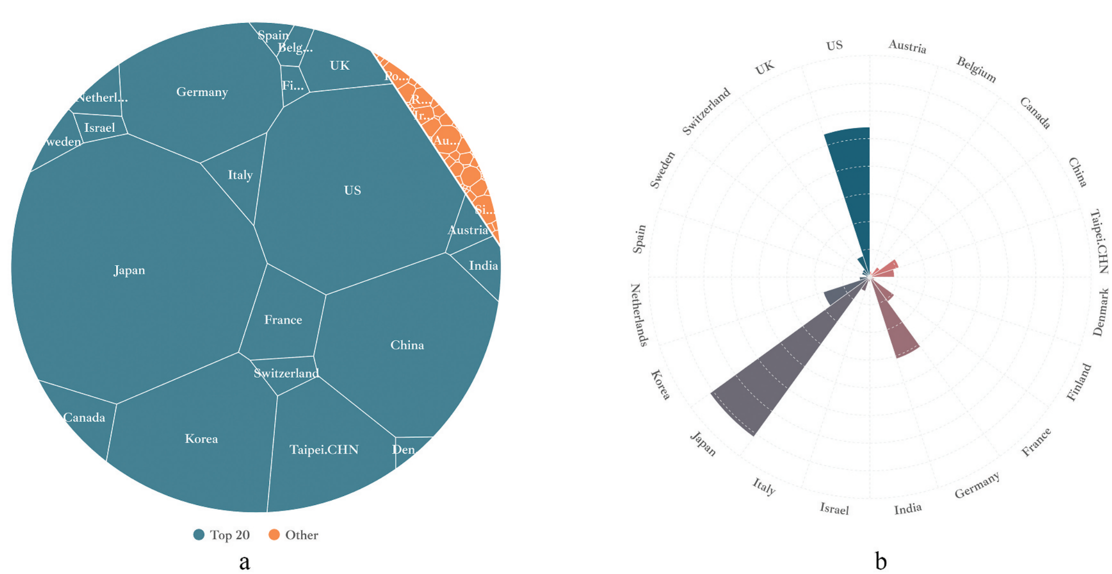

Figure 1. Visualizations on patent volume of countries. (a) Shows the patent proportions of the top 20 and the other countries in 2018. (b) Shows total patent volume from 1989 to 2018 of the top 20 countries.

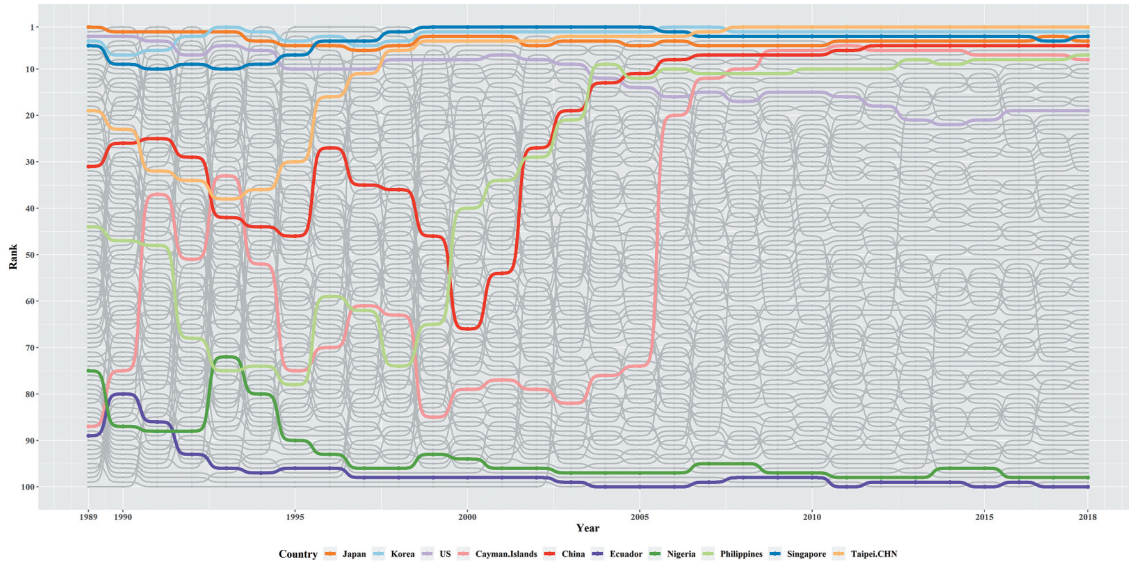

Figure 2. Countries' trajectories of GPYC rank changes from 1989 to 2020.

'Reform and Opening-up' policy, China's technology is booming. Relying on the advantages of the scale effect and the integrated industrial system, China completes technological accumulations and rapid iterations, promoting the ranking from 31st in 1989 to 5th in 2018.

Surprisingly, countries like the Philippines and the Cayman Islands are also high-ranking. The Philippines has a strong strength in terms of semiconductors and electrical machinery as well. Although its total annual patent volume is much smaller than that in developed countries, the Philippines's patent concerning highly complex technologies makes up a large proportion of 95% each and every year. The Cayman Islands has the same advantage as well for its advances in computer technology and digital communication during the last dozen years.

In stark contrast, the ranking of the US plummets to 20th in 2018 from 3rd in 1989. Since the range of technology domains where the US invest is quite wide, its comparative advantages on particular sophisticated domains are dispersed.

Countries ranking 1st to 10th and 81st to 100th in 2018 are listed with their GPYCs in Table 1. The GPYC ranks of 100 countries, in 1989 and 2018, are shown with their rank changes in Figure 3.

Table 1. The top 20 and the last 20 countries in terms of GPYC.

Country	GPYC	Rank	Country	CPYC	Rank
Taipei, China	0.002356523	1	**Cyprus**	0.002214	81
Korea	0.002313476	2	**Latvia**	0.002214	82
Singapore	0.002270177	3	**Argentina**	0.002214	83
Japan	0.002267346	4	**Uruguay**	0.002213	84
China	0.002264124	5	**New Zealand**	0.002213	85
Malaysia	0.002252745	6	**Jordan**	0.002213	86
Philippines	0.002249338	7	**Costa Rica**	0.002213	87
Cayman Islands	0.002243667	8	**Bermuda**	0.002213	88
Egypt	0.002234916	9	**Chile**	0.002213	89
N.Korea	0.002231748	10	**Tunisia**	0.002213	90
Finland	0.002230652	11	**Zimbabwe**	0.002213	91
Romania	0.002228195	12	**Morocco**	0.002213	92
Belarus	0.002225724	13	**Colombia**	0.002212	93
HK, China	0.002225663	14	**Iceland**	0.002212	94
BA	0.002225314	15	**Andorra**	0.002212	95
Germany	0.002223617	16	**Guatemala**	0.002211	96
Netherlands	0.00222356	17	**Kenya**	0.002211	97
Sweden	0.002223538	18	**Nigeria**	0.00221	98
Malta	0.002223485	19	**Cuba**	0.002209	99
US	0.002223252	20	**Ecuador**	0.002209	100

GPYT

Table 2 shows the technology domains ranking 1st to 5th and 31st to 35th in 2018. The patent volume and GPYT of each domain can be seen in Figure 4. Not surprisingly, the top 5 technology domains are all high-end and familiar to everyone. Technologies like semiconductors, audiovisual technology, etc., are monopolized by a mere number of countries. These technology domains are naturally led by countries and territories with higher generalized knowledge complexity, say Chinese Taipei, Japan, etc.

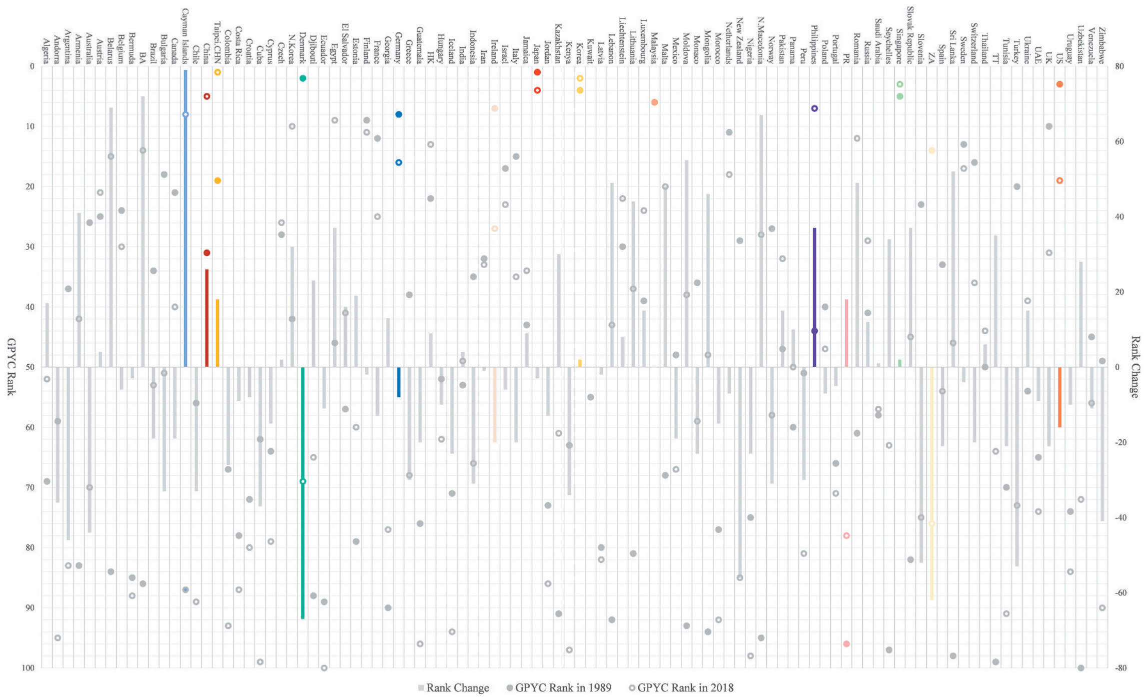

Figure 3. GPYC rank in 1989/2018 and rank change of each country.

Table 2. The top 5 and the last 5 technologies in terms of GPYT.

Code	Technology domain	Rank	Code	Technology domain	Rank
WIPO 8	**Semiconductors**	1	WIPO 11	**Analysis of biological materials**	31
WIPO 2	**Audiovisual technology**	2	WIPO 14	**Organic fine chemistry**	32
WIPO 5	**Basic communication processes**	3	WIPO 15	**Biotechnology**	33
WIPO 9	**Optics**	4	WIPO 16	**Pharmaceuticals**	34
WIPO 3	**Telecommunications**	5	WIPO 18	**Food chemistry**	35

The last-ranked domains are mostly related to biological or pharmaceutical industry, which implies that such industries are involved by a colossal number of countries throughout the globe. As a matter of fact, there still exists robust demand for the bio- or pharma-industry worldwide. Some countries gradually decide to make these industries offshore in the past decade, as more and more countries gradually acquired necessary technologies, thanks to transference of industrial chain.

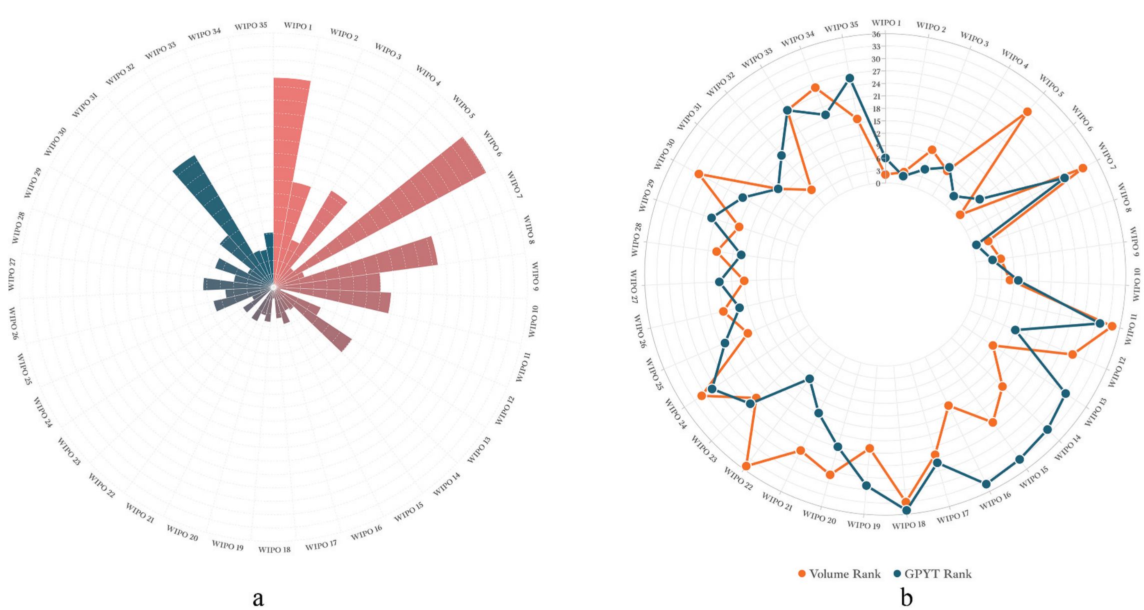

Figure 4. Visualizations on patent volume and GPYT rank of each technology. (a) Shows the volume of 35 technologies in 2018. (b) Shows the ranks of GPYT and total patent volume from 1989 to 2018 of 35 technologies.

IV. Conclusion

This article aims to figure out generalized knowledge complexities at both national and technological levels, through constructing two indicators based on the bipartite network. The indicators, GPYC and GPYT, tend to measure the 'density' or 'intensity' of knowledge within a country or a technology domain instead of the 'amount'. The former may make more sense than the latter in that real GDP per capita is more important rather than GDP.

Calculations in this article indicate that the total volume of patents is not correlated with knowledge complexities directly, neither with countries nor with technologies; knowledge complexities of countries and those of technologies are truly interrelated.

Last but not least, imperfections do exist in this article. For instance, first, setting 1989 as the base time means that the information before is neglected. Second, technologies will be outdated as time passing by, and it is thus not quite appropriate to merely do summations from the base time. Third, no measurement on complexities of technology alone may also lose sight of some information.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Shenshen Zhang http://orcid.org/0000-0003-3271-3796
Paul Zhang http://orcid.org/0000-0002-7420-655X

References

Asheim, B. T., and M. S. Gertler. 2006. "The Geography of Innovation: Regional Innovation Systems." Oxford. *Oxford University Press*. doi:10.1093/OXFORDHB/9780199286805.003.0011.

Balland, P.-A., and D. L. Rigby. 2017. "The Geography of Complex Knowledge." *Economic Geography* 93: 1–23. doi:10.1080/00130095.2016.1205947.

Broekel, T. 2019. "Using Structural Diversity to Measure the Complexity of Technologies." *Plos One*, n. pag 14: e0216856. doi:10.1371/journal.pone.0216856.

Hidalgo, C. A. and R. Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences* 106: 10570–10575. doi: 10.1073/pnas.0900943106.

Morrison, G., S. V. Buldyrev, M. Imbruno, O. A. Doria Arrieta, A. Rungi, M. Riccaboni, and F. Pammolli. 2017. "On Economic Complexity and the Fitness of Nations." *Scientific Reports* 7. n.pag. doi:10.1038/s41598-017-14603-6.

Olav, S., J. W. Rivkin and L. Fleming. 2002. "Complexity, Networks and Knowledge Flow." *Harvard Business School: Strategy Unit Working Paper Series*: n. pag. doi: 10.2139/ssrn.310001.

Pugliese, E., A. Zaccaria, and L. Pietronero. 2016. "On the Convergence of the Fitness-Complexity Algorithm." *The European Physical Journal Special Topics* 225: 1893–1911. doi:10.1140/epjst/e2015-50118-1.

Rivkin, J. W. 2000. "Imitation of Complex Strategies." *Management Science* 4: 824–844. doi:10.1287/MNSC.46.6.824.11940.

Sciarra, C., G. Chiarotti, F. Laio, and L. Ridolfi. 2018. "A Change of Perspective in Network Centrality." *Scientific Reports* 8. n. pag. doi:10.1038/s41598-018-33336-8.

Sciarra, C., G. Chiarotti, L. Ridolfi, and F. Laio. 2020. "Reconciling Contrasting Views on Economic Complexity." *Nature Communications* 11. n. pag. doi:10.1038/s41467-020-16992-1.

Tacchella, A., M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero. 2012. "A New Metrics for Countries' Fitness and Products' Complexity." *Scientific Reports* 2. n. pag. doi:10.1038/srep00723.