Bibliographically Coupled Patents: Their Temporal Pattern and Combined Relevance

Chung-Huei Kuan[1,3], Dar-Zen Chen[2,3]

[1] *maxkuan@mail.ntust.edu.tw*
National Taiwan University of Science and Technology, Graduate Institute of Patent, No. 43, Sec. 4, Keelung Rd., Taipei, Taiwan (R.O.C.)

[2] *dzchen@ntu.edu.tw*
National Taiwan University, Department and Graduate Institute of Mechanical Engineering and Graduate Institute of Industrial Engineering, No. 1, Sec. 4, Roosevelt Rd., Taipei, Taiwan (R.O.C)

[3] National Taiwan University, Center for Research in Econometric Theory and Applications (CRETA), No. 1, Sec. 4, Roosevelt Rd., Taipei, Taiwan (R.O.C.)

Abstract
Bibliographic coupling (BC) is one of the most common indicators in detecting and measuring patent relatedness. Patents that are bibliographically coupled reveal a temporal pattern involving their ages (how long ago they are issued) and spans (their distances in time). BC is more frequently found between patents issued more recently and closer in time, and their coupling strength also tends to be stronger. Aged or long-spanned patent pairs are not only fewer but also inherently limited in their coupling strength. These patent pairs therefore may be overlooked when a threshold is applied, even though their patents are highly related. An improved measure, referred to as *combined relevance*, is proposed to provide fairer treatment to these aged or long-spanned patent pairs when assessing the relatedness between their patents. Combined relevance is as simple as the conventional measures, both conceptually and computationally. More importantly, a fixed threshold may be safely applied with a reduced possibility of erroneously removing the aged or long-spanned patent pairs.

Introduction
A lot of patent bibliometric works involve the detection and measurement of relatedness between patents. Based on their relatedness, structures may be derived from thousands of seemingly disorganized patents. Then, collections of related patents may be modelled and monitored as a unit; evolving trends may be detected by observing related patents or patent clusters in their chronological orders. The cooperation/competition relationship and knowledge exchange between firms, institutions, counties, and fields may be examined and inferred based on the relatedness between their patents.

There are three major categories of approaches in detecting and measuring patent relatedness. The text-based approaches extract textual information from patents' specifications (cf. Arts, Cassiman, & Gomez, 2018; Moehrle & Gerken, 2012; Niemann, Moehrle, & Frischkorn, 2017; Ortiz-de-Urbina-Criado, Nájera-Sánchez, & Mora-Valentín, 2018; Yoon & Magee, 2018). The classification-based approaches evaluate the overlapping of classification symbols assigned to patents (cf. Angue, Ayerbe, & Mitkova, 2014; Chang, 2012; Jaff, 1986; Kuan et al., 2018; Petruzzelli, 2011; Wang, Hou, & Hung, 2018). The third category, citation-based approaches, generally involve three common indicators: direct citation (DC) (Trajtenberg, 1990), bibliographic coupling (BC) (Kessler, 1963), and co-citation (CC) (Small, 1973). These mechanisms may be used individually or together. Kuusi and Meyer (2007) employed BC alone to cluster some related patents and identify an emerging technological paradigm. Lo (2007) also employed only BC to identify technological connections between major research organizations. Von Wartburg, Teichert, and Rost (2005) combined DC and BC in a multistage analysis to reveal technological change. Chen, Huang, Hsieh, and Lin (2011), Yeh, Sung, Yang, Tsai, and

Chen (2013), and Kuan, Huang, and Chen (2018) used both DC and BC to construct more comprehensive citation networks. Citation-based approaches may also be applied together with the other two approaches. Leydesdorff, Kushnir, and Rafols (2014) and Kuan et al. (2018) integrated DC with patent classification codes, Nakamura, Suzuki, Sakata, and Kajikawa (2015) combined DC and co-word analysis, and Park, Jeong, Yoon, and Mortara (2015) used BC and patent text semantic analysis to locate potential R&D collaboration partners.

DC, BC, and CC are all valid relatedness indicators, as evidenced by a large body of works where the aforementioned are only a few samples. These indicators, however, may conflict when one suggests relatedness whereas another specifies otherwise. In a previous study investigating one such conflict involving patents that do not cite each other but are strongly bibliographically coupled, the authors found that this phenomenon is not coincidental (Kuan et al., 2018). For DC to occur, the cited patent has to be published earlier so that it is visible and citable to the applicant or examiner of the citing patent. Therefore, DC rarely occurs between patents whose application processes are highly overlapped, as their applicants or examiners are blind to each other, even though the two patents are indeed related. BC is not handicapped as such and may effectively reveal the patents' relatedness.

During the previous study, the authors also noticed that bibliographically coupled patents reveal a temporal pattern involving their ages (how long ago they have been issued) and spans (their distances in time) and this pattern would place an inherent bound on their coupling strength. In the following sections, this pattern, its cause, and its implication on the traditional measurement and relatedness assessment are described and discussed. Then, an improved measure is proposed to reduce the impact of ages and spans in assessing bibliographically coupled patents.

Temporal Pattern and Cause

A. Phenomenon

To demonstrate this temporal pattern, the same dataset from the previous study is borrowed here, which includes 34,083 US utility patents issued between 1976/1/1 and 2017/3/31 in the field of carbon dioxide capture, storage, recovery, delivery, and regeneration. Among the 34,083 patents, there are 154,505 pairs of cited and citing patents, 1,609,549 pairs of bibliographically coupled patents, and 644,376 pairs of co-cited patents, hereinafter respectively referred to as DC pairs, BC pairs, and CC pairs. Each of the DC, BC, and CC pairs includes patents P_E and P_L respectively issued at an earlier date t_E and a later date t_L ($t_E \leq t_L$). Then, the *span* and *age* of a DC, BC, or CC pair are respectively defined as $t_L - t_E$ and $t_{NOW} - t_E$, where t_{NOW} denotes the cut-off date of the patent data collection (e.g., 2017/03/31 for this dataset).

Figure 1 shows the frequency distributions of all DC, BC, and CC pairs according to their ages (horizontal axis) and spans (vertical axis) in years, where more reddish or bluish points reflect higher or lower counts. BC pairs, in contrast to DC and CC pairs, are particularly concentrated in the lower left corner, meaning significantly more BC pairs have shorter spans and smaller ages, or BC is more frequently found between patents issued more recently and closer in time. DC and CC do not reveal such propensity.

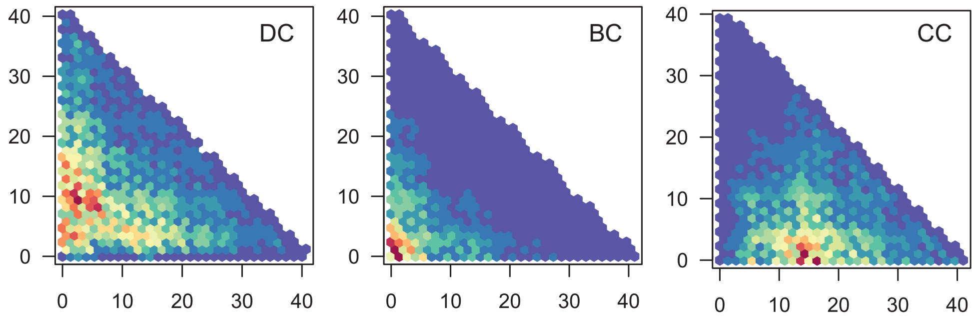

Figure 1. Frequency distributions of DC (left), BC (middle), and CC (right) pairs according to ages (x) and spans (y).

Not only that, the bibliographic coupling strength (BCS) of the BC pairs also reveals a similar pattern. Figure 2 shows the average BCS and average co-citation strength (CCS) for BC and CC pairs across ages and spans, where BCS and CCS are measured as the number of cited and citing patents in common. Again, BC pairs having higher BCS (i.e., more reddish points) are particularly concentrated in the lower left corner, whereas CCS does not show any significant pattern.

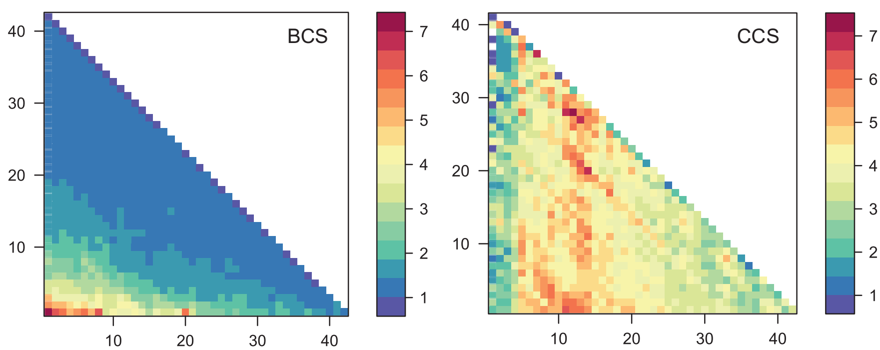

Figure 2. Average BCS (left) and average CCS (right) according to ages (x) and spans (y).

Swanson (1971) and Jarneving (2007a) indicated that only BC pairs having BCS above a threshold are truly related. Indeed, among the 1,609,549 BC pairs, up to 1,167,794 (72.55%) of them have the smallest BCS of 1. To avoid that the pattern shown in Figure 2 is resulted from a large volume of noises, Figure 3 shows the distributions of average BCS after removing those BC pairs having BCS not greater than 1 and the overall average BCS (2.74 or μ). The same pattern is still preserved.

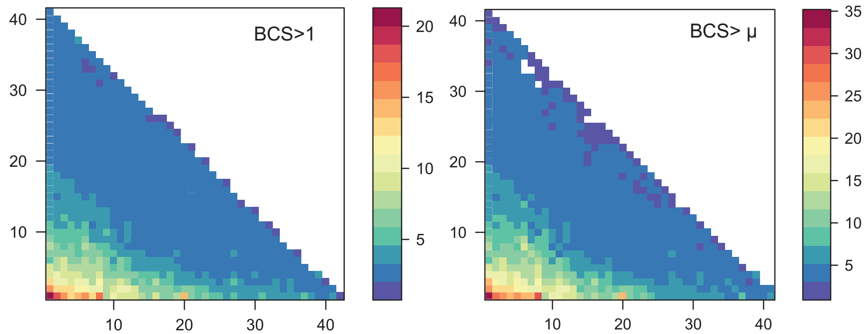

Figure 3. Average BCS for pairs with BCS > 1 (left) and BCS > μ (right) according to ages (x) and spans (y).

B. Patent and Reference Expansion

This unique pattern is related to a field's continuously increasing numbers of accumulated patents and cited references, referred to as *citable patent expansion* and *cited reference expansion* subsequently. Figure 4 depicts the two expansions using the same case data. The horizontal axes show the ages of patents in 2-year intervals. For patents issued earlier in the past or more recently, their data are plotted more to the right or to the left. In the left diagram, the bars show the numbers of patents issued within respective intervals relative to the left scale; the curve depicts the numbers of accumulative patents or citable patents up to each interval. The curve in the right diagram shows the average numbers of cited references at respective intervals.

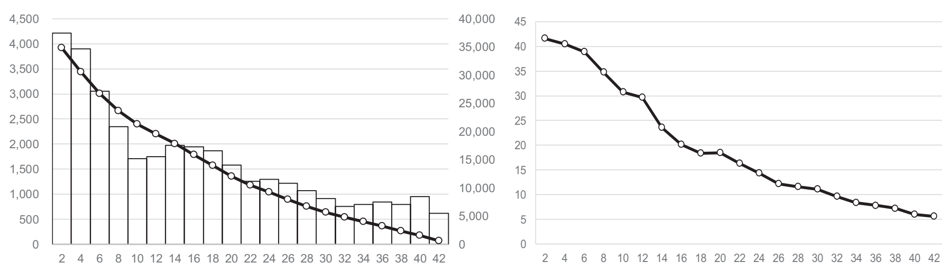

Figure 4. A field's expanding numbers of citable patents (left) and cited references (right).

As illustrated in the left diagram, the citable patent curve rises monotonically from right to left as patents of the field are issued and accumulated over time. Then, later patents have more citable patents and, therefore, may have more references than earlier patents do. The average cited reference curve of the right diagram, as such, also rises monotonically from right to left.

The citable patent expansion should be applicable to patents from any field and from any patent office. The universal applicability of cited reference expansion is, however, questionable. Researchers did notice that the number of references made per U.S. patent and the number of U.S. patents issued both increase over time (Hall, Jaffe, & Trajtenberg, 2001; Zhang, Huang, & Chen, 2018). However, the authors speculates that cited reference expansion should hold for U.S. patents, as U.S. regulation obligates applicants to disclose (i.e., cite) all information (e.g., prior patents) known to be relevant to the applications' patentability (Bicknell, 2008). U.S. applicants therefore tend to cite more when there are more citable patents.

Facing the citable patent and cited reference expansions, Hall, Jaffe, and Trajtenberg (2001) and Zhang, Huang, and Chen (2018) were concerned about the "devaluation" of citations (i.e., a patent's "later citations are less significant than earlier ones"). This study, however, is concerned about their implication on BC and patent relatedness based on BC.

Implication on Conventional Assessment

A. Inherent Bound

Figure 5 depicts four scenarios between patents P_E and P_L from a same field, where the bars are the numbers of patents issued within respective intervals from the left diagram of Figure 4. P_E and P_L have references REF_E and REF_L respectively drawn from citable patents CP_{t_E} and CP_{t_L} accumulated up their issue dates t_E and t_L. P_E and P_L should satisfy the following Eqs. (1) and (2):

$REF_E \subseteq CP_{t_E}$, $|REF_E| \le |CP_{t_E}|$, and (1)

$REF_L \subseteq CP_{t_L}$, $|REF_L| \le |CP_{t_L}|$. (2)

Then, Eq. (3) may be derived according to the citable patent expansion:

$CP_{t_E} \subseteq CP_{t_L}$, $|CP_{t_E}| \le |CP_{t_L}|$. (3)

CP_{t_E} is denoted by dark grey bars, and CPP_{t_L} is denoted by both dark and light grey bars in Figure 5. If the cited reference expansion is applicable, then

$|REF_E| \le |REF_L|$, and (4)

$|REF_E \cap REF_L| \le \min(|REF_E|, |REF_L|) = |REF_E| \le |CP_{t_E}|$. (5)

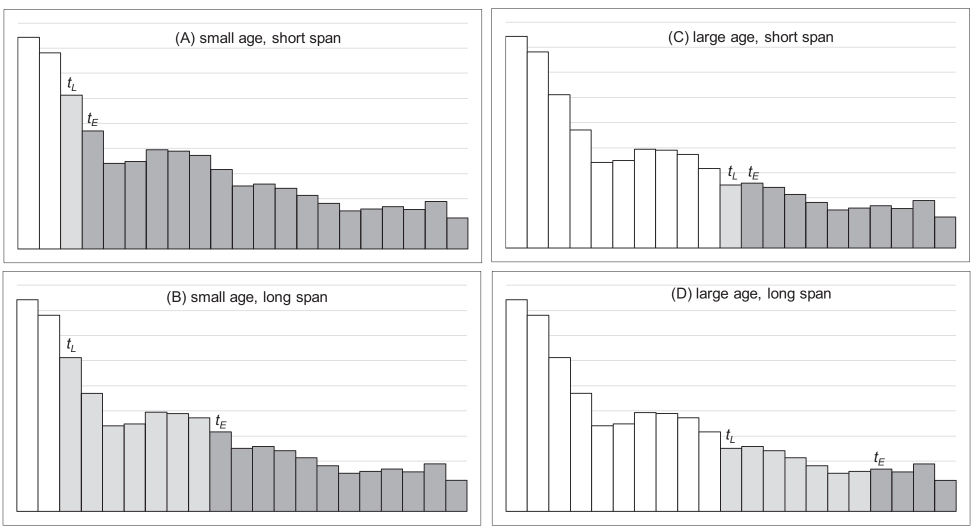

Figure 5. Four scenarios between patents P_E and P_L from a same field.

For P_E and P_L to form a BC pair, they need to have non-empty intersection or $REF_E \cap REF_L \neq \emptyset$. Under cited reference expansion, $|REF_E \cap REF_L|$ cannot be greater than $|REF_E|$, which in turn is bounded $|CP_{t_E}|$. CP_{t_E} is the set of CPs common to both P_E and P_L denoted by the dark grey bars. Therefore, larger CP_{t_E} implies a greater chance in achieving non-empty intersection and thus forming a BC pair. For the four scenarios, whether t_L is more recent as in diagrams (A) and (B) or earlier in the past as in diagrams (C) and (D), $|CP_{t_E}|$ is greater when P_E and P_L have a shorter span. On the other hand, whether P_E and P_L have a shorter span in diagrams (A) and (C) or a longer span as in diagrams (B) and (D), $|CP_{t_E}|$ is greater when t_L is more recent. This is why more BC pairs have shorter spans and smaller ages as illustrated in Figure 1. Similarly, for a BC pair involving P_E and P_L, its BCS, $|REF_E \cap REF_L|$, is also bounded by $|CPP_{t_E}|$, and larger CPP_{t_E} provides a greater chance in achieving higher BCS. This is why BC pairs having shorter spans and smaller ages tend to have greater BCS as illustrated in Figures 2 and 3.

For BC pairs to deliver the unique temporal behavior, the cited reference expansion or Eq. (4) should hold. This is indeed true for a major portion of the case's BC pairs. Table 1 lists the shares of BC pairs satisfying Eq. (4), $|REF_E| \leq |REF_L|$, among those whose BCS crosses ten different thresholds. For all BC pairs (BCS>0), despite a large number of BC pairs that may be noises, there is still more than 65% of the BC pairs satisfying $|REF_E| \leq |REF_L|$. At a higher threshold, there are even greater portions satisfying $|REF_E| \leq |REF_L|$.

Table 1. Shares of BC pairs satisfying $|REF_E| \leq |REF_L|$.

BCS>	*0*	*2*	*4*	*6*	*8*	*10*	*12*	*14*	*16*	*18*
All BC pairs	1,609,549	214,324	94,090	59,260	43,381	34,989	29,357	25,882	23,283	21,284
% of $|REF_E| \leq |REF_L|$	65.46%	66.44%	67.70%	68.46%	69.54%	70.62%	71.70%	72.49%	73.16%	73.59%

B. Problem with Conventional Measure and Threshold

Using a threshold to filter BC pairs, despite a common practice, may not be appropriate for a field revealing the temporal pattern. For example, if a threshold for BCS is set to the overall average (2.74) of the case's all BC pairs, pretty much the BC pairs denoted by the greenish or bluish points in Figure 2 are filtered out, even though some of these aged or long-spanned BC pairs may reflect true relatedness.

This filtering-without-distinction problem not only due to the fixed threshold use, but also due to the BCS measure's overlooking the age and span factors. The frequently used BCS measures may be classified into two broad categories: *intersection-based* and *vector-based* measures with Jaccard coefficient (Jaccard, 1901) and coupling angle (Glänzel, & Czerwon, 1996) as representatives. If Jaccard coefficient or coupling angle is applied to a field with the temporal pattern, all BCS would be bounded by $|REF_E|$, the size of the earlier patent P_E, as derived from the following Eqs (6) and (7):

$$\frac{|REF_E \cap REF_L|}{|REF_E \cup REF_L|} \leq \frac{|REF_E|}{|REF_E \cup REF_L|} \leq |REF_E| \text{ , and} \tag{6}$$

$$\frac{\overrightarrow{REF_E} \cdot \overrightarrow{REF_L}}{|\overrightarrow{REF_E}||\overrightarrow{REF_L}|} = \frac{|REF_E \cap REF_L|}{|\overrightarrow{REF_E}||\overrightarrow{REF_L}|} \leq \frac{|REF_E|}{|\overrightarrow{REF_E}||\overrightarrow{REF_L}|} \leq |REF_E| \text{ ,} \tag{7}$$

where $\overrightarrow{REF_E}$ and $\overrightarrow{REF_L}$ are REF_E and REF_L expressed in binary vectors of equal dimension.

Bibliometric researchers had long noticed the age and span problem. For example, Jarneving (2007b) indicated that "an increase of the distance in time between bibliographically coupled articles leads to a diminishing pool of shared references as there is a tendency to cite the more current articles" That is why usually an observation window is set up so that bibliographically coupled research articles published closer (i.e., about the same age) within the window (i.e., within limited span) are collected and compared together (cf. Jarneving, 2007b; Glänzel, & Czerwon, 1996).

Combined Relevance

However, to observe the knowledge flow or to develop a representative trajectory among patents across a long period of time, where all relevant BC pairs have to be considered, a BCS measure as much immune to their ages and spans as possible would be desirable.

This study therefore proposes a new and simple BCS measure, referred to as *combined relevance* (CR), in Eq. (8):

$$\left(\frac{|REF_E \cap REF_L|}{|REF_E|}\right)\left(\frac{|REF_E \cap REF_L|}{|REF_L|}\right) = \frac{|REF_E \cap REF_L|^2}{|REF_E||REF_L|}. \quad (8)$$

The idea behind Eq. (8) is straightforward. Imaging that REF_E and REF_L respectively represent the information gathered by a BC pair's earlier and later patents P_E and P_L, and that $REF_E \cap REF_L$ is the piece of information shared between them, the two factors in Eq. (8) measure how much this shared information accounts for P_E and P_L's gathered information, or this shared information's *individual relevance* to P_E and P_L. Then, P_E and P_L are considered highly related if their shared information is relevant to both of them.

Figure 6 shows the averages of the two factors for BC pairs whose BCS is greater than 1 in two separate diagrams. The BC pairs are limited to those having BCS > 1 so that the observation is not impaired by a large volume of noises. As illustrated in the left diagram, P_E's factor tends to have higher values for those aged or long-spanned BC pairs located farther away from the lower left corner. This is because their P_E occurs earlier in the field and REF_E is more limited (see Figure 4, right diagram). Then, P_E's factor would be close to 1 as $|REF_E \cap REF_L| \sim |REF_E|$. On the other hand, P_L's factor tends to have higher value for those aged but short-spanned BC pairs located closer to the lower right corner. This is because their P_L also occurs earlier in the field, and REF_L is not only more limited but also close to REF_E, therefore P_L's factor would be close to 1 as $|REF_E \cap REF_L| \sim |REF_L|$.

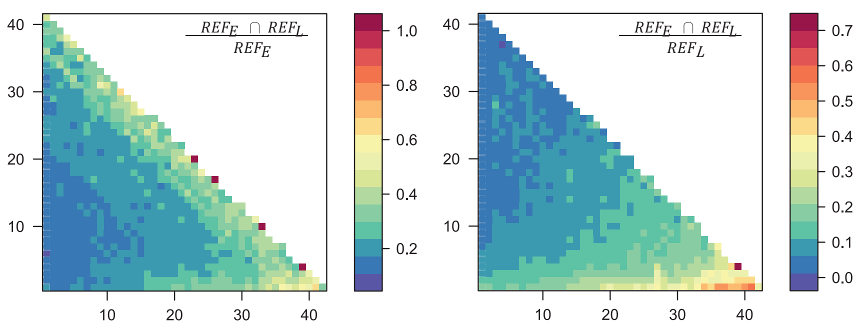

Figure 6. Average relevances of P_E (left) and P_L (right) for pairs with BCS > 1 according to ages (x) and spans (y).

Figure 7 shows the distribution of the average CR for BC pairs whose BCS is greater than 1. As illustrated, CR, as the product of one factor favoring short-spanned pairs and the other favoring long-spanned pairs, has diminished span effect. The age effect remains but to a lesser extent compared to what is revealed in the right diagram of Figure 6. Both phenomena are perhaps due to that P_E's factor does not strictly incline towards aged pairs or long-spanned pairs, but more towards pairs having greater (age+span) values. This is why pairs having higher P_E's factor are more often distributed along the diagonal of the left diagram of Figure 6. This is also why the span effect is not entirely cancelled, and the age effect is lessened. As CR is more uniformly distributed over the ages and spans of BC pairs, a fixed threshold may be applied without causing significant discrimination against aged or long-spanned pairs.

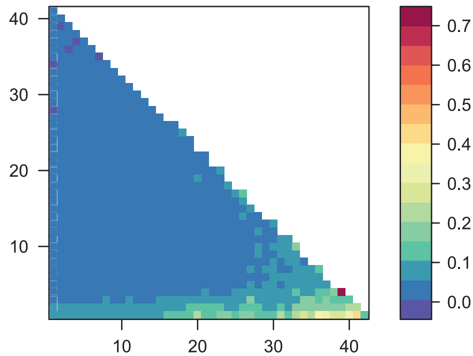

Figure 7. Average CR for pairs with BCS > 1 according to ages (x) and spans (y).

In addition to CR's relatively more uniform distribution across ages and spans, CR also retains more aged and long-spanned pairs. For all BC pairs with BCS>1, their average BCS is 7.32 with standard deviation 29.40, and their average CR is 0.043 with standard deviation 0.14. Figure 8 then provides two diagrams, one showing the frequency distributions of 49,873 BC pairs having above average BCS (left) and the other one showing 55,954 BC pairs having above average CR (right). As illustrated, BCS filters out most pairs having spans above 30, whereas a number of them are still retained by CR. BCS also has fewer larger-aged and longer-spanned pairs. The two sets of BC pairs have 27,369 pairs in common, accounting for 55% of BC pairs in the left diagram, and 49% of BC pairs in the right diagram. In other words, those in the left diagram are not a subset to those in the right diagram. For about half of the BC pairs considered to have reflected relatedness by BCS (or CR), they are indicated otherwise by CR (or BCS).

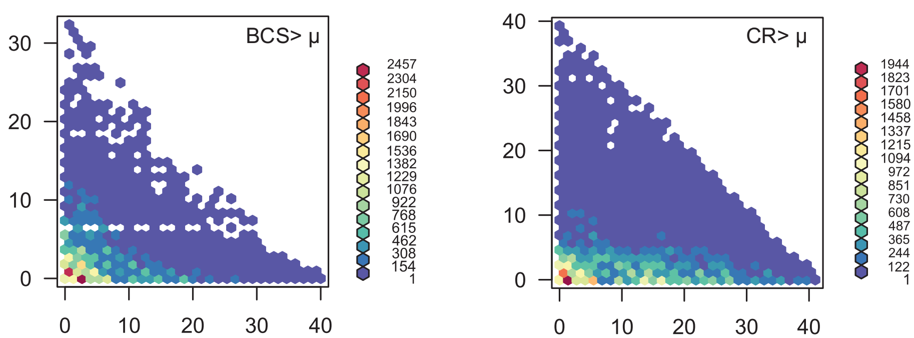

Figure 8. Frequency distributions of BC pairs having above average BCS (left) and CR (right) according to ages (x) and spans (y).

Conclusion

This study describes a temporal pattern found in bibliographically coupled patents, which indicates that BC pairs' temporal characteristics, age and span, may affect their BCS measurement, in addition to the relatedness of their patents. This study therefore proposes a new measure, combined relevance (CR), to reduce such impact.

CR is not ideal as observed above, but it is as simple as the conventional measures, both conceptually and computationally. More importantly, due to its more uniform distribution across various ages and spans, a fixed threshold may be safely applied with a reduced possibility of erroneously removing BC pairs involving truly related patents.

This study does not claim that conventional BCS measures should be replaced by CR, or CR is superior to conventional measures in every respect. When BC pairs are collected from patents issued within a time window, as their ages and spans are confined simultaneously in the same period of time, conventional measures are still viable tools. But for observing long-term knowledge dissemination or tracing overall development trajectory, CR may be a promising new alternative.

This study points out that citable patent expansion and cited reference expansion, the increasing numbers of patents issued and references cited per patent, are the two factors contributing to the temporal pattern. The citable patent expansion is particularly applicable to U.S. patents, as U.S. requires full and obligatory disclosure from patent applicants. However, there is a lack of evidence that non-U.S. patents would undergo cited reference expansion of comparable degree. Therefore, one cannot conclude confidently about non-US patents. Nonetheless, CR's division of relatedness into factors respectively reflecting the relevance of shared references to the patents is a reasonable design, and should still be applicable to non-U.S. patents as well.

This study may be improved in a number of ways. Firstly, a theoretical framework for the temporal pattern should be established, and the temporal pattern should be further verified through rigorous statistical analysis, in addition to the above general observation. Secondly, CR itself may have room for improvement. For one example, the concept of present value may be borrowed and the denominator $|REF_L|$ in $\frac{|REF_E \cap REF_L|}{|REF_L|}$ may be discounted to the date t_E to compensate the cited reference expansion. CR may also be applied to real case data, and the result is compared to that by a conventional measure to see how they differ in an application setting. Finally, the particular threshold picked for removing noise requires further investigation for its validity and influence.

Acknowledgments

This work was financially supported by the Center for Research in Econometric Theory and Applications (Grant no. 108L900204) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan, and by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. MOST 108-3017-F-002-003-.

References

Arts, S., Cassiman, B., & Gomez, J. C. (2018). Text matching to measure patent similarity. *Strategic Management Journal*, *39*(1), 62-84.

Angue, K., Ayerbe, C., & Mitkova, L. (2014). A method using two dimensions of the patent classification for measuring the technological proximity: an application in identifying a potential R&D partner in biotechnology. *The Journal of Technology Transfer*, *39*(5), 716-747.
Bicknell, G. M. (2008). To Disclose or Not to Disclose: Duty of Candor Obligations of the United States and Foreign Patent Offices. *Chi.-Kent L. Rev.*, *83*, 425.
Chang, S. B. (2012). Using patent analysis to establish technological position: Two different strategic approaches. *Technological Forecasting and Social Change*, *79*(1), 3-15.
Chen, D. Z., Huang, M. H., Hsieh, H. C., & Lin, C. P. (2011). Identifying missing relevant patent citation links by using bibliographic coupling in LED illuminating technology. *Journal of Informetrics*, *5*(3), 400-412.
Choi, S. S., Cha, S. H., & Tappert, C. C. (2010). A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, *8*(1), 43-48.
Glänzel, W., & Czerwon, H. J. (1996). A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*, *37*(2), 195-221.
Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.
Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de la Société vaudoise des Sciences Naturelles*, 37, 547-579.
Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value.
Jarneving, B. (2007a). Bibliographic coupling and its application to research-front and other core documents. *Journal of Informetrics*, *1*(4), 287-307.
Jarneving, B. (2007b). Complete graphs and bibliographic coupling: A test of the applicability of bibliographic coupling for the identification of cognitive cores on the field level. *Journal of Informetrics*, *1*(4), 338-356.
Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *Journal of the Association for Information Science and Technology*, *14*(1), 10-25.
Kuan, C. H., Huang, M. H., & Chen, D. Z. (2018). Missing links: Timing characteristics and their implications for capturing contemporaneous technological developments. *Journal of Informetrics*, *12*(1), 259-270.
Kuusi, O., & Meyer, M. (2007). Anticipating technological breakthroughs: Using bibliographic coupling to explore the nanotubes paradigm. *Scientometrics*, *70*(3), 759-777.
Leydesdorff, L., Kushnir, D., & Rafols, I. (2014). Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC). *Scientometrics*, *98*(3), 1583-1599.
Lo, S. C. (2007). Patent coupling analysis of primary organizations in genetic engineering research. *Scientometrics*, *74*(1), 143-151.
Moehrle, M. G., & Gerken, J. M. (2012). Measuring textual patent similarity on the basis of combined concepts: design decisions and their consequences. *Scientometrics*, *91*(3), 805-826.
Nakamura, H., Suzuki, S., Sakata, I., & Kajikawa, Y. (2015). Knowledge combination modeling: The measurement of knowledge similarity between different technological domains. *Technological Forecasting and Social Change*, *94*, 187-201.
Niemann, H., Moehrle, M. G., & Frischkorn, J. (2017). Use of a new patent text-mining and visualization method for identifying patenting patterns over time: Concept, method and test application. *Technological Forecasting and Social Change*, *115*, 210-220.
Ortiz-de-Urbina-Criado, M., Nájera-Sánchez, J. J., & Mora-Valentín, E. M. (2018). A Research Agenda on Open Innovation and Entrepreneurship: A Co-Word Analysis. *Administrative Sciences*, *8*(3), 34.
Park, I., Jeong, Y., Yoon, B., & Mortara, L. (2015). Exploring potential R&D collaboration partners through patent analysis based on bibliographic coupling and latent semantic analysis. *Technology Analysis & Strategic Management*, *27*(7), 759-781.
Petruzzelli, A. (2011). The impact of technological relatedness, prior ties, and geographical distance on university-industry collaborations: A joint-patent analysis. *Technovation*, *31*, 309-319.
Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, *24*(4), 265-269.

Swanson, D. (1971). Some unexplained aspects of the cranfield tests of indexing performance factors. *The Library Quarterly*, *41*(3), 223-228.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, 172-187.

Von Wartburg, I., Teichert, T., & Rost, K. (2005). Inventive progress measured by multi-stage patent citation analysis. *Research Policy*, *34*(10), 1591-1607.

Wang, A. P., Hou, C. E., & Hung, S. W. (2018). Exploration of the evolution of nanotechnology from a patent co-classification perspective. *Nanotechnology Reviews*, *7*(3), 233-245.

Yeh, H. Y., Sung, Y. S., Yang, H. W., Tsai, W. C., & Chen, D. Z. (2013). The bibliographic coupling approach to filter the cited and uncited patent citations: a case of electric vehicle technology. *Scientometrics*, *94*(1), 75-93.

Yoon, B., & Magee, C. L. (2018). Exploring technology opportunities by visualizing patent information based on generative topographic mapping and link prediction. *Technological Forecasting and Social Change*, *132*, 105-117.

Zhang, L., Huang, Y., & Chen, L. (2018). Patent Citation Inflation: The Phenomena, Measurement Methods, and Proposed Solutions Based on Relative Citation Indicators. In *23rd International Conference on Science and Technology Indicators (STI 2018), September 12-14, 2018, Leiden, The Netherlands*. Centre for Science and Technology Studies (CWTS).