

# Navigating the innovation trajectories of technology by combining specialization score analyses for publications and patents: graphene and nano-enabled drug delivery

Seokbeom Kwon<sup>1</sup>  · Alan Porter<sup>1,2</sup> · Jan Youtie<sup>3</sup>

Received: 24 June 2015 / Published online: 5 January 2016  
© Akadémiai Kiadó, Budapest, Hungary 2016

**Abstract** In this study, we combine the specialization scores for publications and patents (the latter is a new indicator of cross-disciplinary engagement) to achieve more comprehensive navigation of the innovation trajectory of a technology. The patent specialization score draws upon counterpart research publication indicator concepts to measure patent diversity. Two nano-based technologies—Nano-enabled drug delivery (NEDD) and Graphene—provide contrasting explorations of the behavior of this indicator, alongside research publication indicators. Results show distinctive patterns of the two technologies and for the respective publication and patent indicators. NEDD research, as evidenced by publication and citation patterns, engages highly diverse research fields. In contrast, NEDD development, as reflected in patent International Patent Classifications (IPCs), concentrates on relatively closely associated fields. Graphene presents the opposite picture, with closely linked disciplines contributing to research, but much more diverse fields of application for its patents. We suggest that analyzing the field diversity of research publications and patents together, employing both specialization scores, can offer fruitful insights into innovation trajectories. Such information can contribute to technology and innovation management and policy for such emerging technologies.

**Keywords** Nano enabled drug delivery · Graphene · Specialization score · Interdisciplinarity · Innovation trajectory · Bibliography analysis

---

✉ Seokbeom Kwon  
skwon61@gatech.edu

<sup>1</sup> School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332-0345, USA

<sup>2</sup> Search Technology, Inc., Atlanta, GA, USA

<sup>3</sup> Enterprise Innovation Institute, Georgia Institute of Technology, Atlanta, GA 30332-0640, USA

## Introduction

Emerging technologies have the potential to reconfigure the industrial and the technological landscape in various application fields (Danneels 2007; Galunic and Rodan 1998; Gruber et al. 2008). Many of these technologies represent combinations of different research fields and patented technologies. Thus, it is important to examine the extent of diversity of the underlying research fields and patented technologies in order to understand the inherent composition of these fields. One aspect of interest is technological diversity in patenting and its counterpart, interdisciplinary research (IDR). The significance of a patent's technological diversity and/or the underlying IDR in fostering innovation is (Stirling 2006) indicated by the number of studies that have, with extensive effort, quantitatively measured the “diversity” and “interdisciplinarity of” or “distance” among scientific disciplines or technology fields engaging a topical domain.

Much research has been involved in analyzing various diversity measures representing published journal articles or patents. Using bibliographic information from published articles, some studies have analyzed the citation or category information in a set of publications to identify the relevant fields of study and examine how diverse fields are associated (Glänzel et al. 1999; Morillo et al. 2003). Other works have employed metrics to measure the diversity of research fields where the underlying disciplines in a set of published journal articles are related. They estimate the diversity with a single measurement (Carley and Porter 2012; Porter and Rafols 2009; Porter et al. 2008; Rao 1982), multiple metrics (Chavarro et al. 2014; Stirling 2006, 2007), and/or an analysis of other dimensional metrics such as coherency or intermediation (Rafols 2014; Rafols et al. 2012).

Measuring the interdisciplinarity of research fields within journal articles provides a perspective into how different bodies of scientific knowledge are interrelated. It also provides way to look at the potential impact of a particular field of study across several different fields. For example, Porter et al. (2008) have examined two core indicators of interdisciplinarity. One indicator, the integration score (comparable to the Rao-Stirling's diversity index (Rao 1982; Stirling 2007), employs subject categories (SCs), which are assigned by the Web of Science (WoS), to calculate the diversity of sources referenced in a given paper (or set of papers). The other indicator, the specialization score, evaluates the diversity of SCs assigned to the journal in which the publication (or set of publications) is printed. SCs that are highly concentrated in a few fields will produce a higher score, with 1 being the maximum value. A comparison of scores across different fields of research reveals how SCs have varying levels of diverse knowledge integration and specialization. Carley and Porter (2012) developed another measurement called the diffusion score (a forward diversity index). This measurement employs the analogous form of the integration score to calculate the forward diversity of a paper or set of papers, but it uses the SCs of journals that contain articles citing the paper or papers in question. First, they retrieve the SCs of the papers that cite the focal article(s). Then, the forward diversity index calculates a diversity index that is analogous to the integration score. Therefore, the indicator measures how many other fields of study are being affected by the body of knowledge in a given paper. The forward diversity index evaluates, from a disciplinary perspective, the level of knowledge diffusion for scientific papers published in 1995 in biotechnology, electrical engineering, mathematics, and medical fields to name a few tested by the authors.

Visualization, has been another approach for understanding the overall landscape of relatedness between scientific disciplines and examining their interdisciplinarity. This

method typically uses cross-citation patterns among journals or categories of journals (Moya-Anegón et al. 2004, 2007; Leydesdorff and Rafols 2009; Rafols et al. 2010).

In a similar vein, other scholars use patent information, particularly patent classes, to measure firm-level or individual patent-level technological diversity and distance. Jaffe (1986, 1989) maps firm-level technological similarity by measuring to what extent the patent classification distribution of a firm's patent portfolio overlaps with other firms' portfolios. Almeida and Kogut (1997) count firms' patents in particular patent classes in order to quantify the diversity of technological fields shared by the firms. Similarly, Breschi et al. (2003) estimate firms' technology diversity by counting technological fields that are represented by particular sets of IPC categories in patent applications. Hinze et al. (1997) estimate the technological distance between technology fields by examining the co-assignment of IPC categories. The more that IPC categories are assigned together, the shorter the technological distance is between their representative technologies. Technological distance, as estimated by IPC co-occurrence patterns, is also employed by the following studies to reveal the innovation dynamics of emerging technology fields. Leydesdorff et al. (2015) utilize IPC categories to calculate the Rao-Stirling diversity index. They also measure IPC category assignments for a set of patents in order to trace the technological change of photovoltaic cells. Kay et al. (2014) employ a transformed IPC (T-IPC), which addresses the unbalanced IPC population problem, in which more patents are assigned to bioscience IPCs than to other, particularly non-high-tech IPCs, and present a visualization of the technological distance between the T-IPCs using cited-citing patterns between T-IPCs in the EPO patents.

Using bibliographic information analysis to represent diversity, interdisciplinarity, and technological distance leads to effective examinations of innovation pathways, innovation trajectories, and technology trends. However, the insights gained from journal article analysis are inherently different from those derived from patent information. Journal articles provide valuable information about behavior, characteristics of scientific knowledge, or relationships between associated research fields. In this sense, interdisciplinarity analysis of academic publications provides insight into how diverse research fields or fields of study advance an underlying scientific discipline. On the other hand, patent analysis provides insight into diverse fields of patenting (in grants or applications). This approach sheds light on the behavior or potential impact of patented technology across various application fields rather than showing the interdisciplinarity of the underlying fields. Thus, analyzing only academic articles or patents produces a partial view of the innovation trajectory: it reveals aspects only of the behavior of scientific knowledge in a research field or of patented technology in an application field. Furthermore, examining the overall innovation trajectory of a technology calls for a combined analysis of scientific knowledge as well as the commercialized technology.

This paper confirms that interdisciplinarity research analysis on journal articles produces different insights than patent diversity analysis when examining the innovation trajectory of a technology. We employ the specialization score (Porter, Roessner, and Heberger 2008). We analyze the specialization scores of two emerging nano-based technologies—NEDD and Graphene—to show how scientific knowledge and the commercialized technology behave differently in terms of interdisciplinarity and diversity. NEDD technology is based on chemistry, biology, and other fields relevant to biotechnology. In contrast, Graphene stems from a single material. For this reason, NEDD research fields, and the related disciplines, are more diversified than those associated with Graphene. On the other hand, NEDDs are limited to biology-relevant applications, such as pharmaceutical or medical uses, while, in contrast, the patented applications of Graphene are

relatively broad. We expect the specialization scores for NEDD patents to be higher than the scores for Graphene patents. This paper also combines the insights from the respective specialization score analysis on journal articles with the analysis of patent information. In our comparative studies of these two domains, we construct a conceptual model based on specialization score analysis, which explains the innovation trajectories of NEDD and Graphene technology, from their research to their patented application.

The paper is structured as follows: In “[Data and Methods](#)” section, we illustrate the data acquisition process for publication and patent records as well as the analysis method. In “[Results](#)” section, we show the specialization score analysis. In “[Discussion](#)” section, we discuss the results and our related insights, and in “[Conclusion](#)” section, we draw our conclusions.

## Data and methods

### Data

We collected NEDD and Graphene patents from the Derwent Innovation Index (DII, <http://apps.webofknowledge.com>), using as our search term strategy a combination of keywords defining Graphene (Shapira, Youtie, and Arora 2012; Kay et al. 2015) and NEDD technology (Zhou et al. 2014). To collect the Graphene patents, we searched for patents that have “Graphene” in their title or abstract. The dataset was collected at the end of March, 2014. For the NEDD patents, we use those patents that were collected by Zhou et al. (2014) through the end of 2012. Likewise, we use the same search methods to gather publication records related to NEDD and Graphene from the Thomson Reuters’ Web of Science, which was collected in January, 2015. We import all the publication and patent records into Vantage Point, a commercialized text mining software package.

### Methods

#### *Publications analysis*

We used the assigned Subject Categories from Web of Science (WoS SCs) to represent research disciplines. For instance, if an article is published in a peer-reviewed journal dealing mainly with Organic Chemistry, the WoS SC for Organic Chemistry is assigned to that paper. Most articles have only one WoS SC, so the majority of the publications would have a specialization score of 1, which does not produce meaningful information (Porter et al. 2008). Therefore, we calculated an aggregated level specialization score based on WoS SC and publication year. Suppose that two journal articles that are published in the same year are associated with WSC1 and WSC2 respectively. For an aggregated level specialization score, we pool the WSC1 and WSC2 together, and compute a specialization score with the pooled WoS lists.

The WoS SCs are not always independent of one another. Some WoS SCs are more strongly correlated with other SCs; others are not. For example, WoS may assign both “Transplantation” and “Surgery” to a journal, but these categories rarely appear in combination with “Telecommunication” so the SC combinations should not necessarily be viewed as equivalent. We take the distance between SCs into account by employing the cosine similarity measure between WoS SCs (Leydesdorff and Rafols 2009). Although the cosine similarity between WoS SCs might not perfectly represent the distance between

SCs, Carley and Porter (2012) explain that the preferred SC is usually very close to the WoS SC and it is quite robust when analyzed across a large number of articles. We apply the specialization score to large number of journal articles (more than 100), which enables us to use the cosine similarity measure as a proxy for the distance between SCs. The cosine similarity between WoS SCs is obtained from the cross-citation pattern among WoS SCs for the year 2010 articles (Leydesdorff et al. 2013). The pairwise cosine similarity between WoS SC  $i$  and WoS SC  $j$  is calculated by the following formula.

$$\cos(i, j) = \frac{\overrightarrow{CITE}_i \cdot \overrightarrow{CITE}_j}{|\overrightarrow{CITE}_i| \times |\overrightarrow{CITE}_j|}$$

$i, j$  are indexes for category,  $\overrightarrow{CITE}_i$ : that normalize the citing pattern of the category ‘ $i$ ’ to other categories.

The index is 1 if the citing pattern of category ‘ $i$ ’ and category ‘ $j$ ’ perfectly overlap. If the two categories do not share any citing to other WoS SCs, the result is 0. The cosine similarity between WoS SCs quantifies how the two WoS SCs are similarly associated with other WoS SCs. It represents how close or distant two different WoS SCs are ranges from 0 to 1.

The specialization score measures how much a developed research outcome is related to a particular field of research. It employs WoS SCs assigned to a set of journals. As a publication set contains more diverse WoS SCs, the specialization score of the publication records takes a lower value. The specialization score is calculated with the following formula:

$$S_k = \frac{\sum(f_i \times f_j \times \cos(i, j))}{\sum(f_i \times f_j)}$$

$f_i$  is the frequency of category ‘ $i$ ’ among all the assigned categories to the publication or patent ‘ $k$ .’ The  $\cos(i, j)$  represents the technological (disciplinary) cosine similarity between categories ‘ $i$ ’ and ‘ $j$ ’.

Suppose that a set of graphene publications in one publication year has 6 SC1’s and 2 SC2’s. The cosine similarity between SC1 and SC2 is given as 0.2. The frequency of SC1 and SC2 in the pooled WoS SCs are 6/8 and 2/8 respectively. The self-cosine similarity (i.e.,  $\cos(\text{SC1}, \text{SC1})$  and  $\cos(\text{SC2}, \text{SC2})$ ) is set to 1. According to the formula, the aggregated specialization score of the set of papers is computed as  $(3/4)^2 + 0.2 \times (3/4) \times (1/4) + 0.2 \times (3/4) \times (1/4) + (1/4)^2 = 0.7$ .

Based on the formula, we compare the aggregated level specialization score for NEDD and Graphene publications for each publication year.

### Patent analysis

The specialization score analysis for the patent dataset requires steps similar to those in the journal article analysis. However, in patent analysis, simply using the IPC category causes two problems. First, the analysis is biased by the “IPC population problem”. Some IPC categories (such as in bioscience areas) have a large number of patents, but others have very few. If we do not balance the IPC population, the metric calculated with the IPC categories will be biased in favor of the large IPC categories. Thus, a more appropriate grouping of IPC categories is needed to evenly represent the number of patents across the patent system (Kay et al. 2014). To address the IPC population problem, we use the IPC category folding algorithm that was suggested by Kay et al. (2014). This algorithm folds

**Table 1** Publication and patent specialization score analysis

Specialization score for research publications	Specialization score for patents
WoS SCs in a set of publication records	T-IPCs (Transformed IPCs) in a set of patents
Distance between WoS SCs based on 2010 WoS articles' cross citation pattern	Cosine similarity between T-IPCs estimated by the citing-cited T-IPCs pattern in a given set of EPO patents.
Aggregated level specialization score analysis by publication year	Aggregated-level specialization score analysis by patent publication year
Interdisciplinarity between Scientific Research fields (Scientific knowledge)	Diversity of technology fields related to the patents

the small IPC categories into a general group, subclass, or class until the folded IPC's population reaches a given cutoff. The folded IPC categories (hereafter "transformed IPC", or "T-IPC") are created from the IPC categories within the sampled EPO patents. The IPC categories that are not part of the sampled EPO patents could not be transformed into the T-IPC. Also, if a T-IPC cannot reach the given population, even after being transformed to its highest level, this IPC is dropped from the T-IPC group. In our research, we take the following steps: (a) collect all IPC categories that appear in the Graphene and NEDD patents, (b) check whether the collected IPCs can be transformed into the existing T-IPCs, (c) investigate if there are missing IPC categories that cannot not be matched into the T-IPCs and, subsequently, (d) fold them into the existing T-IPCs.

The second problem with using just the IPCs is that separate IPC categories do not always point to different technology fields. IPC categories are often assigned together in a patent because technological fields are often close to each other (short technology distance). For example, A61K 39/395, the IPC code for "antibiotics", is likely to be assigned to patents that also have the C12N 15/xx designation for "genetic engineering". To quantify this technological distance, we employ cosine similarity between two T-IPCs categories, which is based on the citing-to-cited pattern in the base set of EPO patents (Kay et al. 2014). This approach profiles the distribution of cited T-IPCs in a set of patents that have a particular T-IPC. In this way, each T-IPC has a vector that presents the frequency of cited T-IPCs. The citing-cited patterns of T-IPCs are employed to compute the cosine similarity between T-IPCs. The cosine similarity is computed by the formula given in the previous section.

The patent specialization score measures how much a developing technology is related to a particular technological application. As patents have more diverse T-IPC categories, the patent's specialization score takes a lower value. Calculating the patent specialization score employs a similar formula to that of the publication specialization score. As with the publication records analysis, we conduct the patent analysis at the aggregated level of specialization scores for NEDD and Graphene. Table 1 summarizes our method to analyze the specialization score of patents and publications.

## Results

### Descriptive statistics

We collected 61,451 NEDD publications from WoS. The publication years for NEDD journal articles range from 2000 to 2012. Data cleaning, including identifying those with

**Table 2** Descriptive Statistics

Technology	NEDD		Graphene	
	Publication	Patents	Publication	Patents
Statistics \ type of records				
Number of records	59,798	7796	24,381	4340
Aggregated specialization score	0.12	0.51	0.30	0.15
Max (specialization score)	1	1	1	1
Min (specialization score)	0.21	0.15	0.27	0.15
Number of SCs (TIPCs)/record	1.85	7.27	1.99	2.74
Publication year	2000–2012	2000–2012	2004–2012	2004–2012
Analysis coverage <sup>a</sup>	97.3 %	99.8 %	99.9 %	99.7 %

<sup>a</sup> No. of records that have valid T-IPC or WoS SCs/(total population of publication or patents)

valid bibliographic information, yielded a dataset comprised of 59,798 publication records, or 97.3 % of the collected NEDD articles. The total number of assigned WoS SCs in the NEDD articles is 110,433, which is equivalent to 1.85 SCs per individual article. We obtained 24,389 Graphene publications and found that 24,381 of these publications or 99.9 % had valid bibliographic information. The publication years for the Graphene articles ranged from 2004 to 2012. The total number of WoS SCs assigned to the Graphene articles is 48,634. On average, each Graphene publication has about 1.99 WoS SCs.

We initially obtained 7905 and 6459 patents for NEDD and Graphene, respectively. We dropped patent records that lacked publication years. For family patents, the patent that was first published was selected and we dropped the other patents in the family.<sup>1</sup> In order to synchronize the time window with the publication records, we only used the patents that were published between 2000 and 2012 for NEDD and between 2004 and 2012 for Graphene. The result is 7814 NEDD and 4383 Graphene patents. Among these patents, 7796 NEDD patents and 4340 Graphene patents have valid T-IPC information, covering 99.8 % of the given NEDD and 99.7 % of the Graphene patents. The total number of T-IPCs assigned to the NEDD patents is 56,642. On average, one NEDD patent has 7.27 T-IPCs. The Graphene patents have 11,909 T-IPCs. On average, one Graphene patent has 2.74 T-IPCs. Table 2 summarizes the descriptive statistics of the datasets.

## Annual trends of the specialization scores

### *Publication analysis*

We began by collecting the WoS SCs for each record and counting them by publication year. We represented these data through a vector denoting the number of WoS SCs that are assigned to Graphene and NEDD journal articles, respectively, per year. Using this vector, we calculated a specialization score. This approach is consistent with the method used by Porter and his colleagues (2008) in that we grouped the articles and calculated a specialization score for the set of papers across years and technology domains rather than for individual publications. Table 3 summarizes annual trends in the specialization scores for NEDD and Graphene publications.

<sup>1</sup> This study analyzes the publication records by the “publication year” rather than “submitted year”. To align the time with the publications, we use the publication year of patents rather than the application year.

**Table 3** Publication specialization score annual trend and number of WoS SCs per article

Publication year	No. of records		No. of WoS SCs/article		Aggregated specialization score	
	NEDD	Graphene	NEDD	Graphene	NEDD	Graphene
2000	1419	–	1.77	–	0.18	–
2001	1552	–	1.69	–	0.17	–
2002	1812	–	1.73	–	0.16	–
2003	2124	–	1.74	–	0.16	–
2004	2793	161	1.7	1.81	0.15	0.33
2005	3449	205	1.69	1.77	0.15	0.26
2006	4122	355	1.79	1.66	0.13	0.3
2007	4943	809	1.81	1.75	0.12	0.32
2008	6112	1454	1.78	1.76	0.12	0.33
2009	6528	2334	1.9	1.85	0.13	0.32
2010	7828	3807	1.92	2.01	0.13	0.32
2011	9261	6092	1.96	2.08	0.14	0.3
2012	7855	9164	1.92	2.05	0.14	0.29
Overall	59,798	24,381	1.85	1.99	0.12	0.3

We derived the following two observations from Table 3: First, the aggregated specialization score of Graphene articles is larger than the aggregated specialization score of NEDD publications across the given time horizon (beginning in 2004). Second, the aggregated specialization scores of NEDD and Graphene articles are stabilized *over* the course of a year. This pattern shows that the research fields associated with the journals publishing Graphene and NEDD articles have not changed much over time. The difference between the aggregated specialization scores for NEDD and Graphene is likely to be due to a fundamental difference between the two research fields rather than stemming from time trends *per se*.

### Patent analysis

Table 4 summarizes the patent analyses. The number of Graphene patent applications has grown over the years. In contrast, the number of NEDD patent applications has been relatively stagnant. Considering that the number of nanotechnology patent applications has a similar pattern with the NEDD during the comparable period across the jurisdictions (Organization for Economic Cooperation and Development 2015) and emerging technology is featured by fast growth (Rotolo et al. 2015), Graphene presents characteristics of an emerging technology that has spun-off from nanotechnology, while NEDD seems more of a focused technological area.

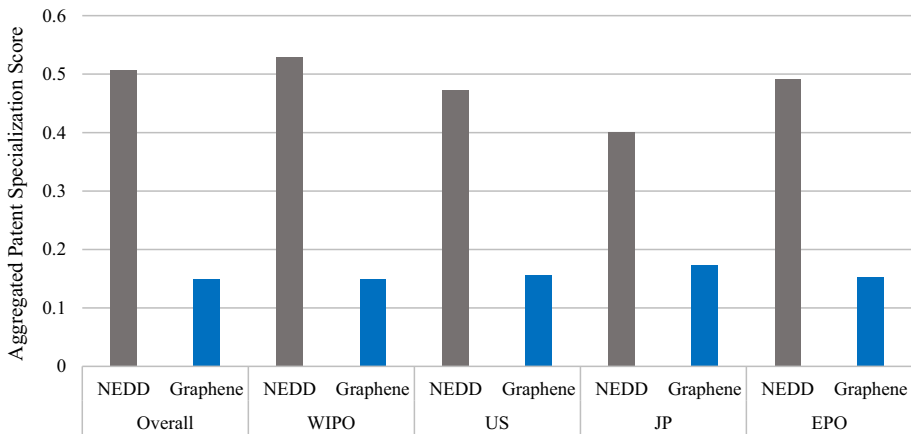
The patent specialization score pattern shows several things. First, the annual specialization scores of NEDD patents and Graphene patents are stable (Table 4, last two columns).

Second, the overall specialization scores of Graphene patents are lower than the scores for NEDD. The yearly comparison indicates that NEDD patents have a higher specialization score than Graphene patents across the given period. This difference is confirmed by comparing specialization scores across the four major destinations of patent filing (WIPO, US, Japan, and EPO) in Fig. 1.



**Table 4** Patent Specialization Score Analyses

Patent publication year	No. of patent applications		No. of T-IPCs/patent		Aggregated level specialization score	
	NEDD	Graphene	NEDD	Graphene	NEDD	Graphene
2000	381	–	8.76	–	0.58	–
2001	686	–	9.57	–	0.59	–
2002	649	–	9.13	–	0.57	–
2003	422	–	7.88	–	0.53	–
2004	438	30	7.72	3.57	0.50	0.23
2005	544	21	7.78	4.43	0.51	0.18
2006	662	23	8.15	3.13	0.52	0.15
2007	642	43	7.58	3.70	0.47	0.15
2008	673	80	7.29	3.59	0.49	0.14
2009	656	204	6.57	3.58	0.50	0.14
2010	742	472	6.77	3.28	0.49	0.15
2011	808	1144	4.37	2.87	0.46	0.16
2012	493	2323	3.79	2.45	0.45	0.16
Overall	7796	4340	7.27	2.74	0.51	0.15



**Fig. 1** Patent specialization score comparison by country

Another finding is that the number of T-IPCs assigned to an NEDD patent is greater than those assigned to a Graphene patent across the time window (Table 4). According to the specialization score formula, the specialization score decreases as more T-IPCs are assigned, all else held constant. The specialization score for the Graphene patents is significantly lower than that of NEDD, although the average number of T-IPCs assigned to a Graphene patent is much smaller than those assigned to an NEDD patent. This observation indicates that the T-IPCs assigned to a Graphene patent are more technologically distant from one another than is the case with the NEDD patents.

As a complementary analysis, we examined how many backward citations were made by and how many subsequent patents were citing the Graphene and NEDD patents

**Table 5** Forward and backward citation trends

Patent publication year	Ratio of mean (forward) and mean (backward)		Forward citation		Backward citation	
	NEDD	Graphene	NEDD	Graphene	NEDD	Graphene
2000	0.73	–	8.02	–	10.97	–
2001	0.69	–	5.48	–	7.99	–
2002	0.52	–	4.89	–	9.41	–
2003	0.54	–	5.48	–	10.20	–
2004	0.54	0.90	5.19	4.80	9.67	5.33
2005	0.38	0.45	4.30	5.62	11.30	12.38
2006	0.35	1.58	3.63	8.22	10.44	5.22
2007	0.38	0.55	2.72	7.09	7.13	12.98
2008	0.31	0.81	2.50	5.68	8.20	7.00
2009	0.29	0.72	1.66	5.73	5.70	8.00
2010	0.27	0.68	1.26	3.00	4.71	4.40
2011	0.27	0.37	0.75	1.27	2.77	3.43
2012	0.09	0.22	0.35	0.33	4.01	1.51
Time average	0.41	0.70	3.56	4.64	7.88	6.69

respectively (forward citation). Such patent citation information provides a sense of the industrial impact, economic value, and novelty of the patented invention (Albert et al. 1991; Trajtenberg 1990; Dahlin and Behrens 2005; Hall et al. 2005), and it sheds light on knowledge flows regarding technology development (Jaffe et al. 1993). Table 5 presents the forward and backward citations for Graphene and NEDD patents by year. Across our time window, Graphene patents received more citations from subsequent patent applications than did NEDD patents. However, the number of backward citations are not significantly different. The ratio of forward citations to backward citation indicates the degree of contribution to developing new technology with the given knowledge input from prior art. This ratio is significantly higher for Graphene than for NEDD patents across all years.

## Discussion

Our analysis is notable in that it reveals opposing patterns in the comparison of specialization scores for NEDD and Graphene in publications versus patents. The aggregated specialization score for NEDD publications is significantly lower than the score for Graphene articles. The difference appears to stem from the closer relatedness among Graphene WoS SCs and their representative research fields. Assuming that the WoS SCs were assigned based on the research fields most closely related to the journal wherein the article is published, we have to conclude that scientific knowledge about NEDD is born from more diversified disciplines than is scientific knowledge about Graphene. Indeed, the most relevant research fields for the journals publishing Graphene articles relate to (applied) material science. However, the body of knowledge for NEDD comprises not only material science but also chemistry and biology.

Interestingly, the patent specialization score analysis produces the opposite result. NEDD patents have significantly higher specialization scores than Graphene patents. This pattern indicates that NEDD technology is more likely to specialize in a particular technological application field than Graphene technology, as combinations of the IPC category represent the potential technological application fields or industrial fields where the patented invention or its function can be used (Hinze, Reiss, and Schmoch 1997; Schankerman 1998; Schmoch et al. 2003).<sup>2</sup> Indeed, the primary purpose of NEDD technology is to enhance the effectiveness of drug delivery to better cure diseases by (a) enabling targeted delivery to organs and tissues, individual cells and organelles and (b) releasing the drug at a controlled rate (Zhou et al. 2014). On the other hand, the lower specialization score associated with the Graphene patents, in comparison to NEDD, implies that Graphene is targeting a broader range of applications. Graphene patents have relatively diverse T-IPC categories that are not so closely related to each other. Indeed, Graphene can be used for multiple purposes, including enhancing the performance of photon sensors, solar cells, display screens, composites, building materials, and water-purification (Segal 2009; Shapira et al. 2012).

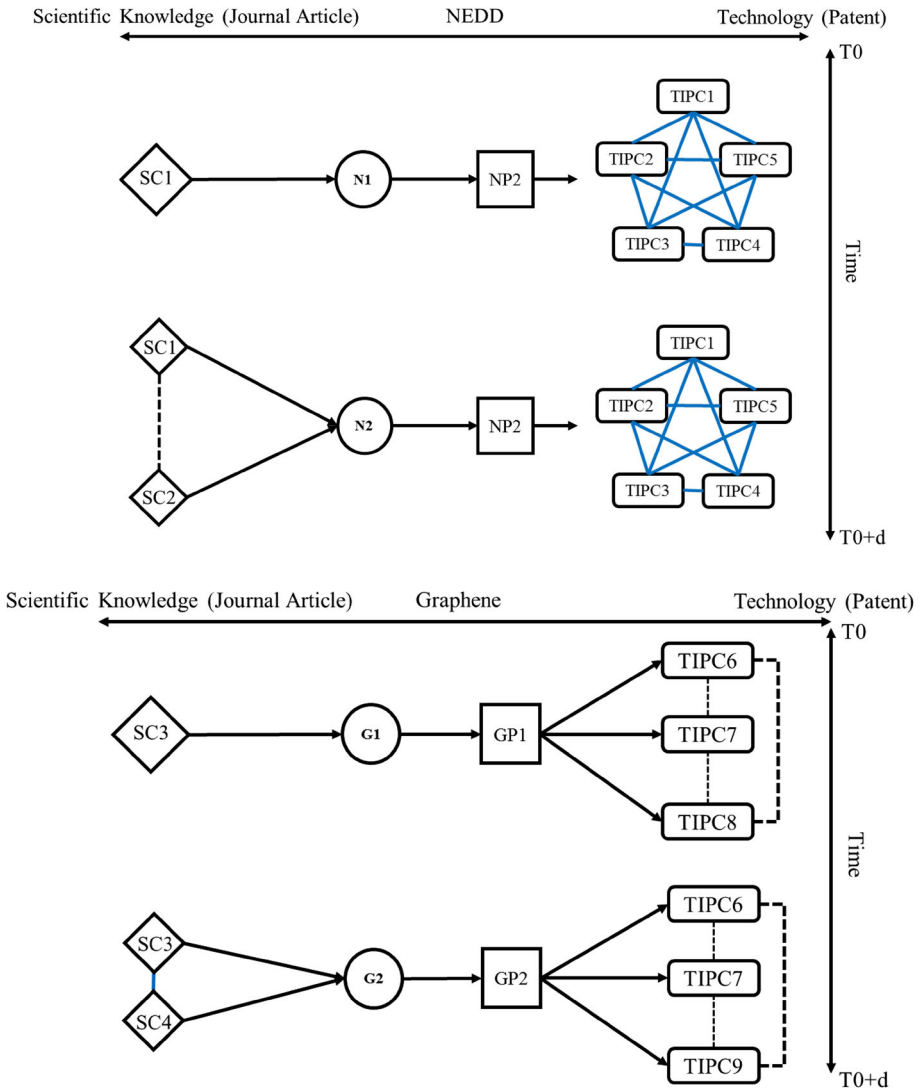
Fewer T-IPCs in Graphene patents, and less obvious relatedness among the T-IPCs, implies that Graphene is very new to the world in terms of its application. The IPC categorization system might not have proper IPC categories that match with the Graphene technology because it is an emerging invention with an application lexicon that may not be fully captured yet in IPC classes. On the other hand, NEDD patents cover technological applications (e.g., medications and pharmaceutical industry applications) that are assigned to extant IPC categories. Therefore, NEDD technology can be mapped more easily than Graphene to IPCs. Graphene's newness can be also observed via measurement of "technological radicalness" by time-lag bias controlled forward citation (Nagaoka et al. 2010). The result shows that, in general, Graphene patents receive more forward citations than do NEDD patents. The higher rate of forward citation for Graphene, in comparison to NEDD, reconfirms our hypothesis, assuming forward citation as a proxy of the technological importance of the technology for subsequent development as well as for innovation radicalness (Albert et al. 1991; Trajtenberg 1990; Dahlin and Behrens 2005; Hall et al. 2005).

Combining the patents and publication specialization scores provides grounds for looking at the innovation trajectories of the two technologies, ranging from the degree of specialization of their scientific origination to their application. Figure 2 visualizes the innovation trajectory of the Graphene and NEDD based on our analysis.

In the NEDD model, the number of assigned WoS SCs gradually increases (SC1 to SC1 and SC2), but the scientific distance between SC1 and SC2 is long (dashed line). The created scientific knowledge goes through the commercialization process, which results in NEDD patented technologies (NP1 for  $t = T_0$ , NP2 for  $t = T_0 + d$ ). The number of assigned T-IPCs for NEDD patents (TIPC1–TIPC5) and the specialization scores are stabilized across each year because the number of T-IPCs and the interrelationships among the T-IPCs do not change much over time. The technological distance among the T-IPCs in NEDD patents is shorter (blue solid lines) than it is for Graphene, which indicates that commercialized NEDD technology covers more closely related application fields.

---

<sup>2</sup> Our research uses T-IPCs rather than original IPCs in order to address the IPC population problem without manipulating the actual meaning of the categories. Thus, we can consider the T-IPCs as the aggregated level IPCs that still contain information about the technology field or application area where the patent invention can be used.



**Fig. 2** Model of innovation trajectory of NEDD (Top) and Graphene (Bottom). Blue Solid (black dashed): Short (long) Technological (Disciplinary) distance

In the Graphene model, the number of WoS SCs gradually increases (SC3 to SC3 and SC4), but their discipline interrelatedness is closer than that for NEDD (blue solid line between SC3 and SC4). The interrelatedness among assigned T-IPCs in a Graphene patent is smaller and they are less closely correlated than those in NEDD patents (connected with dashed black lines). Even though the interrelatedness among the T-IPCs slightly changes over time (TIPC6, TIPC7, TIPC8 to TIPC6, TIPC7, TIPC9), the diversity of the technological application fields, as represented by the T-IPCs, is still greater than with the diversity for NEDD.

## Conclusions

This paper shows how to employ specialization scores in navigating the innovation trajectory of a technology, from identifying its origins in research papers to determining the broadness of its fields of application in patents. Our analyses lead to the following conclusions.

First, we show that the specialization score, which is an advanced diversity measurement analysis of journal articles and patent data, yields different implications when interpreting the interdisciplinarity (or diversity) of a given body of knowledge. Analysis of journal articles provides information about the research fields of the article's body of knowledge. Patent analysis, however, provides implications for the technological fields where the patented invention can be used. Our case study on Graphene and NEDD reinforces this insight. Those who examine diversity/interdisciplinarity indices such as specialization scores should be aware of this difference.

Second, we combine insights gained from the specialization score analysis for NEDD and Graphene with their publication records and patents. This gives a more comprehensive perspective on the innovation trajectories of the two technologies. Such a comprehensive view would not have been possible if only the publication or patent specialization scores had been employed.

We believe that policymakers or firm managers who are interested in evaluating innovation pathways of an emerging technology could benefit from this IDR/diversity analysis with publications and patents. For example, policymakers who decide on national R&D investment can be better informed of the range of impact and origin of a targeted technology by employing dual analyses. When a government tries to encourage interdisciplinary research to help generate technology that can be used broadly, it might analyze the specialization scores of a set of journal articles and patents that are related to the technology in unison. By jointly analyzing the two indicators, the government can assess the technology in a more systematic manner. Firms also can utilize the combinatorial analyses for their technology investment decisions. If a target technology draws on contributions of various research fields, the firms might need to invest in those relevant research areas. If the targeted technology has broad application fields, the firms should find a coherent business strategy that targets those broad application fields.

Our research has several limitations. First, we compared only two technologies for our case study. Future work would apply our approach to more technological innovations to verify robustness. Second, we only analyze the “level of specialization” in order to determine the innovation trajectory of a technology. However, a full technological innovation trajectory cannot be illustrated with just specialization scores. Analysis of multiple metrics would add more insight in establishing the innovation trajectory of a technology, much like Chavarro et al. (2014) Rafols et al. (2012) and Stirling (2006, 2007) attempted to do in their research. Third, the relatedness between WoS SCs and T-IPCs changes over time. We applied a relatedness measure (cosine similarity between the categories) over a single period, 2010. We believe that we would be able to obtain a more dynamic view regarding the innovation trajectory of technology if we captured the changing technological relatedness over time.

**Acknowledgments** We acknowledge support from the US National Science Foundation under Award #1064146—“Revealing Innovation Pathways: Hybrid Science Maps for Technology Assessment and Foresight” and Award # 0937591—the Center for Nanotechnology in Society at Arizona State University.

The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

- Albert, M., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy*, 20(3), 251–259.
- Almeida, P., & Kogut, B. (1997). The exploration of technological diversity and geographic localization in innovation: Start-up firms in the semiconductor industry. *Small Business Economics*, 9(1), 21–31.
- Breschi, S., Lissoni, F., & Maler, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, 32(1), 69–87.
- Carley, S., & Porter, A. L. (2012). A forward diversity index. *Scientometrics*, 90(2), 407–427.
- Chavarro, D., Tang, P., & Rafols, I. (2014). Interdisciplinarity and research on local issues: Evidence from a developing country. *Research Evaluation*, 23(3), 195–209.
- Dahlin, K. B., & Behrens, D. M. (2005). When is an invention really radical?: Defining and measuring technological radicalness. *Research Policy*, 34(5), 717–737.
- Danneels, E. (2007). The process of technological competence leveraging. *Strategic Management Journal*, 28(5), 511–533.
- Galunic, D. C., & Rodan, S. (1998). Resource recombinations in the firm: knowledge structures and the potential for schumpeterian innovation. *Strategic Management Journal*, 19(12), 1193–1201.
- Glänzel, W., Schubert, A., & Czerwon, H.-J. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics*, 44(3), 427–439.
- Gruber, M., MacMillan, I. C., & Thompson, J. D. (2008). Look before you leap: Market opportunity identification in emerging technology firms. *Management Science*, 54(9), 1652–1665.
- Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *The Rand Journal of Economics*, 36(1), 16–38.
- Hinze, S., Reiss, T., & Schmoch, U. (1997). Statistical analysis on the distance between fields of technology. *FhG-ISI: ISE Working Paper*.
- Jaffe, A. (1986). Technological opportunity and spillovers of R & D: Evidence from firms' patents, profits, and market value. *The American Economic Review*, 76(5), 984–1001.
- Jaffe, A. B. (1989). Characterizing the “technological position” of firms, with application to quantifying technological opportunity and research spillovers. *Research Policy*, 18(2), 87–97.
- Jaffe, A. B., Trajtenberg, M., & Hend, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3), 577–598.
- Kay, L., Porter, A. L., Youtie, J., Rafols, I., & Newman, N. (2015). Mapping graphene science and development: Focused research with multiple application areas. *Bulletin of the American Society for Information Science and Technology*, 41(2), 22–25.
- Kay, L., Youtie, J., Porter, A. L., Rafols, I., & Newman, N. (2014). Patent overlay mapping: Visualizing technological distance. *Journal of the Association for Information Science and Technology*, 65(12), 2432–2443.
- Leydesdorff, L., Alkemade, F., Heimeriks, G., & Hoekstra, R. (2015). Patents as instruments for exploring innovation dynamics: geographic and technological perspectives on “photovoltaic cells”. *Scientometrics*, 102(1), 629–651.
- Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, 94(2), 589–593.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.
- Morillo, F., Bordons, M., & Gómez, I. (2003). Interdisciplinarity in science: A tentative typology of disciplines and research areas. *Journal of the American Society for Information Science and Technology*, 54(13), 1237–1249.
- Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Muñoz-Fernández, F., & Herrero-Solana, V. (2007). Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology*, 58(14), 2167–2179.
- Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., & Muñoz-Fernández, F. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics*, 61(1), 129–145.

- Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the Economics of Innovation* (pp. 1083–1127). Oxford, UK: Elsevier.
- Organization for Economic Cooperation and Development. (2015). *OECD Library*. Retrieved May 13 from OECD.Stat: [https://stats.oecd.org/Index.aspx?DataSetCode=PATS\\_IPC](https://stats.oecd.org/Index.aspx?DataSetCode=PATS_IPC).
- Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, *81*(3), 719–745.
- Porter, A. L., Roessner, D. J., & Heberger, A. (2008). How interdisciplinary is a given body of research? *Research Evaluation*, *17*(4), 273–282.
- Rafols, I. (2014). Knowledge integration and diffusion: Measures and mapping of diversity and coherence. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly Impact: Methods and Practice* (1st ed., pp. 169–190). Switzerland: Springer International Publishing.
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business and Management. *Research Policy*, *41*(7), 1262–1282.
- Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, *61*(9), 1871–1887.
- Rao, R. (1982). Diversity: Its measurement, decomposition, apportionment and analysis. *The Indian Journal of Statistics*, *44*(1), 1–22.
- Rotolo, D., Hicks, D., & Martin, B. (2015). What Is an Emerging Technology? *SPRU Working Paper Series*.
- Schankerman, M. (1998). How valuable is patent protection? Estimates by technology field. *The Rand Journal of Economics*, *29*(1), 77–107.
- Schmoch, U., Laville, F., Patel, P., & Frietsch, R. (2003). *Linking Technology Areas to Industrial Sectors—Final Report to the European Commission*. DG Research1. DG Research.
- Segal, M. (2009). Selling graphene by the ton. *Nature Nanotechnology*, *4*, 612–614.
- Shapira, P., Youtie, J., & Arora, S. (2012). Early patterns of commercial activity in graphene. *Journal of Nanoparticle Research*, *14*(4), 1–15.
- Stirling, A. (2006). On the Economics and Analysis of Diversity. *SPRU Electronic Working Papers Series No.28*.
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal Of The Royal Society*, *4*(15), 707–719.
- Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics*, *21*(1), 172–187.
- Zhou, X., Porter, A. L., Robinson, D. K., Shim, M. S., & Guo, Y. (2014). Nano-enabled drug delivery: A research profile. *Nanomedicine*, *10*(5), 889–896.