A Hybrid Method to Trace Technology Evolution Pathways: A Case Study of 3D Printing

Ying Huang[1], Donghua Zhu[1], Yue Qian[1], Yi Zhang[1, 2], Alan L. Porter[3, 4], Yuqin Liu[5], Ying Guo[1*]

1. School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

2. Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia

3. School of Public Policy, Georgia Institute of Technology, Atlanta, GA, 30332, USA

4. Search Technology, Inc., Norcross, GA, 30092, USA

5. Academy of Printing & Packaging Industrial Technology, Beijing Institute of Graphic Communication, Beijing 102600, China

Corresponding author E-mail: guoying_bit@163.com

Abstract

Whether it be for countries to improve the ability to undertake independent innovation or for enterprises to enhance their international competitiveness, tracing historical progression and forecasting future trends of technology evolution is essential for formulating technology strategies and policies. In this paper, we apply co-classification analysis to reveal the technical evolution process of a certain technical field, using co-word analysis to extract implicit or unknown patterns and topics, and main path analysis to discover significant clues about technology hotspots and development prospects. We illustrate this hybrid approach with 3D printing, referring to various technologies and processes used to synthesize a three-dimensional object. Results show how our method offers technical insights and traces technology evolution pathways, and then help decision-makers guide technology development.

Keywords

Tech Mining; Technology Innovation; Technology Evolution; Main Path Analysis; 3D Printing

Introduction

Currently, fierce market competition drives decision makers to identify possible directions and promising options for technology emergence—to help choose the right path for further R&D plans and research activities. The ability to analyze and monitor the history and the current stage of a particular technology is considered a critical asset for gaining competitive advantage and capturing significant opportunities (Choi and Park 2009). Specifically, mastering technology evolution provides decision support for scientific research management, such as scheming science and technology development plans, optimizing government investment in R&D projects, helping forecast the trend of technology development, and discovering the key technological players positioned to pursue specific technologies.

To reveal the process of a specific technology's evolution, many researchers have attempted to identify current technology structures and to trace technological trends by performing patent analyses (J. Yoon and Kim 2011; Chang et al. 2010; Chen et al. 2005; Ganguli 2004; Lacasa et al. 2003). Early patent analytic methods mostly compared the numbers of patents assigned to different entities over time (e.g., nations, affiliations, inventors and technological fields) (von Wartburg et al. 2005; Bengisu 2003; Harhoff et al. 2003). Though such indicators provide a convenient way to draw the landscape, they cannot describe micro-level technology changes—especially for New and Emerging Science & Technologies (NESTs) that have limited history, fast changing, and developing, and have relatively limited applications in the marketplace (Robinson and Propp 2008).

Compared to journal articles that bring most original research contributions to identify likely sources of information about future developments, patents represent disclosure of an invention, nominally at a later stage of innovation (Martino 2003). The interactions that represent previous knowledge underlying a specific inventive step among patents constitute the citation network. The related analyses, so called patent citation analyses, have been widely used for a broad range of applications, e.g., to investigate knowledge flows and technology diffusion (Liu et al. 2013; Murray 2002; Kajikawa and Takeda 2009; Sorenson et al. 2006; Duguet and MacGarvie 2004; Breitzman and Thomas 2002), to confirm rapid growing demands (Kajikawa et al. 2008), to explain the development trajectory of a specific technology (Verspagen 2007), and to track the role of science in a technological innovation (Shibata et al. 2010; Fleming and Sorenson 2004; M. S. Meyer 2001; Tijssen 2001; McMillan et al. 2000). It is helpful to extract patents to represent key technologies in a citation network and to help gain insights on the evolutionary pattern. Main path analysis (MPA) is one of the most attractive methods to determine the critical developing paths to discover citation trajectories (Hummon and Doreian 1989). However, previous studies ignore the essential role of patent families when constructing citation networks. This is essential because consolidating data into patent families not only helps avoid duplicate data retrieval during a search of patents across patent authorities' databases, but it also helps show the geographical focus of the patentee and the patentee's evaluation of the value of the patent (Simmons 2009).

For such requirements, *tech mining* and patent citation tools have to be well tailored to capture the evolution pathways within a complex evolution progression. Tech mining is a multi-step process to analyze Science, Technology, and Innovation (ST&I) information resources by using text mining, visualization, and communication tools (Porter and Cunningham 2005). It provides empirical knowledge to help researchers and managers assess technological maturation (Porter et al. 1991). Particularly, during the process of identifying technology evolution pathways, tech mining offers a macro perspective covering technology development levels and future trends, and also makes R&D information accessible for micro-level probes as needed. For example, tech mining techniques can be applied to analyze relations among specific actors and technologies within a given research-development-innovation system (Hopkins and Siepel 2013; Porter et al. 2002). Moreover, by combining with semantic-TRIZ, tech mining can help obtain better understanding of developmental trends over a relatively short time series (Vicente Gomila and Palop Marro 2013; Zhang et al. 2014b).

We apply the tech mining approach to address several key questions concerning R&D activities—i.e., *when*, *what*, *where*, and *who*. Each question, in turn, can be answered through pertinent indicators. Some indicators are conceptually straightforward (e.g., the trend in research activity on this emerging technology). In this paper, we also extend tech mining methods to further analysis on patent classification and text, and we call them co-classification analysis and co-word analysis. The combination of these methods can, to some extent, mitigate their respective drawbacks and make use of their strengths in (1) obtaining technical core terms in domain areas; (2) identifying influential nodes of a directed citation network; and (3) discovering significant clues about technology hotspots now and technology development prospects for decision making.

This paper is organized as follows: Section 2 overviews main ideas on how to combine co-classification analysis, co-word, and main path analysis as the foundation for this study; Section 3 details the process to identify patent evolution pathways; as an application of the suggested approach, some insights in evolution pathways for 3D printing are presented in Section 4; Finally, in Section 5, we conclude with the summary, discussion, and further research ideas.

Methodology

The main analytic approaches of this paper include co-classification analysis, co-word analysis, and main path analysis of patent citations. Therein, co-classification analysis and co-word analysis are mainly based on patent classifications, and titles and/or abstracts, respectively. Thus, we classify them into the scope of tech mining. The relationships among these three methods can be indicated as Figure 1. In this process, we use a professional desktop text mining software—VantagePoint (http://www.theVantage-Point.com)—to help identify the fields from raw data and show results through a combination of statistics.

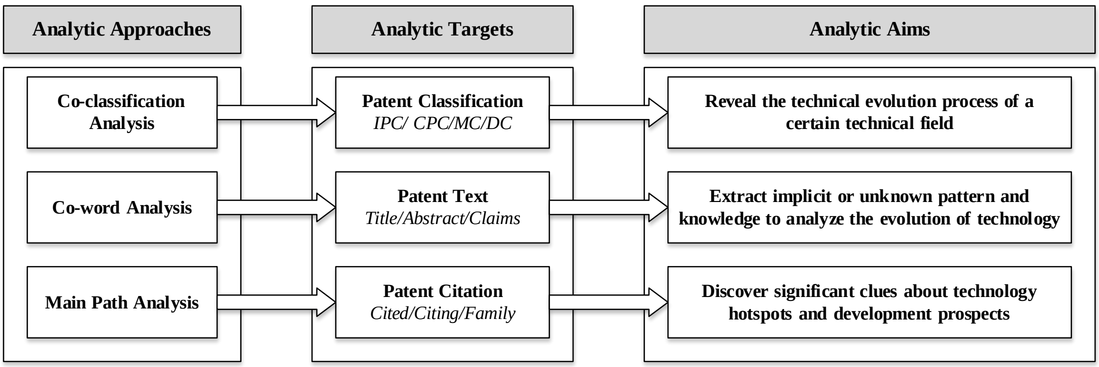

Figure 1. Main Framework of Analyzing Evolution Pathways.

Co-classification Analysis based on Patent Classification

Decision-making activities of knowledge-intensive enterprises depend heavily on the successful classification of patents, which is a reflection of patent technology (Wu et al. 2010). Analyzing patent classification information with some statistical methods along the time axis can reveal the technical evolution process of a certain field. It’s noteworthy that some researchers have utilized the structured information in patent descriptions to analyze the evolution of technology development and to forecast technology development trends (Jun and Lee 2012).

Compared to the International Patent Classification (IPC), the Cooperative Patent Classification (CPC), is a new classification (in effect since January 2013). CPC covers all EPO and U.S. classified documents. The CPC system is based on the IPC structure, considering also three classifications: The European Classification System (ECLA), the In Computer Only (ICO) code, and the U.S. Patent Classification (USPC). This classification contains 250,000 classes—the highest number of subdivisions; thus it is the most granular and precise classification among those in the English versions (Montecchi et al. 2013). Therefore, employing the CPC allows analysis of the parallel development with unprecedented discernment, which so far has been rarely used (Mueller et al. 2015). In this paper, there are three steps to conduct co-classification analysis based on CPCs.

The first step is to build the co-classification matrix. As we know, most patents are related to more than one technology field, so as to belong to multiple patent classifications in one classification system. Thus, if one patent has 6 CPCs in the patent application document, we call the co-occurrence of these 6 CPCs a co-classification relationship. As a result, we can then make the CPC co-classification matrix based on the co-occurrence of CPCs.

The second step is to standardize the co-classification matrix. Different from Salton's cosine and the Pearson correlation, the Jaccard index abstracts from the shape of the distributions and focuses on only the intersection and the sum of the two sets (Leydesdorff 2008). Therefore, the Jaccard coefficient appears to offer a better choice to deal with the co-citation or, more generally, the co-occurrence-matrix. In this paper, we apply the Jaccard coefficient to carry on standardization processing for the co-classification matrix. The rows and columns of the matrix are composed of the frequencies that sub-technologies share in one patent according to its CPCs. As a result, we can then calculate the intensity matrix whose elements measure the diversity among technologies of the co-classification affinity matrix, shown as Table 1.

Table 1. The co-classification intensity matrix.

	CPC_1	**CPC_2**	**…**	**CPC_n**
CPC_1	C_{11}	C_{12}	…	C_{1n}
CPC_2	C_{21}	C_{22}	…	C_{2n}
…	…	…	…	…
CPC_n	C_{n1}	C_{n2}	…	C_{nn}

The formula of calculation of the co-classification intensity matrix is as follows:

$$C_{ij} = \frac{CoC_{ij}}{CoC_i - CoC_j - CoC_{ij}} \quad (1)$$

In formula (1), C_{ij} indicates the co-classification intensity between two technological classifications CPC_i and CPC_j, and the value ranges from 0 to 1 -- the bigger the number is, the stronger the similarity between them. CoC_{ij} is the frequency of co-occurrence between CPC_i and CPC_j, while CoC_i and CoC_i separately indicate the individual frequencies of CPC_i and CPC_j.

The third step is to construct a technology network based on the Girvan-Newman algorithm that is aimed to detect communities by progressively removing edges from the original network. After getting the co-classification matrix of technology intensity, we transform it into a network. Generally, we think the main classifications are located in a max-connected network, which presents a visual, unambiguous technology network. Subsequently, we adopt the Girvan-Newman algorithm (Girvan and Newman 2002) to generate the sub-networks with maximum connectivity, but with less relation among the different sub-networks.

Co-word Analysis based on Patent Textual Information

Co-classification offers an effective way to present the technical evolution process of a certain technical field. However, using patent classification analysis makes it difficult to understand the detailed process of technical evolution, failing to penetrate the patent text; so results often tend to be macroscopic, superficial, and not intuitive. Text mining techniques not only help structure the patent landscape for topical analyses, but also facilitate other analyses, such as patent classification, organization, knowledge sharing, and prior art searches (Tseng et al. 2007). Therefore, in this paper, text mining techniques are introduced to analyze such a corpus to extract intelligence regarding potential technological evolution.

One way of monitoring the trend of a technology is to trace the frequency of specific terms within a given research area. These technical terms are extracted from the abstract fields with special attention to growth in frequency. Additionally, calculating textual similarity based on shared terms goes deeper, which is also related to strong citation links and prior art analysis, infringement analysis, or patent mapping (Moehrle 2010). Some research indicates that the overall relationship among patents provides richer information and thus enables deeper analyses, since it takes more diverse keywords into account (B. Yoon and Park 2004). This method can be used in analyzing up-to-date trends of high technologies and identifying promising avenues for new product development.

In a previous study, one solution to offer detailed insight depends on the "terms" derived from Natural Language Processing (NLP) techniques; however, phrases and terms retrieved in this way are large and "noisy," making them difficult to manually categorize. Using bibliometric and text mining techniques, this paper applies the semi-automatic "Term Clumping" steps, which generate better term lists for achieving competitive technical intelligence (Zhang et al. 2014a). The selected steps of the term clumping process are shown as Figure 2.

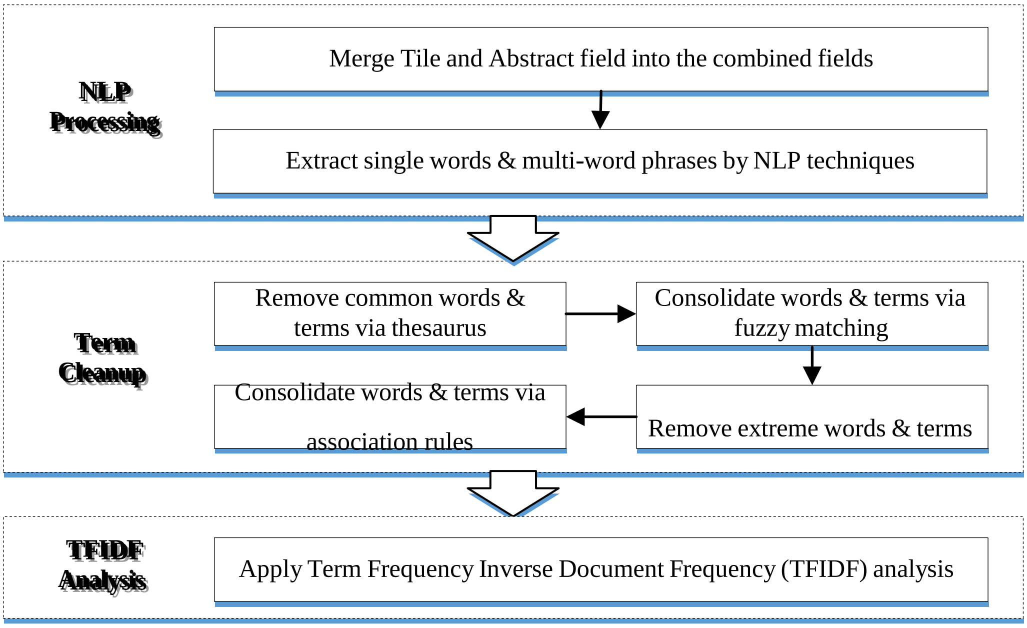

Figure 2. The main process of Term Clumping

First, we combine the abstract field and title field to compress more topical content into one field. We have focused on terms and phrases for quite a long time and have come to determine that in patent intelligence analysis, single words alone are too general in meaning or too ambiguous to indicate a clear concept, and that multi-word phrases could be more specific and desirable. Thus, except for the important single words, multi-word phrases are also extracted by NLP techniques with the help of the VantagePoint software.

Second, we introduce four steps to clean and consolidate the extracted terms: (1) to remove common terms via a thesaurus, (e.g. technology, tool); (2) to consolidate terms via fuzzy matching, where stem and singular & plural forms of English words are recognized; (3) to remove extreme words, [i.e., very common (top 5%) and very rare (occurrence in single records) terms]; and (4) to consolidate terms via association rules, (i.e., sharing words and co-occurrence frequency).

Third, we apply Term Frequency Inverse Document Frequency (TFIDF) analysis to screen the cleaned terms. Identifying the important terms to build the linkage with the evolution of a technology is not completely reliable on the terms' occurrence frequency, but we instead take their emergence in different documents into consideration. The TFIDF involves adding an additional score to the terms that occur in the text under analysis, and can boost scores for neologisms, making them more even with the scores of other terms (Yatsko 2013). Based on the classical formula (Salton and Buckley 1988), we log normalization to the TF to adjust the concise paragraph-size segments of text, such as abstracts. The formula we use in this paper is shown below:

$$\mathrm{TFIDF}_{ij} = \log(TF_{i,j}) \times IDF_{i} = \log(n_{ij}/\textstyle\sum n_{kj}) \times \log(D/d_{i}) \qquad (2)$$

For each term i and document j, $n_{k,j}$ is the number of occurrences of term k in document j, D is the total number of documents in the corpus, and d_i is the number of documents in which term i occurs.

Additionally, expert knowledge is then engaged to refine the outputs of the term clumping process, where some weakly correlated terms are removed and some keywords that indicate the same technological focus are merged. The final keywords reflecting the technology foci are obtained to construct a technology evolution roadmap, building on previous analysis experiences.

Main Path Analysis based on Patent Citation Network

Technological change typically follows along ordered and selective patterns, shaped jointly by technological and scientific principles, and economic and other societal factors (Fontana et al. 2009). Patent text analysis reveals more implicit information in detail since its in-depth character and visualization methods help researchers understand or explain the results more intuitively and clearly. However, such methods are complex and time-consuming; and their results, sometimes, are even vague and not easy to further refine. Patent citation itself represents the evolutionary relations between patent technologies in a certain extent, so mining the patent citation network can study the process of technological evolution and make predictions through exploring such relations (Érdi et al. 2013).

In patent citation analysis, a crucial factor is that patent citations can be included by the applicant and also can be added by the patent examiner responsible for judging the degree of novelty of the patent. Some scholars hold the view that examiner citations are "disturbing noise" and should be removed from patent citations, since these citations sometimes cannot represent the technical spillover between inventors (Jaffe et al. 1993). However, some studies show that there are no evident differences of target area between the two kinds of citations (Alcácer et al. 2009), or that examiner citations to a patent are stronger predictors than inventor citations (Hegde and Sampat 2009). Our current research shows that patents included in an examiner citation network are more specialized in relatively narrow technological fields. Although examiner citation cannot reverse the patent structure of main pathways acquired by analyzing the applicant citation network, it contributes some unique patent nodes that have the potential to activate the process of technological innovation in a target technology field. Therefore, in this paper, we take both examiner citations and inventor citations into consideration to build a more comprehensive and effective patent citation network for MPA.

The main path is defined as a path from a source vertex to a sink vertex with the highest traversal weights on its arcs (De Nooy et al. 2011). Many researchers have used MPA to explore the path of technological development by using bibliographical citation data and/or patent citation data. In our study, four steps are conducted to obtain the critical technology trajectories.

First, merge patents into record families. As mentioned above, a patent family is the collection of patents in different countries referring to the same technical topic (Ho et al. 2014). Citation behavior is different among patent authorities and between parent and child patents; thus, global technology trends cannot be understood with only the analysis of patent data issued by a single authority. For the sake of statistics, the first step is to merge patent documents of a family into a single family record. The family of patents is usually identified by the claim of priority or disclosure, and here, one patent family is marked by the earliest published patent. Meanwhile, all cited patents of a family's members are merged to form the cited patents of the family record.

Second, construct the patent citation network. A general directed network (also called a Bayesian Network) consists of vertices and arcs that link two vertices (nodes). A citation network is a standard directed network that can also be represented as a citation matrix. Its columns and rows stand for the nodes, and each value in the matrix is defined as the strength of citation between two nodes (Choi and Park 2009). While conducting MPA for a given field of technology based on the patent citation network, only citations between patents within the technology field need to be taken into consideration. These effective citations are extracted from the merged family records. In the network, nodes stand for the individual family records, and arcs between two nodes are citations.

Third, calculate the weights of each citation link. How to measure the weights of each citation link from a set of starting vertices to the ending vertices is an important step in MPA. Several indices have been proposed, and the most widespread algorithms, proposed by Hummon and Doreian, are Node Pair Projection Count (NPCC), Search Path Link Count (SPLC), and Search Path Nodes Pair (SPNP) (Hummon and Doreian 1989). In 2003, Batagelj proposed a new traversal count, namely the Search Path Count (SPC), concluding that SPC performs a bit better than SPLC and SPNP, because of its nice properties—even though these indices always obtain almost the same results (Batagelj 2003). However, subtle differences exist among them. In this paper, we do not elaborate on the pros and cons of applying each of the traversal counts but follow the recommendation and apply SPC throughout to count the weight of each citation link.

Fourth, find main paths of the patent citation network. Based on previous phases, technology evolution pathways are finally constructed by identifying the important patents, which locate on the "main trajectory" at different stages. After getting the SPC weight of each node, we need to choose an algorithm to figure out the main path. Most of the traditionally proposed algorithms represent a "local" approach, which repeatedly chooses the link with the largest traversal count emanating from the current starting node. Such local algorithms highlight significance at a particular point in time and track the most significant citation link at every possible splitting point, whereas the global algorithm emphasizes the overall importance and delivers the path with the largest overall traversal count (Ho et al. 2014). In other words, in contrast to the local main path that highlights significance in local progress, the global main path emphasizes the overall importance in knowledge flow (Liu and Lu 2012). Nevertheless, both the local and the global main path may miss the links with the largest traversal count. Liu and colleagues introduced a new method called the "key-route" to enhance MPA; this viewed a main path as an extension of a specific key route and began a search from both ends of that key route (Liu et al. 2013; Liu and Lu 2012). Based on the key-route algorithm, we extract several key routes to determine the most crucial paths in the overall development. The global key-route method not only provides multiple paths (from which we can find the knowledge diffusion trajectory comprehensively), but also contains almost all the important connections and makes the results much more comprehensive. In this paper, we conduct the global key-route method to obtain more technological insights.

Case Study: 3D Printing

The 3D printing technology is used for both prototyping and distributed manufacturing with applications in architecture, industrial design, and biotech (human tissue replacement). The development of 3D printing can be traced back to the mid-1980s. Charles Hull applied for a patent related to stereolithography, and the first commercial rapid prototyping technology, commonly known as 3D printing, emerged in 1985 (Hull 1986). Certainly, the benefits of 3D printing are manifold; for example, it may give rise to production revolution, stimulate creativity, and decrease our environmental problems. In view of these respects, we are eager to know what trend this technology will have in a few years through the path it follows.

A wide range of patent databases has become available [e.g., Derwent Innovations Index (DII); the United States Patent and Trademark Office (USPTO)]. We contend that Thomson Innovation (https://www.thomsoninnovation.com) brings together the world's most comprehensive international patent coverage and powerful Intellectual Property (IP) analysis tools. Compared to the Thomson Innovation, DII lacks the citation of patent information and the USPTO lacks patent family tabulation; thus, we collect data from Thomson Innovation (that incorporates Derwent patent information).

The search query we used is "TABD= (((3D OR 3-D OR (3 ADJ dimension*) OR (three ADJ2 dimension*) OR additive) NEAR (print* OR fabricat* OR manufactur* OR product*)))", which was directed to search the title and abstract fields. Besides, in consideration of the time lag for when patents are filed, we refined the publication period to 1985 through 2014, while we performed the search on January 9, 2016. Ultimately, we received 7,975 records. The reason for setting 1985 as the beginning year for the acknowledged and first published 3D printing related patent is that EP171069 was applied to the 3D system in 1985.

In this stage, we first disassembled the IPC subclass of all targeted patents to get a glimpse of the technological area distribution. The result is that B29C (shaping or joining of plastics; shaping of substances in a plastic state, in general; after-treatment of the shaped products) is mentioned in 2,785 patents, accounting for 34.92% of the total 3D patents. The result is followed by B22F (working metallic powder; manufacture of articles from metallic powder; making metallic powder), which occupies 9.78% of the dataset (780 records). G06F (electric digital data processing), H01L (semiconductor devices; electric solid state devices) and B41J (typewriters; selective printing mechanisms) take up the next highest proportion. Furthermore, we recombine the IPC categories to reflect a finer distribution of patents by introducing patent overlay mapping (Kay et al. 2014). What stands out among those of the key component research fields is "Plastics"—especially in plastics shaping. The Luminescent field follows—especially in metallic powder (see Figure 3—with larger nodes reflecting more patents). In fact, this result is in accordance with our subjective judgment. We can also note that the fields of Chemistry, Semiconductors, and Foods and Drugs warrant attention.

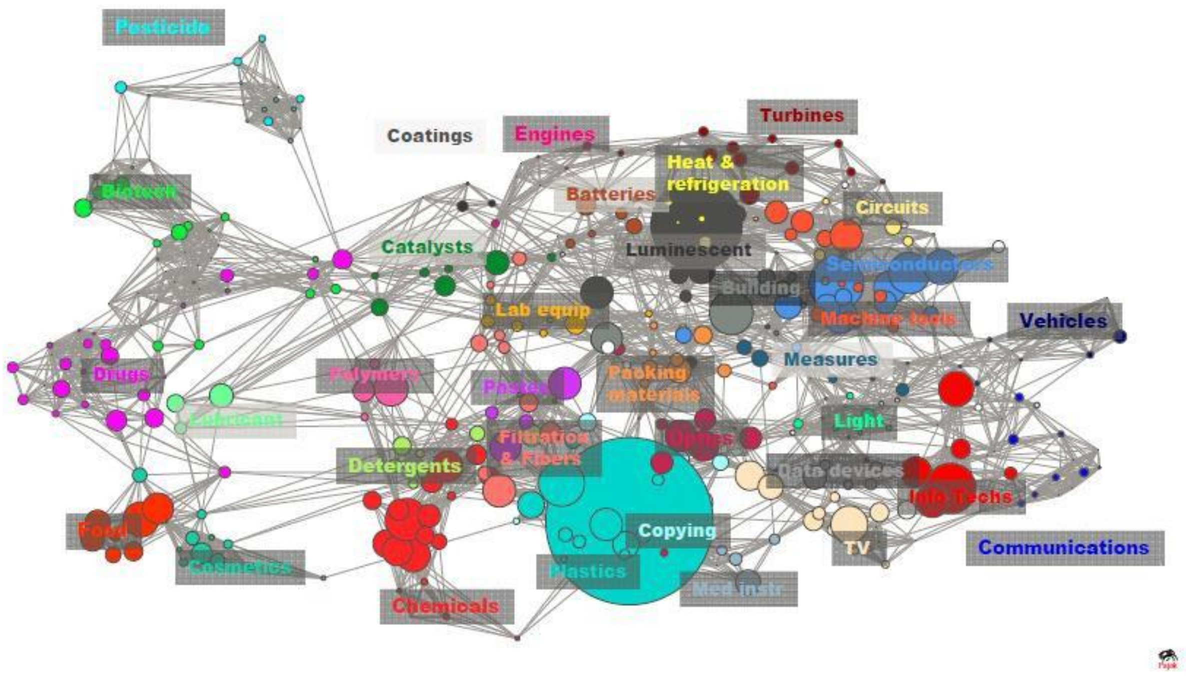

Figure 3. Patent overlap mapping of 3D printing in 1985-2014 by research fields.

As there are high costs for patent application and maintenance, patents pursued in multiple countries tend to have higher technical advantage and perceived commercial potential. Therefore, we chose as our target sub-dataset the patent families that have multiple application countries to capture the leading countries with strong technological strength; only 28.20% of the 3D printing patents (2,249 records) are ultimately selected. Figure 4 uses the Aduna cluster map technique to compare the top ten priority countries and territories by measuring and visualizing ownership ranges to reflect a country's patent performance as a whole. The number after the country code indicates the total number of corresponding assignee countries, and the linkages present the co-applied relationships among patent assignees between countries. This shows the United States (US) as the leading assignee country, followed by Germany (DE) and Japan (JP). Chinese patent assignees owned the most inventions (2,970 records), but only 82 inventions have applications in other countries too. We can discern that China's assignees applied for most of the 3D patents in their home country, while the United States' assignees would rather focus on the global impacts of 3D printing and apply for priority protection worldwide. Therefore, the United States has more advantages to win more potential markets' shares for its competitive technological superiority.

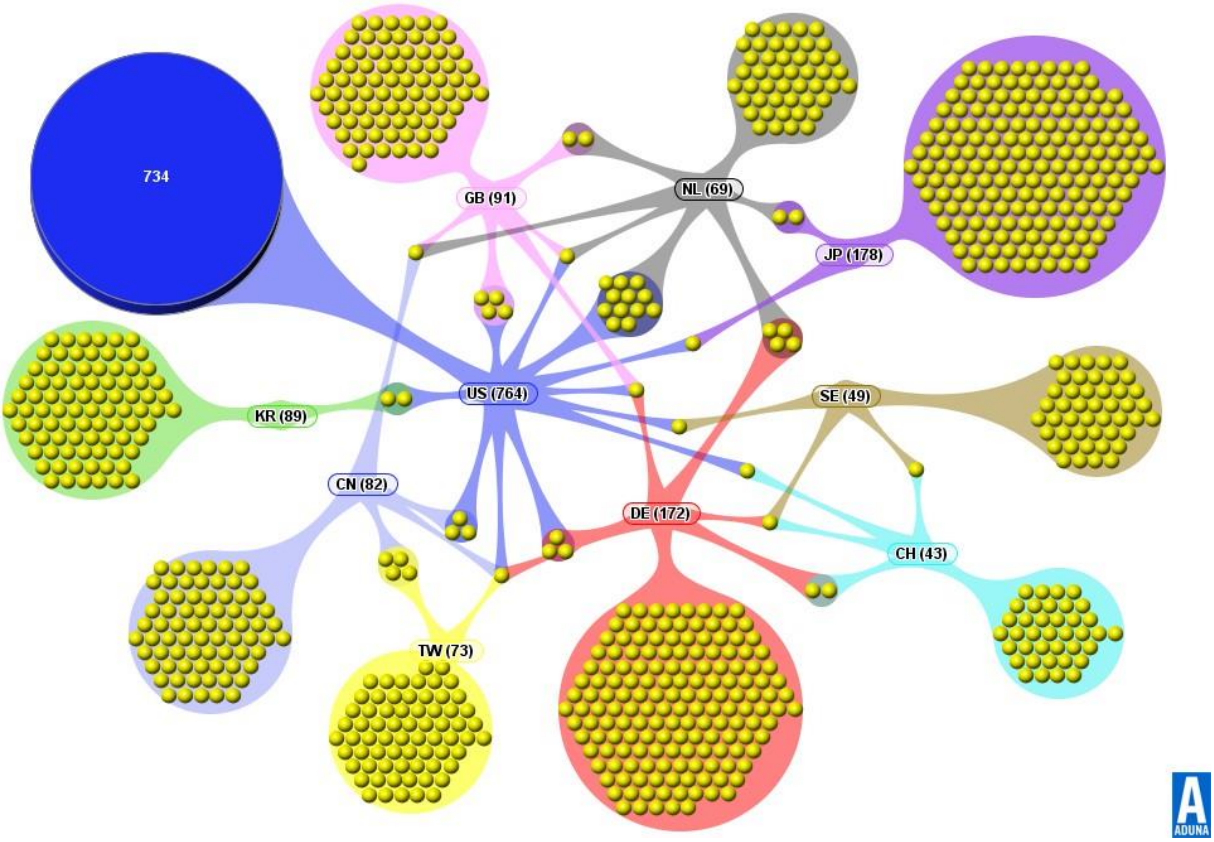

Figure 4. Top 10 priority countries of 3D printing, 1985–2014.

In the early stage of technology development, a few powerful patentees play a vital role. Over the course of technology development, the market grows, competition grows fierce, and the leading organizations lose their absolute dominance. Table 2 shows the top 10 assignees of 3D printing for the period from 1985 to 2014. When we take the whole of published 3D printing patents into consideration, it shows 3D System Inc. as the earliest company devoted to the research of 3D printing, yet it does not make an outstanding performance in terms of patent application numbers. In contrast, some Chinese patent assignees, (e.g. the Institute of Chemistry Chinese Academy of Sciences, Print-Rite Unicorn Image Products Co., Ltd of Zhuhai and Xi'an Zkmt Electronic Technology Equipment Co., Ltd) have a sudden increase in the number of 3D printing-related patents in 2014, which attracts our attention. However, when we turn to the multiple family country patents, the situation changes. Stratasys Inc., the well-known manufacturer of 3D printers and 3D production systems for office-based rapid prototyping and direct digital manufacturing solutions, keeps its leading role in both evaluation criteria. The top four assignees are all from the United States, making up 50% among the top assignees (including the Massachusetts Institute of Technology). Chinese assignees focus more on the domestic market. On the contrary, the assignees who come from the Taiwan region stand out for their multiple family country patents, which reveals that they have intention to pursue international markets.

Table 2. Top 10 assignees of 3D printing during 1985–2014.

NO	Patent Assignees (All)	Records	Patent Assignees (Multiple Family Country)	Records
1	Stratasys (US)	102	Stratasys (US)	56
2	Xi'an Zkmt Electronic Technology Equipment (CN)	75	United Technologies (US)	56
3	Seiko Epson (JP)	71	3D Systems (US)	45
4	Matsushita Electric Works (JP)	69	Hewlett-Packard (US)	45
5	United Technologies (US)	65	Cal Comp Electronics & Communications (TW)	41
6	Print-Rite Unicorn Image Products (CN)	59	Kinpo Electronics (TW)	41
7	Samsung Electronics (KR)	59	Samsung Electronics (KR)	41
8	3D Systems (US)	58	Sanwei Int 3D Printing Technology (TW)	38
9	Hewlett-Packard (US)	56	XYZprinting (TW)	37
10	Chinese Acad Sci, Inst Chem (CN)	55	Massachusetts Inst Technology (US)	30

As technology often advances, an ever-growing pace indicates different development tracks; therefore, it is necessary to consider the technology life cycle (TLC) when creating a distinct R&D strategy plan. The technology life cycle comprises a pattern of dynamic characteristics pertaining to technology, in which its innovative and economic outcomes change over time (Park and Heo 2013). A dominant approach is to analyze TLC using an S-shaped growth curve to model technological performance, either over time or in terms of cumulative R&D expenditures (Gao et al. 2013). The S-curve models adopted by this study are the Loglet Lab model (P. S. Meyer et al. 1999). We used Loglet Lab 2 software (http://phe.rockefeller.edu/LogletLab/2.0/) to analyze the growth curve on the accumulated number of 3D printing patents. Shown in Figure 5, we found that from 1985 to 2000 the patent counts were small and slowly increasing. 3D printing is becoming a hot topic again, with the strong growth in patenting implying tremendous investment in the development of 3D printing. Meanwhile, it shows that organizations perceive potential commercial value in 3D printing technology after 2005. It also shows that the midpoint occurred in 2016, the life cycle was 21.4 years, and the saturation point was around 4,034 patents. Based on the logistic model of TLC (Yung et al. 1999) and preliminary analysis, we think 3D printing technology emerged during the period of 1985-2004, and entered the development growth stage from 2005, and we forecast it enters the maturity stage in 2016.

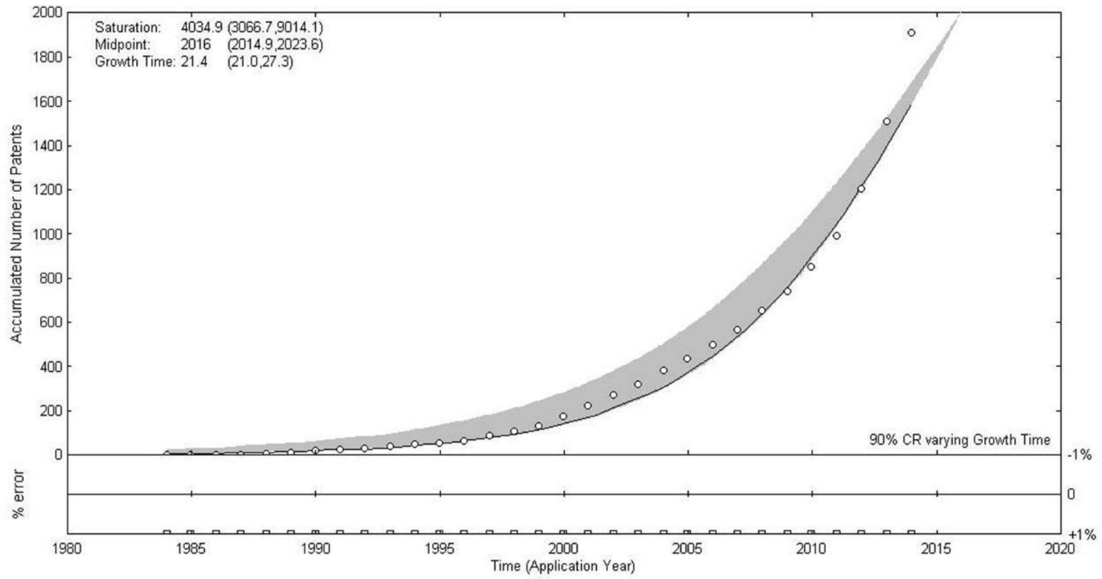

Figure 5. S-curve of 3D printing during 1985–2014.

Based on the methodology we proposed to conduct co-classification analysis based on CPC information, we can map the sub-technology fields network of 3D printing during the period of 1985–2004 (shown as Figure 6) and 2005–2014 (shown as Figure 7) using a software tool for analysis and visualization of large networks—Pajek (http://mrvar.fdv.uni-lj.si/pajek/). From Figure 6, we can see that there are three main sub-technology fields: (1) C12M-B81B-C22F-related fields (the Yellow Cluster), which mainly link to methods or devices of fabricating 3D objects and products; (2) D06M-D06L-C11D -related fields (the Green Cluster), which mainly relate to flexible materials and textiles technology; (3) C07D-C12N-C12R-related fields (the Red Cluster), which is mainly about organic materials applied in 3D printing materials.

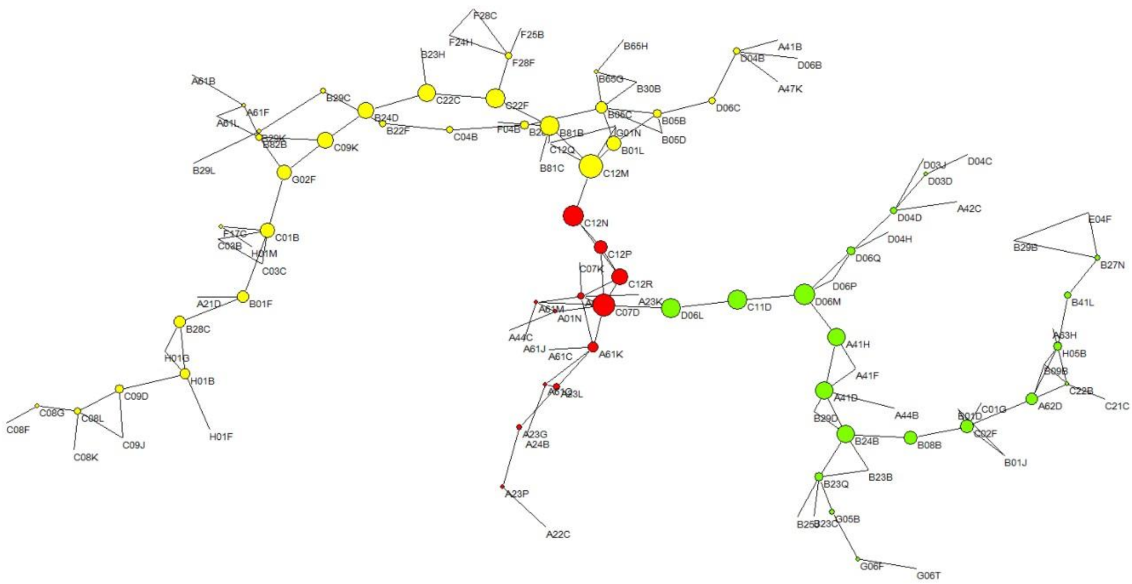

Figure 6. Sub-technology fields network of 3D printing, 1985–2004.

Following similar logic, we can figure out the main sub-technology fields' network of 3D printing during the period of 2005–2014 (shown as Figure 7). There are four essential fields that can be observed: (1) D06M-D06C-D06L-related fields (the Green Cluster), which present different kinds of materials for 3D printing (including fibers, threads, yarns, fabrics, feathers, or fibrous goods made from such materials); (2) B29K-B27N-B26F-related fields (the Blue Cluster), which indicate the close relationship to shaping during the process of fabricating 3D molds or products (including perforating, punching, cutting-out, stamping-out, severing by means other than cutting, and so on); (3) B81B-B60R-B60H-related fields (the Red Cluster), which show some practical applications of 3D printing technology, especially in the transportation area (including the micro-mechanical devices and vehicles); (4) H05B-H02J-related fields (the Yellow Cluster), which refer to circuit arrangements or systems for supplying or distributing electric power. Compared to the emerging stage, the sub-technology fields' network of this stage demonstrates some detailed and sufficient clues for us to further explore the evolution of 3D printing technology.

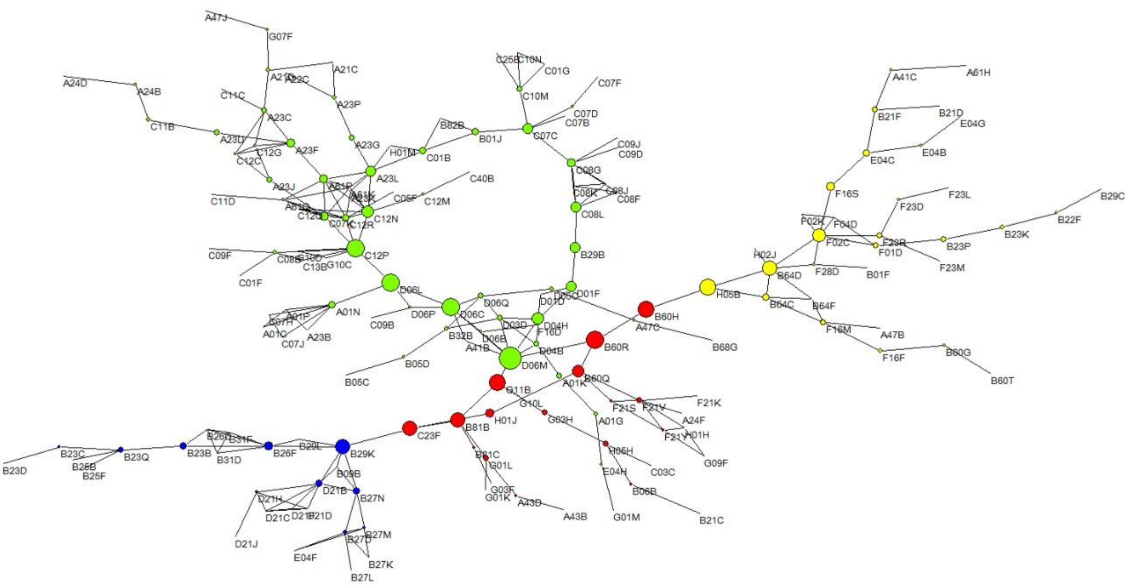

Figure 7. Sub-technology fields network of 3D printing during 2005–2014.

Based on the steps of the cleanup procedure in *Term Clumping*, we initially obtain a list of the top 100 (high-frequency of occurrence) terms for further analyses. After calculating the TFIDF value, we picked up 50 key terms combining previous experience and experts' knowledge. The topographic map of the 50 key terms in 3D printing is shown in Figure 8 for the period of 1985–2004 and in Figure 9 for the period of 2005–2014, which are mapped in the *ItgInsight*, an Intelligence analysis software mainly developed by the laboratory of Knowledge Management and Data Analysis (KMDA) at the Beijing Institute of Technology. The number after the term represents the frequency of a certain term; the color that nodes are embedded with indicate the centrality degree of the nearest term; and the nodes covered in the same mountain share are more similar in topic.

The 50 high TFIDF value terms in 3D printing technology from 1985 to 2004 are clustered into several mountains as follows (shown as Figure 8). The right part seems to be connected more tightly. "polymeric material (13)," and "mound (30)," are the zenith of the upper area, and the former is surrounded by material-related terms (e.g. "polycarbonate (22)") and prototype-related term (e.g. "rapid prototyping (20)"); while the latter is embraced by a processing-related term like "selective laser (31)," "deposition (20)" and so on. In the left area, “raw material (23),” “powder material (19),” and “base material (13)” are outstanding terms that can be applied to track the technological topics. Therefore, we can infer that, in the emerging stage, the block for promoting 3D printing technology is lay in materials. This is mainly due to 3D printing producing component parts layer-by-layer through the additional use of materials. More evidence is shown in the rest of the mountains, of which peaks mostly consist of material-related terms, such as “surfactant (15),” “polypropylene (19),” and so on.

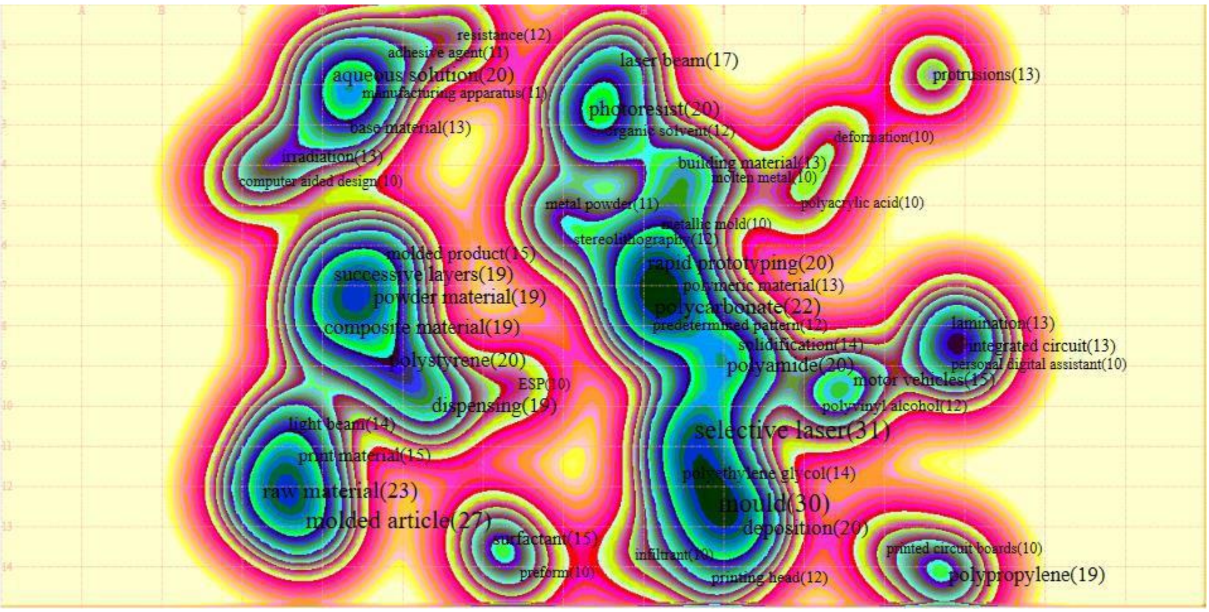

Figure 8. The co-occurrence network of top 50 terms during 1985–2004.

Compared to the first phase, the development of 3D printing has become more advanced and focused. On one hand, current technologies are lacking in accuracy and scale, and also in the ability to produce truly robust parts in a sufficient variety of materials to make desirable consumer or automotive products. Thus, polymeric materials are paid plenty of attention in the past few years, which can be detected from the higher frequency terms, including “polycarbonate (110),” “polyamide (99),” “polypropylene (89),” “polystyrene (87),” “polyvinyl alcohol (81),” “polyethylene (80)” and so on. On the other hand, one of the advantages of using home-based 3D printing to manufacture spare parts is that it enables consumers to manufacture a one-off product at a very low volume with no cost penalty, unlike with traditional manufacturing methods. More and more patent assignees attempt to reduce the manufacturing cost and produce the 3D printing objects and products in a low-cost and environment-friendly way; this trend can be observed by terms such as, “low cost (173),” “efficient manner (85),” and “environment friendly (40).”

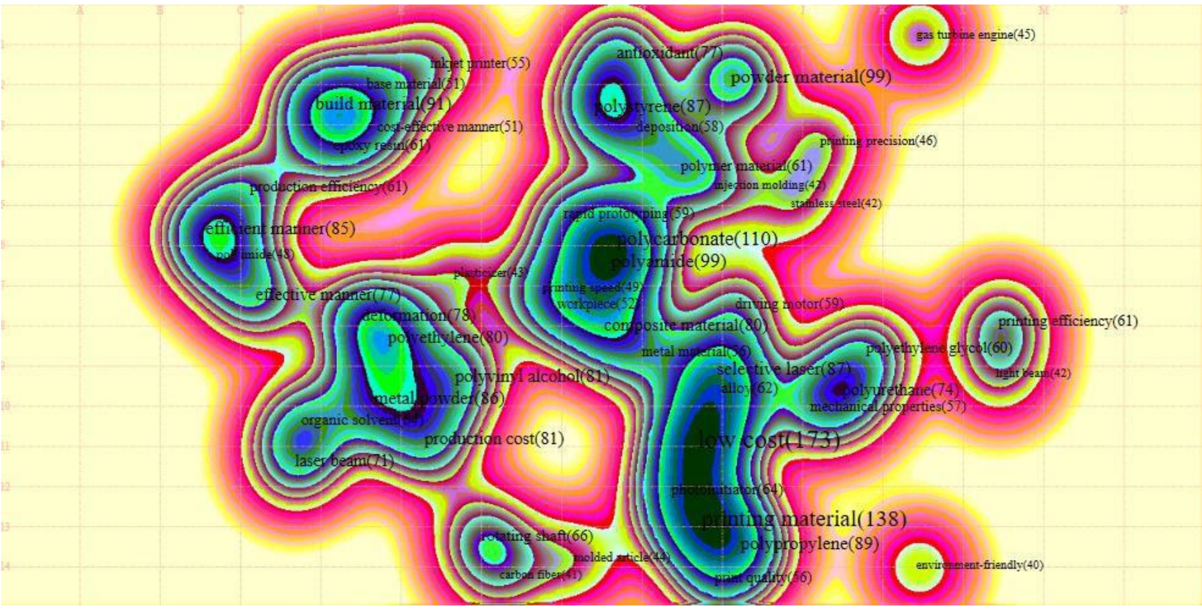

Figure 9. The co-occurrence Network of Top 50 Terms during 2005–2014.

After analyzing the main topics and trends in the 3D printing field, we attempted to discover the evolution pathways in a more detailed way. We used the global key-route MPA. First, we tried different sequential numbers of the routes to seek a reasonable threshold to identify paths exhibiting the greatest weights in the patent citation network, and then created a global main path using these links; the results are shown in Figure 10 and Figure 11.

In terms of the global main path of 3D printing during 1985–2004 (shown as Figure 10), EP171069 is the well-known starting point of the 3D printing technology application, invented by Charles Hull. It presented a 3D object production using stereo-lithography, generating successive cross-sectional laminae of objects at the surface of fluid medium by synergistic stimulation. Thereeafter, EP338751 (rapid prototyping by fabricating a 3D object from representation) and WO1998028124 (stereolithographic system to produce a 3D object) are also proposed by 3D Systems Inc. The fourth node is WO1998009798, whose patent assignee is Z CORP. Z CORP has a co-assignee cooperation with 3D Systems Inc. that introduced a 3D printing material system for offices, consisting of a mixture of particles of adhesive, filler, first fibrous components and solvents. Followed by WO1998009798, there are two linkages that are indicated: one is concerning some materials for use in 3D objects; another describes some potential applications (e.g., representation of bones), involving dispensing homogeneous fluid into dry particulate material.

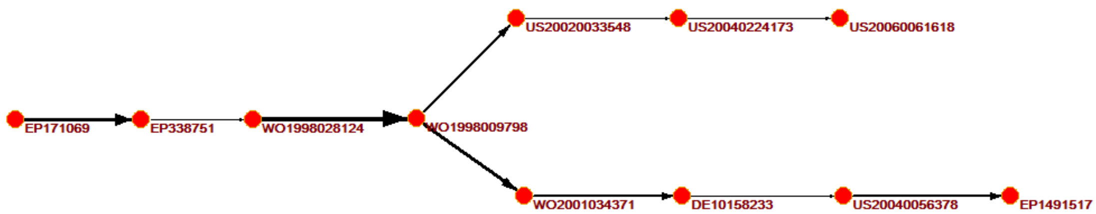

Figure 10. Global Main Path of 3D Printing during 1985–2004.

When we take the whole period into consideration, we have some interesting findings. First, EP171069 and EP338751, as two groundbreaking patent families, play a very important role in pushing 3D printing technology forward for developing a prototype system based on a process known as stereolithography. Second, 3D Systems Inc. and Stratasys Inc., as two leading manufacturers and technology providers in the 3D printing field, have similar citation behaviors and tend to cite their own invitations. Such results possibly tell us another story—that they keep updating their technology innovation to lead the product market. Third, rapid prototyping by fabricating 3D objects is the hot topic in the initial stage (e.g., EP171069, EP338751, WO1998028124, and WO1998009798). As a result, composition materials became a new topic for the materials used in 3D printing of complex structures, which are thought of as a challenging but promising direction.

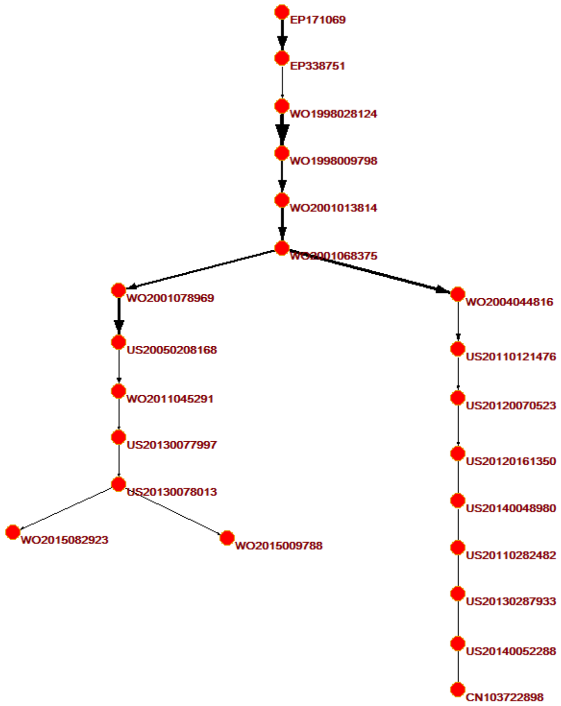

Figure 11. Global Main Path of 3D Printing during 1985–2014.

Conclusion and Discussion

Tracing technology evolution pathways is essential to track innovation progress, but it is a challenge to understand the process of technical evolutions and trends in detail. Many scholars in technology management have sought to extract more technical intelligence to support decision making. However, most previous research focuses on single factors to trace the technology pathway, so it is hard to grasp both comprehensive insights on a macro level and technological features on a micro level. In this paper, we introduce a systematic approach that applies co-classification analysis to reveal the technical evolution process of a certain technical field, co-word analysis to extract implicit or unknown patterns and topics, and main path analysis to discover significant clues about technology hotspots and development prospects.

We conduct a case study in the field of 3D printing. As a "tool-less" and digital approach to production, 3D printing presents companies and consumers with a wide and ever expanding range of technical, economic, and social benefits. Based on proposed analyses we can derive several beneficial insights for technology management. First, after the hobbled development for several years, 3D printing technology is now receiving more attention, and it seems poised to enter the maturity stage presently. The fierce competition in the 3D industry is predictable and the United States' cooperation has a huge advantage over other countries for outstanding patent assignees and powerful technological reserves. 3D Systems Inc. and Stratasys Inc. show at the top of the sector, with the former edging out the latter for supremacy, even though they were not leading players in terms of patent application in recent years. Second, as for the layer-by-layer production characteristic of 3D manufacturing, plastics and shaping are important technological fields for the past twenty years. Rapid prototyping by fabricating 3D objects was the hot topic in the initial stage, and then compounding emerged. More and more patent assignees devoted themselves to reducing manufacturing costs and fabricating products. Reducing machine and material costs along with increased software accessibility will undoubtedly drive growth within the 3D printing market. Therefore, being low cost and environmental friendly are two critical factors on which to focus to win the market competition, particularly for traditional advantage companies. Third, the persistent interest inspired developments in many technologies, such as a wider variety of materials in 3D printing. Different from the traditional manufacturing analogue, composite materials became a new topic for use in 3D printing of complex structures, which are thought of as a challenging but promising direction. The changes in use will lead to different growth rates for different materials; at the same time, the promotion of materials will lead 3D printing to be applied to more sectors and applications.

As an important part of patent analysis, technological evolution analysis can support decision-making for governments' science and technology planning and enterprises' R&D strategy. Furthermore, the systematic approach proposed by this study can be applied to other sectors to reveal the emerging and key technical field, monitor the landscape of patents via topical analysis, and study the process of technological evolution. Compared with the individual approaches, this combo method, to some extent, avoids the drawbacks, and can better locate the leading players and potential market, monitor the historical trajectory of technological development and forecast future development trends by engaging expert knowledge. Clearly, it can handle problems in a wide range of ST&I policy research and provide insights for R&D plan and strategic management.

There are also several limitations of this paper requiring more detailed and specific discussions. On one hand, patent citations are more likely to be incomplete for the time lags between citing and cited patents, which is a great hindrance to know the relations between them overall (especially if we take patent family into consideration). Besides, this research study requires high data quality, but in fact not all patent office and database service companies provide comprehensive historical citation information (Limited organizations have the access to TI database for its high price). On the other hand, both co-classification analysis and co-word analysis are based on the co-occurrence relationship and high frequency terms or classifications. In actual situation, it is important to pay attention to analyzing lower-frequency terms and exploring non-co-occurrence relationship. Unlike scientific publications, inventors tend to comply with the requirements of the Patent Office and disclose the least possible information to protect their intellectual properties, so many terms could not reflect the actual development in the patent. Therefore, we anticipate further study to seek ways to better integrate with patent classification, patent text, and patent citation in detail, both from the theoretical level and practical level.

Acknowledgment

We acknowledge support from the U.S. National Science Foundation (Award No. 1527370), National Natural Science Foundation of China (Grant No. 71503020), MOE (Ministry of Education of the People's Republic of China), Project of Humanities and Social Sciences (Grant No. 13YJC630042), and the Beijing Institute of Technology Basic Research Project (Grant No. 20142142013). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the supporters. We also thank anonymous reviewers who make useful comments to us.

Reference

Uncategorized References

Alcácer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in U.S. patents: An overview and analysis. *Research Policy, 38*(2), 415-427, doi:10.1016/j.respol.2008.12.001.

Batagelj, V. (2003). Efficient Algorithms for Citation Network Analysis. https://arxiv.org/pdf/cs/0309023v1.pdf. Accessed Dec. 13 2015.

Bengisu, M. (2003). Critical and emerging technologies in Materials, Manufacturing, and Industrial Engineering: A study for priority setting. *Scientometrics, 58*(3), 473-487, doi:10.1023/B:SCIE.0000006875.61813.f6.

Breitzman, A., & Thomas, P. (2002). Using patent citation analysis to target/value M&A candidates. *Research-Technology Management, 45*(5), 28-36, doi:10.1080/08956308.2002.11671519.

Chang, P.-L., Wu, C.-C., & Leu, H.-J. (2010). Using patent analyses to monitor the technological trends in an emerging field of technology: a case of carbon nanotube field emission display. *Scientometrics, 82*(1), 5-19, doi:10.1007/s11192-009-0033-y.

Chen, D. Z., Chang, H. W., Huang, M. H., & Fu, F. C. (2005). Core technologies and key industries in Taiwan from 1978 to 2002: A perspective from patent analysis. *Scientometrics, 64*(1), 31-53, doi:10.1007/s11192-005-0236-9.

Choi, C., & Park, Y. (2009). Monitoring the organic structure of technology based on the patent development paths. *Technological Forecasting and Social Change, 76*(6), 754-768, doi:10.1016/j.techfore.2008.10.007.

De Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with Pajek* (Structural Analysis in the Social Sciences). New York: Cambridge University Press.

Duguet, E., & MacGarvie, M. (2004). How well do patent citations measure flows of technology? Evidence from French innovation surveys. *Economics of Innovation & New Technology, 14*(5), 375-393, doi:10.1080/1043859042000307347.

Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., et al. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics, 95*(1), 225-242, doi:10.1007/s11192-012-0796-4.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal, 25*(8-9), 909-928, doi:10.1002/smj.384.

Fontana, R., Nuvolari, A., & Verspagen, B. (2009). Mapping technological trajectories as patent citation networks. An application to data communication standards. *Economics of Innovation & New Technology, 18*(4), 311-336, doi:10.1080/10438590801969073.

Ganguli, P. (2004). Patents and patent information in 1979 and 2004: a perspective from India. *World Patent Information, 26*(1), 61-62, doi:10.1016/j.wpi.2003.10.015.

Gao, L. D., Porter, A. L., Wang, J., Fang, S., Zhang, X., Ma, T. T., et al. (2013). Technology life cycle analysis method based on patent documents. *Technological Forecasting and Social Change, 80*(3), 398-407, doi:10.1016/j.techfore.2012.10.003.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America, 99*(12), 7821-7826, doi:10.1073/pnas.122653799.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy, 32*(8), 1343-1363, doi:10.1016/s0048-7333(02)00124-5.

Hegde, D., & Sampat, B. (2009). Examiner citations, applicant citations, and the private value of patents. *Economics Letters, 105*(3), 287-289, doi:10.1016/j.econlet.2009.08.019.

Ho, M. H. C., Lin, V. H., & Liu, J. S. (2014). Exploring knowledge diffusion among nations: a study of core technologies in fuel cells. *Scientometrics, 100*(1), 149-171, doi:10.1007/s11192-014-1265-z.

Hopkins, M. M., & Siepel, J. (2013). Just how difficult can it be counting up R&D funding for emerging technologies (and is tech mining with proxy measures going to be any better)? *Technology Analysis & Strategic Management, 25*(6), 655-685, doi:10.1080/09537325.2013.801950.

Hull, C. W. (1986). Apparatus for production of three-dimensional objects by stereolithography. Google Patents.

Hummon, N. P., & Doreian, P. (1989). Connectivity in a citation network: the development of DNA theory. *Social Networks, 11*(1), 39-63, doi:10.1016/0378-8733(89)90017-8.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics, 108*(3), 577-598, doi:10.2307/2118401.

Jun, S., & Lee, S.-J. (2012). Emerging technology forecasting using new patent information analysis. *International Journal of Software Engineering & Its Applications, 6*(3), 107-115.

Kajikawa, Y., & Takeda, Y. (2009). Citation network analysis of organic LEDs. *Technological Forecasting and Social Change, 76*(8), 1115-1123, doi:10.1016/j.techfore.2009.04.004.

Kajikawa, Y., Yoshikawa, J., Takeda, Y., & Matsushima, K. (2008). Tracking emerging technologies in energy research: Toward a roadmap for sustainable energy. *Technological Forecasting and Social Change, 75*(6), 771-782, doi:10.1016/j.techfore.2007.05.005.

Kay, L., Newman, N., Youtie, J., Porter, A. L., & Rafols, I. (2014). Patent overlay mapping: Visualizing technological distance. *Journal of the Association for Information Science and Technology, 65*(12), 2432-2443, doi:10.1002/asi.23146.

Lacasa, I. D., Grupp, H., & Schmoch, U. (2003). Tracing technological change over long periods in Germany in chemicals using patent statistics. *Scientometrics, 57*(2), 175-195, doi:10.1023/a:1024133517484.

Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's Cosineversus the Jaccard index. *Journal of the American Society for Information Science and Technology, 59*(1), 77-85, doi:10.1002/asi.20732.

Liu, J. S., & Lu, L. Y. Y. (2012). An Integrated Approach for Main Path Analysis: Development of the Hirsch Index as an Example. *Journal of the American Society for Information Science and Technology, 63*(3), 528-542, doi:10.1002/asi.21692.

Liu, J. S., Lu, L. Y. Y., Lu, W. M., & Lin, B. J. Y. (2013). Data envelopment analysis 1978-2010: A citation-based literature survey. *Omega-International Journal of Management Science, 41*(1), 3-15, doi:10.1016/j.omega.2010.12.006.

Martino, J. P. (2003). A review of selected recent advances in technological forecasting. *Technological Forecasting and Social Change, 70*(8), 719-733, doi:10.1016/S0040-1625(02)00375-X.

McMillan, G. S., Narin, F., & Deeds, D. L. (2000). An analysis of the critical role of public science in innovation: the case of biotechnology. *Research Policy, 29*(1), 1-8, doi:10.1016/S0048-7333(99)00030-X.

Meyer, M. S. (2001). Patent citation analysis in a novel field of technology: An exploration of nano-science and nano-technology. *Scientometrics, 51*(1), 163-183, doi:10.1023/a:1010572914033.

Meyer, P. S., Yung, J. W., & Ausubel, J. H. (1999). A primer on logistic growth and substitution: The mathematics of the Loglet Lab software. *Technological Forecasting and Social Change, 61*(3), 247-271, doi:10.1016/s0040-1625(99)00021-9.

Moehrle, M. G. (2010). Measures for textual patent similarities: a guided way to select appropriate approaches. *Scientometrics, 85*(1), 95-109, doi:10.1007/s11192-010-0243-3.

Montecchi, T., Russo, D., & Liu, Y. (2013). Searching in Cooperative Patent Classification: Comparison between keyword and concept-based search. *Advanced Engineering Informatics, 27*(3), 335-345, doi:10.1016/j.aei.2013.02.002.

Mueller, S. C., Sandner, P. G., & Welpe, I. M. (2015). Monitoring innovation in electrochemical energy storage technologies: A patent-based approach. *Applied Energy, 137*, 537-544, doi:10.1016/j.apenergy.2014.06.082.

Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy, 31*(8-9), 1389-1403, doi:10.1016/s0048-7333(02)00070-7.

Park, J., & Heo, E. (2013). Patent quality determinants based on technology life cycle with special reference to solar-cell technology field. *Maejo International Journal of Science and Technology, 7*(2), 315-328.

Porter, A. L., & Cunningham, S. W. (2005). *Tech mining : Exploiting new technologies for competitive advantage* (Systems Engineering and Management). New York: John Wiley & Sons.

Porter, A. L., Kongthon, A., & Lu, J.-C. (2002). Research profiling: Improving the literature review. *Scientometrics, 53*(3), 351-370, doi:10.1023/a:1014873029258.

Porter, A. L., Roper, A. L., Mason, T. W., Rossini, F. A., & Banks, J. (1991). *Forecasting and Management of Technology* (Engineering Technology Management). New York: John Wiley & Sons.

Robinson, D. K., & Propp, T. (2008). Multi-path mapping for alignment strategies in emerging science and technologies. *Technological Forecasting and Social Change, 75*(4), 517-538, doi:10.1016/j.techfore.2008.02.002.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing and Management, 24*(5), 513-523, doi:10.1016/0306-4573(88)90021-0.

Shibata, N., Kajikawa, Y., & Sakata, I. (2010). Extracting the commercialization gap between science and technology - Case study of a solar cell. *Technological Forecasting and Social Change, 77*(7), 1147-1155, doi:10.1016/j.techfore.2010.03.008.

Simmons, E. S. (2009). “Black sheep” in the patent family. *World Patent Information, 31*(1), 11-18, doi:10.1016/j.wpi.2008.08.005.

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy, 35*(7), 994-1017, doi:10.1016/j.respol.2006.05.002.

Tijssen, R. J. W. (2001). Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. *Research Policy, 30*(1), 35-54, doi:10.1016/S0048-7333(99)00080-3.

Tseng, Y.-H., Lin, C.-J., & Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information Processing & Management, 43*(5), 1216-1247, doi:10.1016/j.ipm.2006.11.011.

Verspagen, B. (2007). Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research. *Advances in Complex Systems, 10*(1), 93-115, doi:10.1142/s0219525907000945.

Vicente Gomila, J. M., & Palop Marro, F. (2013). Combining tech-mining and semantic-TRIZ for a faster and better technology analysis: a case in energy storage systems. *Technology Analysis & Strategic Management, 25*(6), 725-743, doi:10.1080/09537325.2013.803065.

von Wartburg, I., Teichert, T., & Rost, K. (2005). Inventive progress measured by multi-stage patent citation analysis. *Research Policy, 34*(10), 1591-1607, doi:10.1016/j.respol.2005.08.001.

Wu, C.-H., Ken, Y., & Huang, T. (2010). Patent classification system using a new hybrid genetic algorithm support vector machine. *Applied Soft Computing, 10*(4), 1164-1177, doi:10.1016/j.asoc.2009.11.033.

Yatsko, V. A. (2013). Advantages and disadvantages of TF*IDF term weighting. *In the World of Scientific Discoveries, 2013*(42), 229-244.

Yoon, B., & Park, Y. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. *The Journal of High Technology Management Research, 15*(1), 37-50, doi:10.1016/j.hitech.2003.09.003.

Yoon, J., & Kim, K. (2011). Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks. *Scientometrics, 88*(1), 213-228, doi:10.1007/s11192-011-0383-0.

Yung, J. W., Meyer, P. S., & Ausubel, J. H. (1999). The Loglet Lab software: A tutorial. *Technological Forecasting and Social Change, 61*(3), 273-295, doi:10.1016/s0040-1625(99)00023-2.

Zhang, Y., Porter, A. L., Hu, Z. Y., Guo, Y., & Newman, N. C. (2014a). "Term clumping" for technical intelligence: A case study on dye-sensitized solar cells. *Technological Forecasting and Social Change, 85*, 26-39, doi:10.1016/j.techfore.2013.12.019.

Zhang, Y., Zhou, X., Porter, A. L., & Vicente Gomila, J. M. (2014b). How to combine term clumping and technology roadmapping for newly emerging science & technology competitive intelligence: "problem & solution" pattern based semantic TRIZ tool and case study. *Scientometrics, 101*(2), 1375-1389, doi:10.1007/s11192-014-1262-2.