Automation in Construction 87 (2018) 215–224

Contents lists available at ScienceDirect

Automation in Construction

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

The structure and knowledge flow of building information modeling based on patent citation network analysis

Yoo-Na Park[a], Yoon-Sun Lee[b,*], Jae-Jun Kim[a], Tai Sik Lee[c]

[a] Department of Architectural Engineering, Hanyang University, Seoul 04763, Republic of Korea
[b] International Space Exploration Research Institute, Hanyang University, Ansan 15588, Republic of Korea
[c] Department of Civil Engineering, Hanyang University, Ansan 15588, Republic of Korea

ARTICLE INFO

Keywords:
Building information modeling (BIM)
Cooperative patent classification (CPC)
Patent analysis
Social network analysis (SNA)

ABSTRACT

Building information modeling (BIM) creates new business value and innovation in the construction engineering industry. This study identifies BIM technology structure and the characteristics of knowledge flow through a patent citation network. A technology patent database was collected with backward and forward citation patents currently registered in the US Patents and Trademark Office and identified by the Cooperative Patent Classification. Technology citation networks interact and follow the power law distribution. The analysis of individual patent characteristics and intermediaries revealed digital data processing and telemetry systems as core technology fields stimulating innovative BIM technology. This study overcomes the limitations of previous methods of technology forecasting that have depended on expert opinion or peer review by performing a patent big data analysis to ascertain the domain of the BIM industry. This approach could soon be applied to R&D strategy planning and competition in business and the development of cutting-edge technology.

1. Introduction

Building information modeling (BIM) is particularly salient for influential innovations that have emerged in the construction industry during the past decade [1]. BIM is defined as the digital representation of physical and functional facility characteristics. BIM serves as a shared knowledge resource for facility information and forms a reliable basis for decisions during a facility's life cycle from inception and beyond [2,3]. BIM is an integrated concept that manages life cycle management [2], and its characteristics render it an appropriate tool for application in the construction industry. Due to the potential but limited use of BIM in practice [4,5], increasing research has been devoted to theoretically or empirically investigating BIM-related issues in the past decade [2–4].

Analysis of patent information provides insights on the status of industries and technology flow to determine strategic direction to maximize research and development (R&D) performance. As the performance of R&D is applied for patent and carry over, patented technology commercialization, competitive patent creation, and the identification of patent strategy are becoming increasingly important. According to Market Research on Marketing [6] conducted by the Global Industries Corporation across the United States, 70% of executives said patents are needed for innovation, and 87% of top managers said they respect patent rights. As the importance of patents increases, the methodology using data mining and social network analysis (SNA) allows to build networks with information relationships to identify the structure of the industry. Although the SNA on the construction field [7–9] has been recommended, it is limited to understand the relationship between technologies.

Therefore, this paper will effectively ascertain the structure of the construction BIM industry and BIM technology flow through the patent citation network analysis. Because the US is currently the largest manufacturer and consumer of BIM technology [10], we analyzed the domain of BIM through the citation information of BIM patents registered with the United State Patents and Trademark Office (USPTO). Using the Cooperative Patent Classification (CPC), which complies with the current state of technology to effectively search and manage a significant amount of patent information [11], we built a technology field citation network to understand BIM technology structure and the characteristics of technological knowledge flow.

2. Literature review

2.1. Traditional methods and current trends in patent analysis

A patent is composed of the content of technical embodiments,

* Corresponding author.
E-mail address: yoonsunlee@hanyang.ac.kr (Y.-S. Lee).

https://doi.org/10.1016/j.autcon.2017.12.023
Received 14 February 2017; Received in revised form 28 September 2017; Accepted 7 December 2017
Available online 02 January 2018
0926-5805/

technology classification codes, citation information, and owner information. Technology change trends, technology levels, and commercial values can be understood through the analysis of the component factors in the patent [12]. Thus, patent makes it possible not only to conduct quantitative analysis by providing standardized information on the technologies, but also to provide important information to relevant persons in charge of R&D, technology policies, or strategies [13]. Therefore, patent analysis is suitable for understanding a technological flow and development direction in various industries, exploring new research fields and the creation of new technologies [14].

The previous patent analysis studies are covered in construction field using the bibliometric analyses of patents and publications in energy technology [15], the fuzzy inference system [16] and theory of solving inventive problem (TRIZ) [17]. The patent analysis also includes regression analysis and correlation analysis using strategic variables and evaluates the impact of patents by applying patent statistics to economic models or theories. However, there are limitations to this method because it provides only partial information on technological knowledge diffusion. To examine the technological knowledge diffusion process from a comprehensive perspective, more studies have addressed patent citation relations in terms of network theory [18]. Within patent literature, patent citation information have recently been used to analyze a technological evolution such as technology flow, diffusion, and fusion [19,20] as well as forecasting new technologies [21–23]. Co-classification analysis and citation analysis of technology codes are often developed to reflect linkage relationships between different technological knowledge domains [15]. Researchers also understand the relationship between industries and technology information in industries and discover key technology advancements by forming a patent network based on citations [24].

2.2. SNA measures

SNA, a quantitative technique derived from graph theory, identifies how actors influence other actors and the implications of any relationship within a network [25,26]. Various SNA use the patent citation information to identify technology knowledge flow in organic photovoltaic cells [27], organic solar cells [18], and dye-sensitized solar cells [28]. Sternitzke et al. [29] identifies that, through citation networks, the position of applicants within citation networks is useful in explaining applicant behavior in the marketplace such as cooperation or patent infringement trials. Wang et al. [30] examines whether patent quality is predicted by analyzing information embedded in a patent citation network. Yoon and Park [13] propose a network-based patent citation analysis and show the overall relationship among patents of wavelength division multiplexing as a visual network. The analysis assists users in determining the relative importance of individual patents, and it facilitates the analysis of up-to-date trends in high technologies and the identification of promising avenues for new product development.

SNA methods are likewise evolving, and network topological analysis and node centrality analysis are commonly used [27]. Network topological analysis explains topological structure using characteristics and provides a holistic perspective of knowledge flow particularly in the scientific literature and patent citation [31,32]. However, network topological analysis cannot provide quantitative information on the importance and value of individual nodes [27], so centrality analysis is required to assess the value of each node and the measurement of the structural location. The actual value of the center, that is, the extent to which a node is located at the center of the entire network, has been proved [27,33]. This study analyzes each node using the three types of centrality; degree centrality, closeness centrality, and eigenvector centrality. Additionally, we investigate which technology field acts as an intermediary or bridge in the network using the brokerage analysis. Brokerage analysis recognizes every triad and the role of each node in that triad based on the partition vector [34].

3. Materials and methods

The purpose of this paper is to confirm the structure and characteristics of the technological knowledge flows by comparing and analyzing the overall data and data for the last three years. We first collect BIM patent data including citation information. Second, we investigate the overall structure of the patent citation network through the network visualization and topological analysis. It can grasp the general structure of the network using the various SNA measures shown in Table 2. Then, we identify the value and importance of an individual node by centrality analysis. Finally, we perform the broker analysis for the intermediary role or the bridge between the technology fields. This study uses Netminer [35], a SNA program, to analyze a citation network based on patent citations.

3.1. Data acquisition

Patent citation analysis is required to understand technology flow and forecast the direction of technology in any industry. Although it is particularly important, there is still no research that supports construction industry. Thus, in this study, we adopted the BIM patents which are promising issue these days to experimentally verify the technology flow for promising technology forecasts. The research data is composed of the number of patents currently registered in the USPTO through Worldwide Intellectual Property Service (WIPS) [36] which provides patent search service. Some steps were applied to collect and finalize the data. First, the BIM is an integrated concept, therefore, we performed the keyword expansion and searched the title, abstract, and exemplary claim. In the second step, we refined the unnecessary data using the CPC. The CPC system is divided into nine sections, A-H and Y that, in turn, are sub-divided into classes, sub-classes, groups, and subgroups [37]. Therefore, in this study, we excepted irrelevant CPCs such as “A61: Medical or veterinary science; hygiene,” “B29: Working of plastics, substance in a plastic state” and “G01N: Investigating or analyzing materials by determining their chemical or physical properties.” Through this process, finally, this research was conducted in August 2016, and 113 patent data were extracted from 1998 to 2016.

Fig. 1 shows the status of registered BIM patents by applicant origin. By the early 2000s, less than five patents were registered each year, from the late 2000s, 84% of the total data were registered.

The citation information related to 113 patents was extracted to construct a BIM patent citation network with 3120 backward citations and 1803 forward citations. Fig. 2 shows the top 6 technology fields, which occupy 85% (4272 of 5036) of the whole data. The main CPCs on BIM technology fields from the collected patent data are explained as indicated in Table 1.

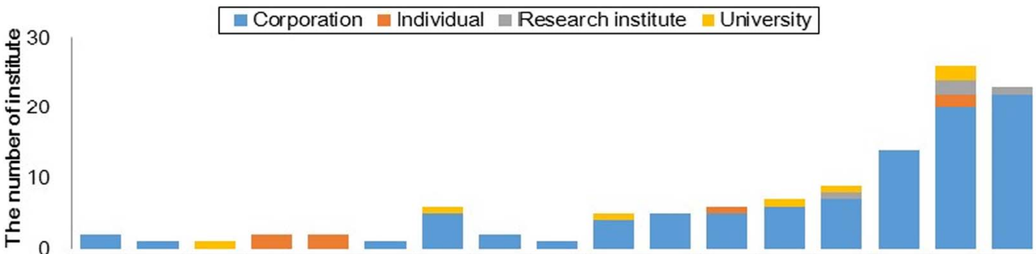
The distribution trends of these key CPCs are shown in Fig. 3. “G06” shows the most critical technology field, which has been cutting-edge BIM technology during the last 10 years; “H04”: represents the next fastest growing proportion. Moreover, continuous expansion of new technology fields shows that BIM technology has been converging into various state-of-the art technologies.

3.2. Network analysis measures

3.2.1. Network visualization and topological analysis

A patent citation network is a directed network and can be constructed in various forms by analytical units [27]. This selected technology field for the analytical unit establishes the patent technology field citation network. In the citation network, a node is a subgroup of the CPC as a technology field, and a link is a citation relationship (citing and cited) among patents.

We first establish the patent technology field citation network, and visualize a network map to assess a general overall structure of the network. Then, we analyze the structural and topological characteristics of the network by using various indices in the Table 2. These



	1998	1999	2000	2001	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
G06	2	1	1		2	1	5	1	1	4	3	6	6	6	9	20	19
H04								1					1	1	1	2	2
G05										1	1			1	2	2	
G01				1			1							1	2		1
Other				1							1					2	1
Total	2	1	1	2	2	1	6	2	1	5	5	6	7	9	14	26	23

Fig. 1. The status of registered BIM patents.

results are compared with those from a random network of the same size [27]. Additionally, a citation network has the characteristics of a scale-free network with a power law degree distribution [32]. Particularly, Choe et al., Hung and Wang [27,38] find a patent citation network roughly follows a power law, where a small number of nodes occupy a large number of links by preferential attachment.

3.2.2. Critical node analysis

Critical node analysis ascertains the value of each node using node centrality, which is an actor-related variable at the individual level under the whole network [39]. Patent citation has the direction of a link, that is, citing or cited. For the citing relationship, a node that cites many other nodes is regarded as in-degree. For the cited relationship, a node that is cited by many other nodes is regarded as out-degree [27]. Based on these node degrees, we use a hubs and authorities analysis [32]. We analyze hub and authority using two types of centrality. First, according to Table 2, degree centrality is limited to the local range in which the node has a direct link, which represents local centrality. On the other hand, the closeness centrality is considered both the direct and indirect links; thus, it has the meaning global centrality. We consider that in- degree centrality is a local hub and in-closeness centrality is a global hub. Therefore, hub analysis can interpret the patent that cites other patent and authority analysis can consider the patent which is cited by other patents.

In addition, the influence node calculated by computing the principal eigenvector that has the biggest eigenvalue among every eigenvector is analyzed using eigenvector centrality. Eigenvector centrality is a more sophisticated version of the degree centrality [44,45]. Whereas degree centrality gives a simple count of the number of connections a node has, eigenvector centrality does not assume that all connections are equal and is the centrality measure that reflects indirect influence. The key idea of this centrality measure is that the prominence of a node is understood to be proportional to the combined prominence of its neighbors [43]. For example, even if a patent cites just one other patent, which subsequently cites many other patents, the first patent in that domain can be considered highly influential. With this centrality, we can consider a technology field's influence based on which patents are being cited for technological development.

3.2.3. Brokerage analysis

Brokerage analysis analyzes every triad and the role of each node in that triad given a partition vector. A number of papers have been published that make use of Gould and Fernandez's methodology [34] and find that actors in brokerage positions may enjoy strategic advantages in a number of activities. These range from the political to the entrepreneurial to academic inventors designated as patent creators. Therefore, we can identify the characteristics of technology fields that act as coordinators, gatekeepers, representatives, consultants, or liaisons, as shown in Table 2, to understand the domain and route of the exchanges between them and other technology fields [46].

4. Results

4.1. Network visualization and topological analysis

To analyze the core technology field, we built a core network composed of the top 100 citation relations according to link weight [32]. Consequently, Fig. 4 shows the core network visualization composed of 80 main node sets and top 100 citation relations.

In this figure, “G06F-0017/30395” constructs the main cluster. “G06F-0017/30395” is iterative query focused on three-dimensional (3D) presentation for interactive geographic information system (GIS). Next, “H04Q-0009/00” and “G06F-0017/30545” form a local cluster. “H04Q-0009/00” is space management method through occupancy

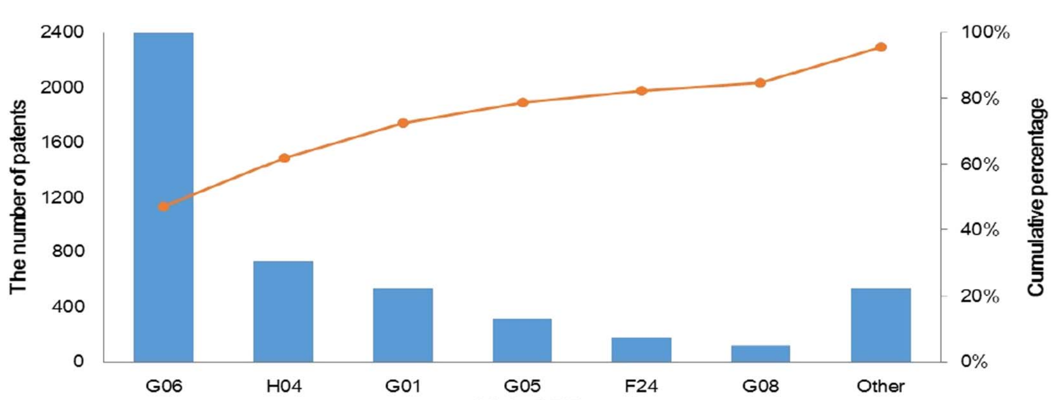


Fig. 2. Key technology fields of the BIM technology fields.

Table 1
Main CPC descriptions related with BIM technology.

Level	Code	Descriptions
Section	F	Mechanical engineering
Class	F24	Ventilating, use of air currents for screening
Sub-class	F24F	Air-conditioning, ventilation, use of air currents for screening
Main-group	F24F11/00	Control or safety systems or apparatus
Sub-group	F24F11/006	Control systems or circuits characterized by type of control, internal processing
	F24F11/0012	Control or safety systems for the Air temperature
	F24F11/0086	Control systems or circuits characterized by other control features
Section	G	Physics
Class	G01	Measuring, testing
Sub-class	G01C	Measuring distances, levels, surveying, photogrammetry or videogrammetry
Main-group	G01C15/00	Surveying instruments or accessories
Sub-group	G01C21/20	Instruments for performing navigational calculations
	G01C21/206	Measuring levels specially adapted for indoor navigation
Sub-class	G01K	Measuring temperature, Measuring quantity of heat
Main-group	G01K13/00	Measuring quantity of heat for adaptations of thermometers for specific purposes
Class	G05	Controlling, regulating
Sub-class	G05B	Functional elements of such systems, monitoring arrangements for such systems
Sub-group	G05B15/02	Systems controlled by electric computer
Sub-class	G05D	Systems for controlling or regulating non-electric variables
Sub-group	G05D23/1904	Control of temperature variable in time
Class	G06	Computing, calculating, counting
Sub-class	G06F	Electrical digital data processing
Sub-group	G06F8/35	Software engineering for model driven
	G06F17/30395	Computing methods for iterative querying based on the preceding query results
	G06F17/30545	Digital data processing about distributed queries
	G06F17/30693	Digital computing methods for reuse of stored results of previous queries
	G06F17/5004	Data processing methods architectural design
Sub-class	G06K	Recognition of data, presentation of data, record carriers, handling record carriers
Sub-group	G06K9/00597	Methods for recognizing printed or written characters, recognizing eyes
Main-group	G06K13/00	Recognition of data conveying record carriers from one station to another
Sub-class	G06Q	Data processing systems, specially adapted for administrative, managerial, supervisory or forecasting purposes
Sub-group	G06Q10/06	Administration of resources, workflows, human or project management
	G06Q10/087	Administration system for inventory or stock management
Sub-class	G06T	Image data processing or generation
Sub-group	G06T7/0004	Image analysis of industrial image inspection
	G06T11/60	2D image generation editing figures and text
Main-group	G06T17/00	3D modeling, e.g. data description of 3D objects
Sub-group	G06T17/05	3D modeling for geographic models
	G06T19/006	Manipulating 3D models or images for computer graphics about mixed reality
Section	H	Electricity
Class	H04	Electric communication technique
Sub-class	H04Q	Selecting (switches, relays, selectors)
Main-group	H04Q9/00	Arrangements in telemetry systems for selectively calling a substation from a main station, in which substation is selected for applying a control signal thereto
Sub-class	H04W	Wireless communication networks
Sub-group	H04W16/20	Network planning for indoor coverage or short ranges network deployment

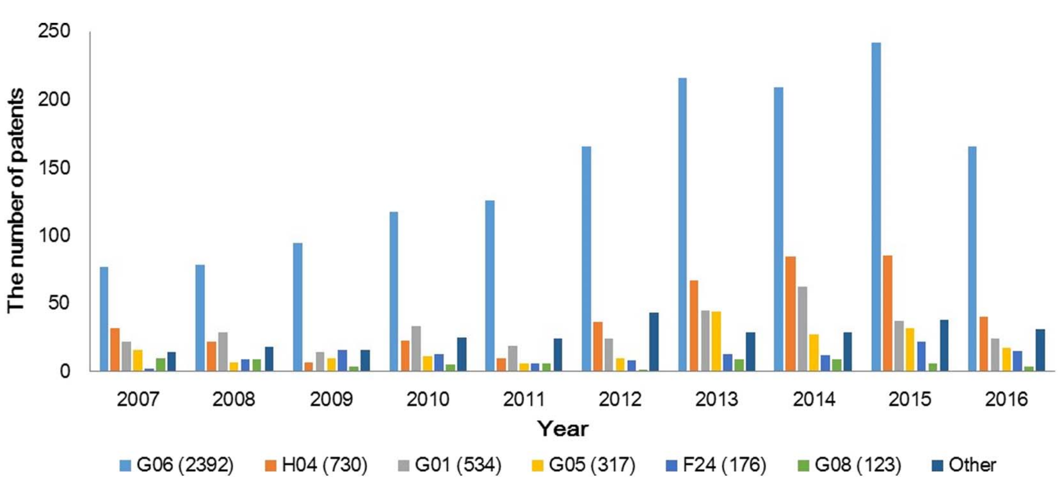


Fig. 3. The distribution of key CPCs.

Table 2
Research measures.

Analysis	Measure	Descriptions	Application
Topological analysis	Power law distribution	The method combined maximum-likelihood fitting methods with goodness-of-fit tests based on the Kolmogorov-Smirnov (KS) statistic and likelihood ratios. If p-value of the goodness-of-fit test is large (close to 1), then it fits the power-law distribution	[8,27,32,38]
	Link	The number of lines	[8,12,26,29,30,40]
	Density	The number of lines in a network expressed as a proportion of the maximum number of possible number of lines	[8,12,26,29]
	Average degree	The average number of links that a node has to other nodes	[27,32]
	Components	The maximal sub-graph in which each pair of node is connected by a semi-path	[41]
	Inclusiveness	The number of connected nodes expressed as a proportion of the total number of nodes. Connected nodes mean the nodes except isolates	[25]
	Clustering coefficient	A measure of the degree to which nodes in a graph tend to cluster together	[8,12]
	Mean distance	Geodesic, mean, diameter	[8,29]
	Diameter	The length of the largest geodesic path in the network	[27]
	Degree centralization	Calculated by finding the total sum of values gained by subtracting the degree centrality of each node from the maximum degree centrality within the network, followed by dividing the total sum by the theoretically possible maximum of degree centrality	[12,32]
Critical node analysis	Degree centrality	Degree centralization index is a measure of variability of individual centrality scores. The larger degree centralization index is the more centralized that network is	[8,12,13,40,42,43]
	Closeness centrality	Closeness centrality expands the definition of degree centrality by measuring how close a node is to all the other nodes. So, it can consider both direct and indirect connections	[8,29,30,43]
	Eigenvector centrality	Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix defining the network. The key idea of this centrality measure is that the prominence of a node is understood to be proportional to the combined prominence of its neighbors	[8,34,43–45]
Brokerage analysis	Coordinator	If node 'a' receives a link from node 'b' in the same partition, and send a link to node 'c' in the same partition, then add 1 Coordinator score to node 'a'	[8,34]
	Gatekeeper	If node 'a' receives a link from node 'b' in different partition, and send a link to node 'c' in the same partition, then add 1 Gatekeeper score to node 'a'	[8,34]
	Representative	If node 'a' receives a link from node 'b' in the same partition, and send a link to node 'c' in different partition, then add 1 Representative score to node 'a'	[8,34]
	Itinerant (Consultant)	If node 'a' receives a link from node 'b' in different partition, and send a link to node 'c' in that partition (same as 'b'), then add 1 Itinerant score to node 'a'	[8,34]
	Liaison	If node 'a' receives a link from node 'b' in different partition, and send a link to node 'c' in another different partition (different from 'b'), then add 1 Liaison score to node 'a'	[8,34]

sensors, and "G06F-0017/30545" is an ontological model of a building automation system (BAS). "G06F-0010/06," "G01C-0021/20, "G05B-0015/02," and "G06T-0017/05" are closely related among the clusters.

Table 3 shows the network topological measures of the technology field citation. The degree centralization index (in-degree: 16.49%, out-degree: 7.52%) is significantly higher than that of a random network of the same size. Particularly, the in-degree centralization index is higher than the out-degree centralization index. This indicates that the degree to which the network is concentrated in a center is higher in a citing network than in a cited network.

The p-value of both degree distributions is greater than 0.1 (in-degree: 0.45, out-degree: 0.13). These values indicate that the network is a scale-free network that follows the power law distribution where the number of links is unevenly distributed. This implies that some of the specific nodes affect the characteristic of the overall network. The network has a small average degree 1.138 with lower density 0.014. Both the average degree and density are lower than random network. While the network has seven components, the inclusiveness is almost 1(0.98). This implies that most technologies in the network are directly or indirectly connected and some nodes influence the whole network as shown in Fig. 4. The technology field citation network has a significantly larger clustering coefficient (0.18) than that of a random network of the same size. This implies that the nodes are more closely connected in a technology field citation network than in a random network. This network has a small mean distance (2.395) and a small diameter (5). These values are significantly smaller than a random network of the same size. This indicates that the knowledge transfer process in the technology citation network is faster than in the random network.

4.2. Critical node analysis

Table 4 shows the results of the critical node analysis for the technology field citation network.

First, for the hub, the top-ranked CPC is similarly shown in both global and local hubs. "G06F-0017/30395" ranked first in both in-degree centrality and in-closeness centrality, which implies that the patents belonging to this CPC cluster most frequently cite other patents in different technology fields either directly or indirectly. Although "G06F-0017/30395" is the most important local and global hub in the technology citation network, there are less citations in the recent three years. Another major hub is "H04Q-0009/00." This hub ranked highest in the recent three years at both local and global hub. "G06T-0017/05" has higher in-closeness centrality than in-degree centrality, which implies that this hub is more related indirectly with other technology fields; the hub concerns the visualization and combination of image information to generate 3D models for the geographic model. "G06F0017/5004," data processing methods for architectural design, is also one of the major hubs. Additionally, there is remarkable recent growth in "G06F-0017/30545", "G06F-0017/30693" and "G0T-0019/006".

In the Authority, "G06Q-0010/06" is ranked highly regardless of time in both the local and global authority, which implies that the patents belonging to "G06Q-0010/06" are cited most frequently either directly or indirectly. In the other words, this is the most important local and global authority in the BIM technology field citation network. "G06F-0017/30395" is ranked first in local authority; however, in the recent three years, it is ranked relatively low. This implies that research in the "G06F-0017/30395"-based technology field is needed. "G06T-0017/05" and "G05B-0015/02" are also major authorities. The technologies containing "G05B-0015/02" reflect building management systems (BMS).

The eigenvector centrality results shown in Table 5 shows a node set with influential nodes.

"G06F-0017/5004" has the highest eigenvector centrality. This implies that "G06F-0017/5004" is related to core technology to greater

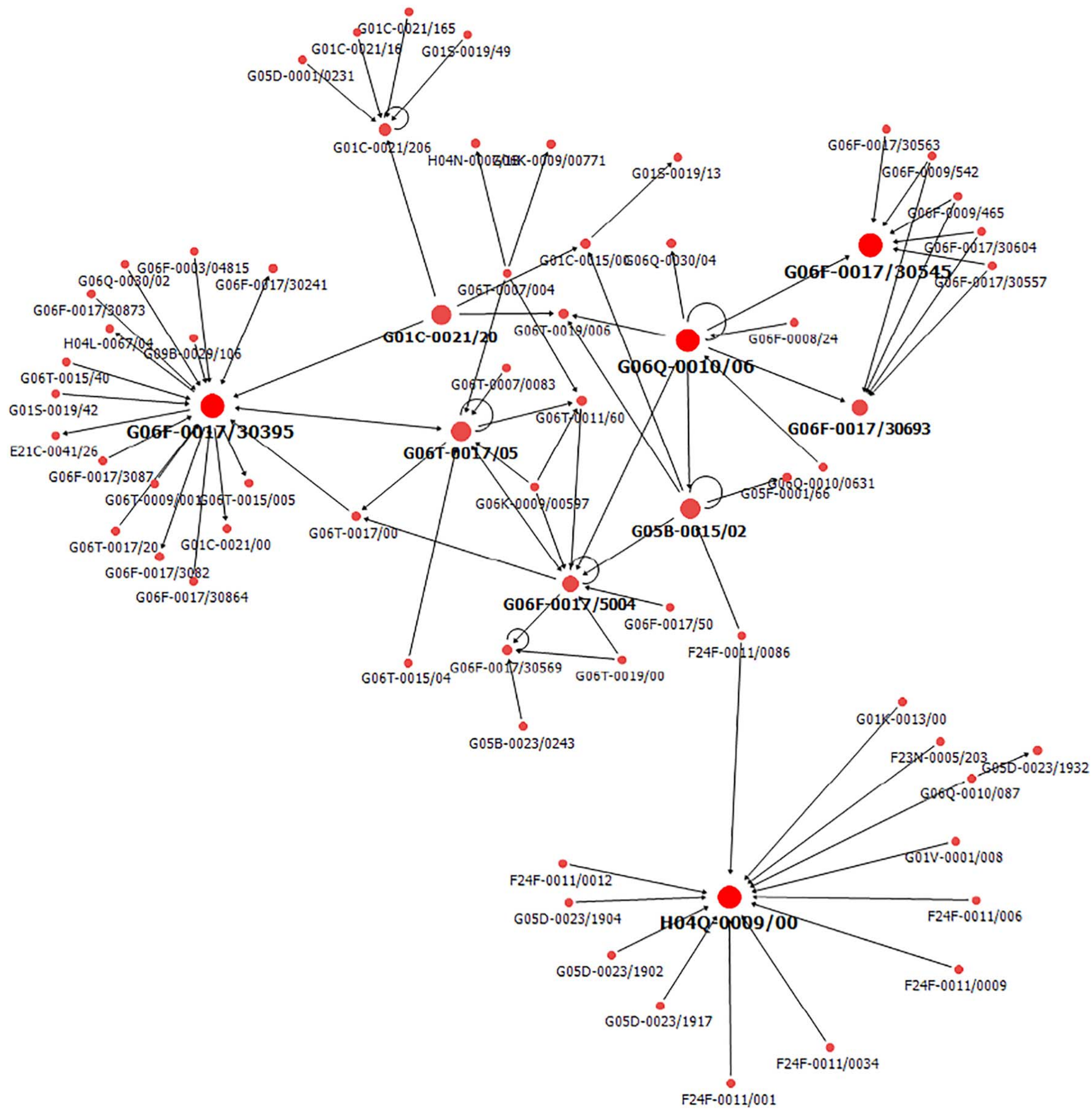


Fig. 4. The core technology fields network.

Table 3
Topological measures of the technology field citation network.

	Links	Density	Average degree	Components	Inclusive-ness	Clustering coefficient	Mean distance	Diameter	Degree centralization	Power-law distribution: KS statistic	Power-law distribution: p-Value
Technology citation network	100	0.014	1.138	7	0.975	0.183	2.395	5	In) 16.488% Out) 7.515%	In) 0.19 Out) 0.286	In) 0.45 Out) 0.13
Random network	100	0.016	1.25	7	0.938	0.041	4.127	12	In) 4.807% Out) 3.525%	–	–

extent than the number of relationships with other nodes. G05 class, which includes control and regulating systems, and the G06 class, which includes computing and calculating, are ranked in eigenvector centrality, the same as "G06Q-0010/06," "G05B-0015/02," "G06T-0017/05," and "G06K-0009/00597." Diverse technology fields have emerged in the last three years and are placed near a high-powered node in the network. Because H section (electricity system) and F section (mechanical engineering) have been revealed in the latest three years, and they connected with core nodes, these sections exercise considerable influence on the network. Table 5 shows a substantial difference between the overall data and the results for the recent three years, which reveals that BIM is a wide-ranging, continuously expanding integrated technology.

4.3. Brokerage analysis

Brokerage analysis analyzes the five types of broker role for each node. We designate a subclass of CPC as partition value. Fig. 5 and Table 6 show the results of the broker analysis.

"G06F-0017/30395" is the main bridge node connecting the

Table 4
Top 5 hubs and authorities.

Hubs	Local hub			Global hub		
	Rank	CPC	In-degree centrality	Rank	CPC	In-closeness centrality
	1	G06F-0017/30395	0.177	1	G06F-0017/30395	0.181
	2	H04Q-0009/00	0.165	2	G06T-0017/05	0.181
	3	G06F-0017/5004	0.089	3	H04Q-0009/00	0.165
	4	G06T-0017/05	0.076	4	G06F-0017/5004	0.154
	5	G06F-0017/30545	0.076	5	G06T-0017/00	0.138
	The recent 3 years					
	1	H04Q-0009/00	0.197	1	H04Q-0009/00	0.198
	2	G06F-0017/30545	0.183	2	G06F-0017/30545	0.194
	3	G06F-0017/30693	0.169	3	G06F-0017/30693	0.183
	4	G01C-0021/206	0.127	4	G06T-0019/006	0.165
	5	G06T-0019/006	0.085	5	G01C-0015/00	0.156

Authorities	Local authority			Global authority		
	Rank	CPC	Out-degree centrality	Rank	CPC	Out-closeness centrality
	1	G06F-0017/30395	0.089	1	G06Q-0010/06	0.103
	2	G06Q-0010/06	0.076	2	G01C-0021/20	0.098
	3	G06T-0017/05	0.051	3	G06F-0017/30395	0.096
	4	G05B-0015/02	0.051	4	G06T-0017/05	0.085
	5	G01C-0021/20	0.051	5	G06T-0007/004	0.080
	The recent 3 years					
	1	G06Q-0010/06	0.099	1	G06Q-0010/06	0.107
	2	G05B-0015/02	0.070	2	G05B-0015/02	0.082
	3	G06T-0017/05	0.056	3	G06T-0017/05	0.070
	4	G01C-0021/20	0.056	4	F24F-0011/0086	0.067
	5	H04Q-0009/00	0.042	5	H04Q-0009/00	0.064

technologies on condition that partition value is “G06F”. First, “G06F-0017/30395” has the highest score in liaison in the technology domain implies that it is highly influential when the “G06F-0017/30395” received links from different partition sends links to another different partition. Fig. 5 shows that the “G06F-0017/30395” cluster acts as a liaison broker among the “G06T,” “G06Q,” “G01C,” and “G05B” clusters. It also scored highly in representative, which implies a ripple effect when the “G06F-0017/30395” receives links from the same partition as G06F and sends links to different partitions.

“G06F-0017/5004,” the same group as “G06F-0017/30395,” is disseminated through “G06F-0017/30395” to “G06T” and “G01C”, as shown in Fig. 5. Only 9% of the total brokerage score corresponds to coordinator. The reason for this low coordinator score is that “G06F” group is highly externally connected. However, “G06F-0017/30395” is an important broker node in comparison with other partition value results.

5. Discussion

This research analyzes a core technology network based on top 100 patent citation relationships to understand the structure and characteristics of the technology flow in the field of BIM and to identify the value of individual nodes. For network analysis, topological analysis, critical node analysis, and brokerage analysis were carried out.

We assessed the overall structure of the network compared with a random network through the topological analysis. The result is that the network is a scale-free network that follows the power law. The topological indices also show that the BIM patent citation network is more compact than a random network. This result suggests that nodes in the citation networks cooperate more closely exchanging information than those in the random network. Thus, the process of knowledge diffusion within the network is slightly more efficient.

In the critical node analysis, the hub and authority analysis were reviewed according to the citation direction based on node degree. The former is related to the technology field absorbing other technologies; the latter is related to the core propagated field.

The hub analysis shows that “G06F-0017/30395,” facilitating communication systems between servers and clients of the interactive GIS, and the virtual tours method of user-defined paths in the context of distributed geospatial visualization have been developed. These patents cite methods such as providing statistically interesting geographical information and apparatus for adaptive hierarchical visibility in a tiled 3D architecture. These patents have high impact within “G06F-0017/30395” on the entire network. In the recent three years, “H04Q-009/00” on electronic communication technologies such as sensors has developed. Additionally, the main patents in these fields, almost submitted by Nest Labs Inc., are updated consistently and are now under the new ownership of Google Inc. This implies that original applicants are acquired by more powerful companies in the BIM technology field during the process of enhancing main technology. Thus, analyzing technology flows using patent citation networks to establish the corporate strategy of construction firms are considered key market drivers.

For authority analysis, the core technology field “G06Q-0010/06” has been highly cited regardless of time. This field represents project management and resource flows, a form of informatization for construction work, and is cited in BIM technology. The individual patents

Table 5
Top 5 eigenvector nodes.

Rank	CPC	Eigenvector centrality	Rank	CPC of the recent 3 years	Eigenvector centrality
1	G06F-0017/5004	0.603	1	H04Q-0009/00	0.719
2	G06Q-0010/06	0.584	2	F24F-0011/006	0.431
3	G05B-0015/02	0.342	3	G05D-0023/1904	0.244
4	G06T-0017/05	0.225	4	F24F-0011/0012	0.216
5	G06K-0009/00597	0.121	5	G01K-0013/00	0.216

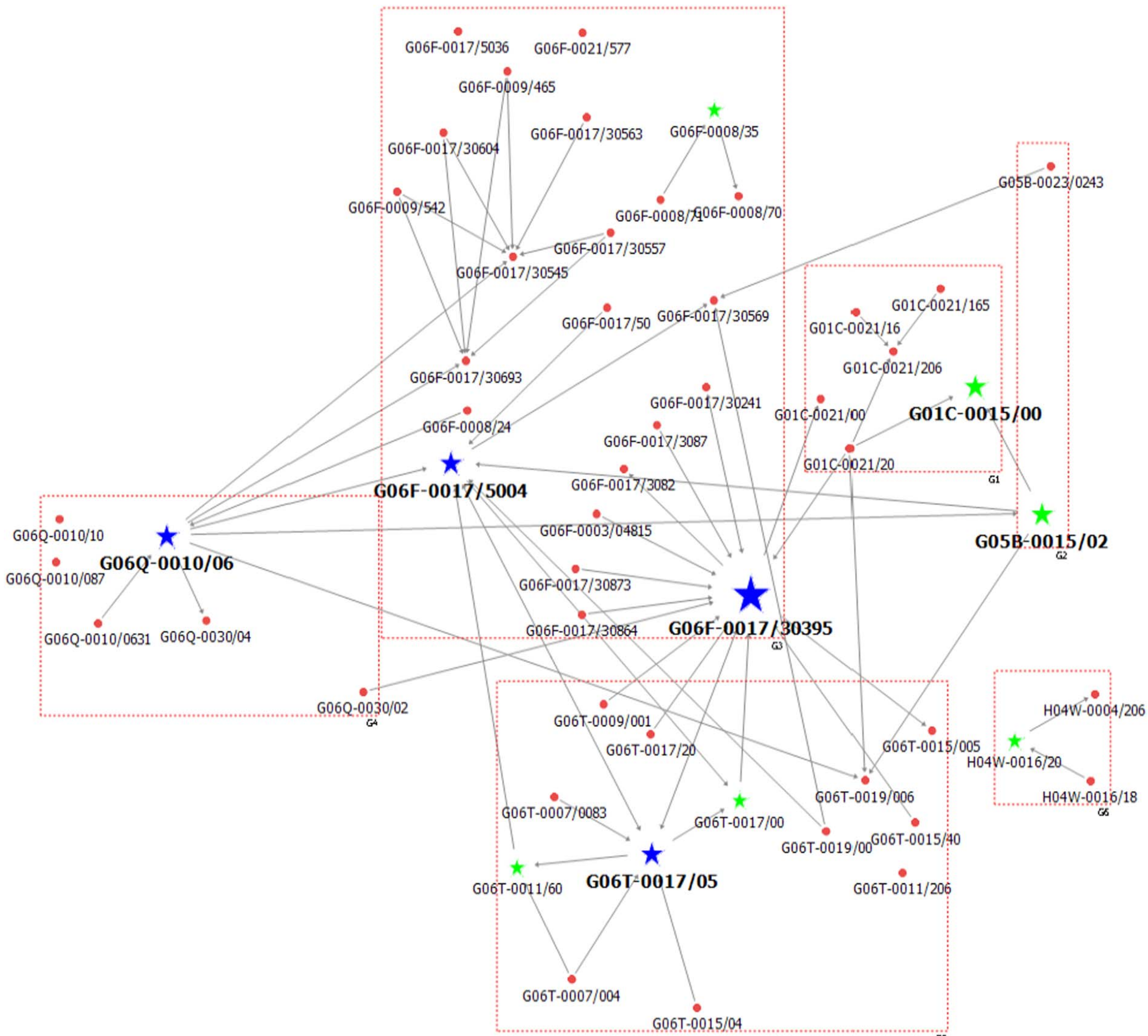


Fig. 5. Brokerage analysis visualization.

Table 6
Top 5 brokers.

Rank	CPC	Partition Value	Coordinator	Gatekeeper	Representative	Itinerant	Liaison	Total
1	G06F-0017/30395	G06F	9	18	25	10	34	96
2	G06T-0017/05	G06T	5	4	6	2	1	18
3	G06F-0017/5004	G06F	1	5	2	4	5	17
4	G06Q-0010/06	G06Q	1	1	5	3	2	12
5	G05B-0015/02	G05B	0	0	0	0	6	6

in these technology fields show a wide range of technologies, a technical delivery framework facilitating the development of complete enterprise service management solutions, and an on-site management system. This system acts as a building support tool that displays a layout creation screen representing a virtual system under construction. In addition to spatial program validation, cost calculation of cash flow modeling and scheduling systems are mainly propagated. Since the cost estimation and scheduling is revealed in an authority, this result implies that the flow from 3D (spatial model with quantity take off), 4D (plus construction scheduling), 5D (plus cost calculation) to multifunctional (project life cycle) BIM are extensively discussed in leading BIM technologies.

"G06F-0017/5004" data management focuses on the architectural design stage, which is ranked first in the eigenvector centrality. This can potentially accelerate the rate at which "G06F-0017/5004" receives information and, in turn, access information on breakthroughs or developments. The main patents in this field focus on energy analysis, library-based generation of the 3D geometric model, existing digital topographic maps of buildings, and a semi-automatic modeling method.

Finally, in the brokerage analysis, "G06F-0017/30395" is the key broker node. This node plays the role of liaison, which implies that all technologies in the "G06F" class move towards other technology class fields for greater convergence. "G06F-0017/30395" is also important as it plays a representative role implying that most of the technology in the "G06F" class, particularly "G06F-0017/5004," the generating energy and 3D geographic model, should almost flow through "G06F-

0017/30395," which is a geospatial iterative query system compared to other technology classes. Overall, this research explains the domain and characteristics of technology flow of current industries analyzing both empirical technology and new technologies emerging through network analysis.

6. Conclusion

Industry 4.0 represents an opportunity to evolve and set a new precedent for possible technology advancements. Thus, big data analytics are receiving remarkable attention and noticeably impacting the civil engineering domain [47]. In this respect, this research performed a patent big data analysis to ascertain the domain of the BIM industry overcoming previous methods of technology forecasting, which depend on expert opinion or peer review [14]. Because patents are based on existing technologies and are valuable assets of technical and economical knowledge [19], patent citation analysis performs an important role in identifying the flow of knowledge and technology. Therefore, this research focused on understanding the structure and direction of BIM technology flow and the role of brokerage in the network considering existing information on CPCs with backward citation and forward citation.

The BIM technology network shows that the knowledge transfer process is faster and the process of knowledge diffusion is slightly more efficient. "G06F-0017/30395," iterative querying system, is the main node in the network. "H04Q-0009/00", telemetry systems, is gradually citing other nodes. "G06Q-0010/06," administration and management data processing systems, is highly regarded in the local and global authority area. We also find that "H04Q-0009/00" had high eigenvector centrality in the last three years while "G06F-0017/5004," digital computing for architectural design, was the highest among the overall data. Finally, the node of the main bridging role is "G06F-0017/30395." This result reveals that an SNA approach is affected by the number of nodes that appear but can identify the node that has an important position in the citation relations. Specifically, although H04L, H04W, G01S have a high frequency ratio in the BIM patents domain, they are not significantly located in the technology field citation network. On the contrary, this study confirmed that H04Q, G06K, G06D, which have a low frequency ratio, occupy an important position comparatively in the network. This means that the BIM technology network proposed by this study can analyze the core source technologies of the BIM technology convergence and understand the flow of knowledge. The results of this study interpreted the latent meaning of the technology relationship in the citation network beyond the number of patent applications. The noteworthy technologies do not correspond with the technologies owned by company with the largest number of patents. These results have the potential to become a basis for the implementation of a supporting tool to establish R&D encouragement strategies and develop cutting-edge technology in the construction industry.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, and Technology (No. NRF-2012R1A2A2A04046958).

References

[1] T.M. Froese, The impact of emerging information technology on project management for construction, Autom. Constr. 19 (2010) 531–538, http://dx.doi.org/10.1016/j.autcon.2009.11.004.
[2] R. Volk, J. Stengel, F. Schultmann, Building Information Modeling (BIM) for existing buildings - literature review and future needs, Autom. Constr. 38 (2014) 109–127, http://dx.doi.org/10.1016/j.autcon.2013.10.023.
[3] M. Yalcinkaya, V. Singh, Patterns and trends in building information modeling (BIM) research: a latent semantic analysis, Autom. Constr. 59 (2015) 68–80, http://dx.doi.org/10.1016/j.autcon.2015.07.012.
[4] D. Cao, H. Li, G. Wang, X. Luo, X. Yang, D. Tan, Dynamics of project-based collaborative networks for BIM implementation: analysis based on stochastic actor-oriented models, J. Manag. Eng. (2015) 1–12, http://dx.doi.org/10.1061/(ASCE)ME.1943-5479.0000503.
[5] D. Cao, G. Wang, H. Li, M. Skitmore, T. Huang, W. Zhang, Practices and effectiveness of building information modelling in construction projects in China, Autom. Constr. 49 ((2015) 113–122, http://dx.doi.org/10.1016/j.autcon.2014.10.014.
[6] I. Ventures, Patent Attitudes in the C-Suite: Market Research Highlights Do American Businesses Value Patents? Patents are not Stifling American Innovation, (2013).
[7] M. Al Hattab, F. Hamzeh, Using social network theory and simulation to compare traditional versus BIM-lean practice for design error management, Autom. Constr. 52 (2015) 59–69, http://dx.doi.org/10.1016/j.autcon.2015.02.014.
[8] Y.-S. Lee, J.-J. Kim, T.S. Lee, Topological competiveness based on social relationships in the Korean construction-management industry, J. Constr. Eng. Manag. (2016) 5016014, , http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001175.
[9] S. Moon, S. Han, P.R. Zekavat, L.E. Bernold, X. Wang, Process-waste reduction in the construction supply chain using proactive information network, Concurr. Eng. (2016), http://dx.doi.org/10.1177/1063293X16667451.
[10] J.C.P. Cheng, Q. Lu, A review of the efforts and roles of the public sector for BIM adoption worldwide, J. Inf. Technol. Constr. 20 (2015) 442–478 http://www.itcon.org/paper/2015/27.
[11] G. Kim, J. Bae, A novel approach to forecast promising technology through patent analysis, Technol. Forecast. Soc. Change. 117 (2017) 228–237, http://dx.doi.org/10.1016/j.techfore.2016.11.023.
[12] J. Choi, Y.S. Hwang, Patent keyword network analysis for improving technology development efficiency, Technol. Forecast. Soc. Change. 83 (2014) 170–182, http://dx.doi.org/10.1016/j.techfore.2013.07.004.
[13] B. Yoon, Y. Park, A text-mining-based patent network: analytical tool for high-technology trend, J. High Technol. Manag. Res. 15 (2004) 37–50, http://dx.doi.org/10.1016/j.hitech.2003.09.003.
[14] J.-H. Han, J.-G. Na, C.-B. Kim, A Study of ICT technology convergence analysis and development direction by using the patent Information: focusing on Kyongbuk area company, J. Intellect. Prop. 10 (2015) 203–238 https://www.kiip.re.kr/en/journal/view.do?bd_gb=ejor&bd_cd=1&bd_item=0&po_d_gb=1&currentPage=5&po_a_no=221.
[15] J.E. Altwies, G.F. Nemet, Innovation in the U.S. building sector: an assessment of patent citations in building energy control technology, Energ Policy 52 (2013) 819–831, http://dx.doi.org/10.1016/j.enpol.2012.10.050.
[16] W.D. Yu, S.S. Lo, Patent analysis-based fuzzy inference system for technological strategy planning, Autom. Constr. 18 (2009) 770–776, http://dx.doi.org/10.1016/j.autcon.2009.03.003.
[17] I.A. Renev, L.S. Chechurin, Application of TRIZ in building industry: study of current situation, Procedia CIRP. 39 (2016) 209–215, http://dx.doi.org/10.1016/j.procir.2016.01.190.
[18] H. Choe, D.H. Lee, H.D. Kim, I.W. Seo, Structural properties and inter-organizational knowledge flows of patent citation network: the case of organic solar cells, Renew. Sust. Energ. Rev. 55 (2016) 361–370, http://dx.doi.org/10.1016/j.rser.2015.10.150.
[19] Z. Liu, D. Zhu, Web mining based patent analysis and citation visualization, Web Min. Web Based Appl. (2009) 19–23, http://dx.doi.org/10.1109/WMWA.2009.33.
[20] M. Karvonen, T. Kässi, Patent citations as a tool for analysing the early stages of convergence, Technol. Forecast. Soc. Change. 80 (2013) 1094–1107, http://dx.doi.org/10.1016/j.techfore.2012.05.006.
[21] M.N. Kyebambe, G. Cheng, Y. Huang, C. He, Z. Zhang, Forecasting emerging technologies: a supervised learning approach through patent analysis, Technol. Forecast. Soc. Change. (2017) 0–1, http://dx.doi.org/10.1016/j.techfore.2017.08.002.
[22] A. Breitzman, P. Thomas, The emerging clusters model: a tool for identifying emerging technologies across multiple patent systems, Res. Policy 44 (2015) 195–205, http://dx.doi.org/10.1016/j.respol.2014.06.006.
[23] Direct validation of citation counts as indicators of industrially important patents, Res. Policy 20 (1991) 251–259, http://dx.doi.org/10.1016/0048-7333(91)90055-U.
[24] T. Montecchi, D. Russo, Y. Liu, Searching in cooperative patent classification: comparison between keyword and concept-based search, Adv. Eng. Inform. 27 (2013) 335–345, http://dx.doi.org/10.1016/j.aei.2013.02.002.
[25] J. Scott, Social network analysis: a handbook, SAGE Publ. 208 (2000), http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Social+Network+Analysis+A+Handbook#6.
[26] C.S. Weng, W.Y. Chen, H.Y. Hsu, S.H. Chien, To study the technological network by structural equivalence, J. High Technol. Manag. Res. 21 (2010) 52–63, http://dx.doi.org/10.1016/j.hitech.2010.02.007.
[27] H. Choe, D.H. Lee, I.W. Seo, H.D. Kim, Patent citation network analysis for the domain of organic photovoltaic cells: country, institution, and technology field, Renew. Sust. Energ. Rev. 26 (2013) 492–505, http://dx.doi.org/10.1016/j.rser.2013.05.037.
[28] Y. Huang, F. Zhu, Y. Guo, A. Porter, D. Zhu, Identifying technology evolution pathways based on patent citation network and tech mining – illustrated for dye - sensitized solar cells, Futur. Technol. Anal. Brussels (2014) 27–28, http://dx.doi.org/10.13140/RG.2.1.3168.1125.
[29] C. Sternitzke, A. Bartkowski, R. Schramm, Visualizing patent statistics by means of social network analysis tools, World Patent Inf. 30 (2008) 115–131, http://dx.doi.

org/10.1016/j.wpi.2007.08.003.
[30] J.C. Wang, C. hsin Chiang, S.W. Lin, Network structure of innovation: can brokerage or closure predict patent quality? Scientometrics 84 (2010) 735–748, http://dx.doi.org/10.1007/s11192-010-0211-y.
[31] P. Erdös, A. Rényi, On random graphs, Publ. Math. 6 (1959) 290–297, http://dx.doi.org/10.2307/1999405.
[32] X. Li, H. Chen, Z. Huang, M.C. Roco, Patent citation network in nanotechnology (1976–2004), J. Nanopart. Res. 9 (2007) 337–352, http://dx.doi.org/10.1007/s11051-006-9194-2.
[33] R.S. Burt, Structural holes: the social structure of competition, Harvard Univ. Press. Cambridge Massachussetts, 1992, pp. 38–40, http://dx.doi.org/10.1177/0265407512465997.
[34] R.V. Gould, R.M. Fernandez, Structures of mediation: a formal approach to brokerage in transaction networks, Sociol. Methodol. 19 (1989) 89–126, http://dx.doi.org/10.2307/270949.
[35] N. Cyram, Cyram Netminer 4, http://www.netminer.com/product/overview.do, (2013).
[36] Hyung Chil Lee, WIPS: Worldwide Intellectual Property Service, http://www.wipscorp.com/main.wips, (1999).
[37] R. Kapoor, M. Karvonen, S. Ranaei, T. Kässi, Patent portfolios of European wind industry: new insights using citation categories, World Patent Inf. 41 (2015) 4–10, http://dx.doi.org/10.1016/j.wpi.2015.02.002.
[38] S.W. Hung, A.P. Wang, Examining the small world phenomenon in the patent citation network: a case study of the radio frequency identification (RFID) network, Scientometrics 82 (2010) 121–134, http://dx.doi.org/10.1007/s11192-009-0032-z.
[39] M.S. Mizruchi, C. Marquis, Egocentric, sociocentric, or dyadic? Identifying the appropriate level of analysis in the study of organizational networks, Soc. Networks 28 (2006) 187–208, http://dx.doi.org/10.1016/j.socnet.2005.06.002.
[40] S.S. Ko, N. Ko, D. Kim, H. Park, J. Yoon, Analyzing technology impact networks for R&D planning using patents: combined application of network approaches, Scientometrics (2014) 917–936, http://dx.doi.org/10.1007/s11192-014-1343-2.
[41] H. Kwak, C. Lee, H. Park, S. Moon, What is Twitter, a social network or a news media? Int. World Wide Web Conf. Comm. (2010) 1–10, http://dx.doi.org/10.1145/1772690.1772751.
[42] B. Gress, Properties of the USPTO patent citation network: 1963–2002, World Patent Inf. 32 (2010) 3–21, http://dx.doi.org/10.1016/j.wpi.2009.05.005.
[43] H. Kim, J. Song, Social network analysis of patent infringement lawsuits, Technol. Forecast. Soc. Change. 80 (2013) 944–955, http://dx.doi.org/10.1016/j.techfore.2012.10.014.
[44] M.M. Fischer, D.A. Griffith, Modeling spatial autocorrelation in spatial interaction data: an application to patent citation data in the european union, J. Reg. Sci. 48 (2008) 969–989, http://dx.doi.org/10.1111/j.1467-9787.2008.00572.x.
[45] E. Costenbader, T.W. Valente, The stability of centrality measures when networks are sampled, Soc. Networks 25 (2003) 283–307, http://dx.doi.org/10.1016/S0378-8733(03)00012-1.
[46] F. Lissoni, Academic inventors as brokers, Res. Policy 39 (2010) 843–857, http://dx.doi.org/10.1016/j.respol.2010.04.005.
[47] A.H. Alavi, A.H. Gandomi, Big data in civil engineering, Autom. Constr. (2017), http://dx.doi.org/10.1016/j.autcon.2016.12.008.