The core/periphery positions in technological network: The affiliation network analysis

Calvin S. Weng1*, Hsien-Che Lai 2
1 Takming University of Science and Technology, Taiwan, R.O.C.
E-mail: calvinweng@takming.edu.tw (corresponding author)
2 National University of Tainan, Taiwan, R.O.C
E-mail: hclai@mail.nutn.edu.tw

Abstract

This paper is to explore the interactions between the focus of R&D and technological areas, and to examine how the technological network shapes the technological position in the technological network. The social network analysis is utilized to investigate the inquiry of how R&D activities interact. The empirical data was from the United States Patent and Trademark Office (USPTO). The affiliation network approach was conducted to study the interaction between patents (R&D activities) and technological classifications (technological areas). From the result of analysis, we explored the differentiated positions of the core and periphery and further identified some key patents as well. Technological implications are discussed in the conclusion.

Key words:
Affiliation network, core/periphery, technological position, social network analysis, patent analysis

1 Introduction

Technological development is seen as systemic activities involving learning processes and many innovations. A considerable body of literature and evidences points out that building out technological position of an organization is a complex endeavor and lasting task with respect to the variables such as absorptive ability, technological competency and knowledge exploitation of firm [1-3]. And these variables are strongly influenced by the social or institutional context in which firms operate [4-8]. As stated by Plolny, et al. [7]: “if we define a technological tie as a link between an antecedent and consequent invention, then an organization’s niche is its position in technology space, as defined by the pattern of technological ties involving its inventions”. Past studies often identified the technological position of firms by the means of statistical method, such as multi-dimensional scaling (MDS) or the clustering technique [9-11]. In recent years, many studies turn to pay attention to the effect of network structure and formation by using social network analysis [12-21].

Social network analysis describes the structure of relationships formed by connected and interacting actors [4]. And the perspective of social networks provides us a powerful tool for the analysis of configurations of technologies in a dynamic technological environment. Nohria [22] stated that “networks are as much process as they are structure, being continually shaped and reshaped by the actions of actors who are in turn constrained by the structural positions in which they find themselves”. Vanhaverbeke et al. argue that a firm with superior network positions has better opportunities to pool more resources to exploit technological development and to explore new technologies [23]. The past literature also reflects a consensus on the fact that the way of embeddedness of a firm in the inter-firm network makes a big difference to its economic and innovative performance [5, 24, 25]. In addition, Stuart [2] argued that high-prestige organizations with previously sponsored successful innovational paths tend to dominate superior positions in the market. In this respect, we further argue that firms occupy market positions through their innovative products or
advance technologies, and their positions emerge from the structure of technological network in which they are embedded.

Besides traditional network analysis, affiliation network is widely discussed in the literature such as project embeddedness [6, 26], hyperlinking of web sites [27], career networking in biotechnology [28], interlocking directors [29], and collaboration networks [4, 30-35]. Although there were a significant amount of studies in various issues, there are still only a few studies exploring the technological network positions by means of the affiliation network approach.

What is an affiliation network? An affiliation network is a kind of 2-mode network consisting of 2-mode data in which sets of relations connect “actors” and “events” [36]. In social network analysis, the affiliation network especially refers to membership or participation data. The advantage of the affiliation network approach is as follows [26, 37, 38]: First, studying an affiliation network allows us to retain all the necessary information available in the affiliation matrix; thus, no data is lost. Moreover, in an affiliation network links between actors are also considered, although not directly but through events. Additionally, we are always able to know which actors attended which events. Lastly, one-mode networks and two-mode networks can be considered complementary. In the other words, the affiliation network offers a dual perspective on network relationships between actors and events [38], rather than direct ties between pairs of actors as in one-mode data. In short, the reasons that we utilized affiliation network analysis to analyze our research issues are threefold: first, only a few studies have attempted to identify social positions using the affiliation networks approach [26, 39]. Second, affiliation networks describe collections of actors rather than simply ties between pairs of actors [36]. Third, the result of the analysis provides a dual perspective of actors and events [39-41].

The main purpose of this study was to examine the interactions between the focus of R&D and technological areas, and to inquire how the technological network shapes the technological position in the technological network. Affiliation network analysis allowed for a novel examination and an exploration of how R&D activities interact. In addition, the network measurements of centralities were investigated. Specifically, we also looked at the network structure by examining the co-membership of technologies and research areas. We selected patent information as our empirical data, because patents identified inventions or technological innovations and further influenced position in the technology network. “Actor” and “Event” are two basic elements of constituting affiliation network. The "actor" can be an individual, technology, organization or group of individuals. In this paper, we treat the patent as an actor. In addition, we argue that a technological classification (United States Patent Classification, USPC) listed in the patent documentation is as a technological area in which the firm participates a technological activity in technological environment. Therefore, a technological classification is assumed as a technological “event” in our analysis. In sum, we defined patents as the actors and USPCs as the events. Patents and USPCs jointly constitute a affiliation network. The usage of affiliation network is justified for this study.

This paper proceeds in the following manner: in section 2, we briefly review the relevant literature on the social network analysis and affiliation network. The methodology and measures are also described in section 2. Data and analysis are presented in section 3-5. Results are then summarized in section 6. Lastly, we discuss our results and offer conclusions in section 7. Our perspectives and empirical results provide suggestions for making R&D decisions.

2 Social network analyses

Social network analysis is not new to technological studies. The concepts of network and networking have gained considerable attention in innovation and invention studies [42-45]. Actor and relation constitute two basic units of network analysis. The definition of an actor depends on the research purpose [46]. An
actor may be represented as a person, an organization or a technology. In this paper, we treat patents as actors in the technological network.

In social network analysis, the term “mode” refers to a class of entities. In the 1-mode case, the mode usually refers to actors whose members have social ties with other members, such as John and Marry who are acquainted with each other. In the 2-mode case, actors are also members of a class, such as John joins a health club. A matrix form is used to represent the network data for analysis purpose. The affiliation network consists of 2 sets of data of actors and events. Therefore, an affiliation network is one kind of 2-mode matrix. A description of affiliation network is as follows.

2.1 The affiliation network

Many social network relations are linked through actors’ joint participation in social activities or membership in collectivities [26, 37]. This common activity that actors participate in is called the “event” and creates a network of ties among actors. The affiliation network often presents a relation from the units in one set to the units in the other set, such as “person-by-event”. Therefore, the affiliation matrix denotes people attending events, joining projects or organization employing people.

Let the set of actors is denoted by \( N = \{n_1, n_2, n_3, \ldots, n_g\} \) and the set of events is denoted by \( M = \{m_1, m_2, m_3, \ldots, m_h\} \), and g actors and h events form a \( gxh \) matrix. The affiliation matrix can be represented by \( A = \{a_{ij}\}_{g \times h} \), where \( a_{ij} \) records the affiliation of each actor \( i \) with each event \( j \). The value of \( a_{ij} \) is also simply a zero or a 1. \( a_{ij} \) is calculable by means of the following formula:

\[
a_{ij} = \begin{cases} 
1 & \text{if actor } i \text{ is affiliated with event } j \\
0 & \text{otherwise}
\end{cases}
\]

In social network analysis, the term “affiliations” is especially referred to the cases when the data consist of some kind of participation or membership [36, 37, 40, 47]. The most classical example of an affiliation network is the “Deep South” data collected by Davis et al. This data set recorded 18 women attending 14 social events over a 9-month period. Davis et al. used this data to investigate the clique structure (social classes) among these women[48].

Furthermore, the example of Faust is provided to illustrate an affiliation network[37]. The affiliation network matrix which consists of six actors ( \( N = \{n_1, n_2, n_3, \ldots, n_6\} \) ) and three events ( \( M = \{m_1, m_2, m_3\} \) ) is presented in Table 1.

![Table 1: Adjacent matrix of affiliation network](image)

Data source: Faust [37]

The affiliation network is one kind of 2-mode matrix containing two sets of nodes. Therefore, the structure of an affiliation network can be represented as a bipartite graph. So there are 6+3 nodes in the bipartite graph. In the bipartite graph, the lines indicate ties of affiliation between actors and events. Figure 1 depicts the bipartite graph of the affiliation network of Table 1 [37].

![Figure 1: Bipartite graph of affiliation network](image)

In sum, affiliation network analysis allows for a novel examination and provides different perspectives [38]. In this study, we are interested in the interrelationships between R&D activities and research areas and looked at the structure of the affiliation network of R&D activities associated with technology.

2.1.1 Co-affiliation matrix
In some cases, the purpose of research is often to understand the pattern of ties within one of the sets, not to understand the pattern of ties between the two sets. In this regard, Faust suggested 3 ways for studying an affiliation network [37]: (1) properties of actors in the one-mode relation of actor co-memberships; (2) properties of events in the one-mode relation of event overlaps; and (3) properties of both actors and events in the two-mode affiliation relation.

Along with the affiliation data, we are also interested in ties among members of the actors set as well as members of the events set, but the affiliation data does not include ties among members of either set, yet it can be converted into co-affiliations [39]. Following the definition by Faust [37], $A = \{a_{ij}\}_{g \times h}$ is an affiliation matrix of the $g$ actors with the $h$ events. $X^N$ denotes the number of memberships shared by each pair of actors, and $X^M$ denotes the number of actors shared by each pair of events. The co-affiliation matrix of actors (co-membership matrix, $X^N$) and overlap matrix of events (event overlap matrix, $X^M$) are then obtained through the following equations:

$$X^N = AA^T$$  \hspace{1cm} (1)

and

$$X^M = A^T A$$  \hspace{1cm} (2)

In this study, the use of co-affiliation matrix is twofold. First, the co-membership matrix is used to investigate the frequency of two patents join the same technological event. Second, the event overlap matrix is used to find out the most active technological event.

2.2 Measures of centrality

In general, centrality refers to a family of network properties of node positions. Measures of centrality focus on the number of ties and length of distance in the network that an actor has with other members of the network [49]. Different aspects of centrality have been discussed extensively by Wassermann and Faust [36], Freeman [50], Knake and Burt [51], Bonacich [52], Faust and Wasserman [53]. In a sense, the use of centrality measures gives us some indication of the effect of the technological network on the relationship between technologies and technological classifications[54]. Four kinds of centrality are discussed below.

2.2.1 Degree centrality ($C_D(n_i)$)

Degree centrality is defined as the number of ties adjacent to a node $i$. $C_D(n_i)$ denote the degree centrality of actor $n_i$ [36], and can be calculated by means of the following formula:

$$C_D(n_i) = d(n_i) = \sum_j a_{ij} = \sum_j a_{ji}$$  \hspace{1cm} (3)

However, it is usual to normalize centrality measures by dividing by the maximum value possible in a network of that size. But for affiliation networks, we must apply two separate normalizations depending on which node set a node belongs to. The higher measure indicates the greater potential for activity within the flow of communication [50].

2.2.2 Betweenness centrality ($C_B(n_i)$)

Betweenness centrality is defined as the “share” of shortest paths in a network that pass through a node $i$ [47]. Assuming there is more than one geodesic between $n_j$ and $n_k$, all geodesics are equally likely to be used. Let $g_{jk}(n_i)$ be the number of geodesics linking the two actors that contain actor $n_i$, and let $C_B(n_i)$ denote the betweenness centrality of actor $n_i$ [36], then

$$C_B(n_i) = \sum_{j \neq k} \frac{g_{jk}(n_i)}{g_{jk}}$$  \hspace{1cm} (4)

Betweenness centrality captures the capacity for an actor to play the role of intermediary in the network [38]. In a sense, the betweenness of a technology or technological classification is a function of paths from one technology to another, from technology to technological classification, and from technological classification to technological classification..

2.3.3 Closeness centrality ($C_C(n_i)$)

Closeness centrality is defined as the sum of geodesic distances from node $i$ to all others in the network [47]. Let $d(n_i,n_j)$ be the number of lines in the geodesic linking actors
\( \sum_{j=1}^{n} d(n_i, n_j) \). Then the total distance that \( n_i \) is from all other actors is \( \sum_{j=1}^{n} d(n_i, n_j) \). \( C_C(n_i) \) denotes the closeness centrality of actor \( n_i \) [36] and is calculable by means of the following formula:

\[
C_C(n_i) = \frac{1}{\sum_{j=1}^{n} d(n_i, n_j)}
\] (5)

To normalize closeness in the affiliation network, we simply multiply \( C_C(n_i) \) by \((g - 1)\) [36]:

Scott indicates that if a node lies at short distance from many other points then a node is “close” [49]. In a sense, closeness indicates the potential independence of an actor from the flow of communication [50].

### 2.2.4 Eigenvector centrality (\( C_E(n_i) \))

Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix of a graph [55]. It may be thought of as a weighted degree measure in which the centrality of a node is proportional to the sum of centralities of the nodes it is adjacent to [26]. So, in eigenvector centrality, a node’s score is proportional to the sum of the scores of its neighbors [47]. Assuming \( S \) contains a pairwise similarity measure between \( n_i \) and \( n_j \). Let \( \lambda \) be the largest eigenvalue and \( X \) be the corresponding eigenvector. \( C_E(n_i) \) denotes as an eigenvector centrality of actor \( n_i \) [36], is measured by the formula:

\[
C_E(n_i) = \frac{1}{\lambda} S X
\] (6)

or

\[
C_E(n_i) = \frac{1}{\lambda} \sum_{j=1}^{n} \alpha_{ij} x_j
\] (7)

In general, eigenvector centrality captures not only how many actors you “know,” but how many actors they “know” as well [38]. Hence, eigenvector centrality is a measure of the “importance” of a node in a network. In this manner, a technology that is connected to many other technologies (with high degree centrality), who are themselves well connected (also with high degree centrality) has a high level of eigenvector centrality.

### 2.3 Core/periphery structure

We employ a core/periphery model proposed by Borgatti and Everett [56] to study the structural differentiation of the technological network. Borgatti and Everett [56] defined the core/periphery structure as follows: “The core/periphery model consists of two classes of nodes, namely a cohesive subgroup (the core) in which actors are connected to each other in some maximal sense and a class of actors that are more loosely connected to the cohesive subgroup but lack any maximal cohesive with the core”. In other words, for the network structure, the “core” is characterized by a high density of interrelations in contrast to the “periphery” of a more loosely connected class of actors [57].

The measurement of how well the observed structure approximates the ideal structure is also proposed by Borgatti and Everett as follows[56]:

\[
\rho = \sum_{i} \sum_{j} \alpha_{ij} \epsilon_{ij}
\] (8)

Where:

\[
\epsilon_{ij} = \begin{cases} 1 & \text{if } i \in \text{core or } j \in \text{core} \\ 0 & \text{otherwise} \end{cases}
\] (9)

In the equations, \( \alpha_{ij} \) and \( \epsilon_{ij} \) indicate the existence or absence of a tie between any two actors \( n_i \) and \( n_j \) in the observed network and a corresponding tie in the ideal core/periphery structure, respectively. The value of \( \rho \) is maximum if the observed network (matrix of \( \alpha_{ij} \)) is identical to the ideal core/periphery structure (matrix of \( \epsilon_{ij} \)). However, it is hard to expect that an observed network can completely match an idealized pattern. We can readily appreciate that observed structures will only approximate idealized pattern to some degree. Hence, we claim that the observed network shows core/periphery structure if the value of \( \rho \) is sufficiently high.

### 3 Data

We chose patent data as our empirical data. Patents are direct outcomes of R&D. Many technological studies used patents as empirical data and patents provide extremely valuable
insight into the technology development process [58, 59]. The patent data used in this study came from the United States Patent and Trademark Office (USPTO). The advantages of using patent data in this analysis are as follows: First, patents identify the inventions or innovative technologies[60] and relationships between technologies[61]. Second, very few studies have taken a network perspective on the technological activities behind patents.

We selected insurance business method patents as our analyzing objects. In general, the technological classification 705 (USPC705) in USPTO is referred to as the “business method” class and is an emerging technology area to fulfill demands of customer in E-Commerce. The forth subclass of USPC705 entitled as “insurance” is pertinent to the insurance sector and is a good example to study. In our analysis, we assume the following; first, the technological classification represents a specific applicable technological area or technological event in which R&D activities are involved. Second, various inventors have different research competencies and skills, etc. von Hippel [62] argues that inventors who are limited and far from routine are a good example of the kind of tacit knowledge. The number of inventors somehow represented in a given patent is determined to be the complexity of labor in that particular R&D activity. In this manner, technological classifications are the common technological events where patents jointly participate. Therefore, the affiliation network is a good way to facilitate knowledge and skills sharing within a technological event and to form a technological network with a particular structure.

In total, 779 patents of the insurance business method were extracted from the USPTO database for the period 1985/01/01-2011/01/01. These patents were included as they represent a significant effort by the inventors to pursue new areas in research and application. Patent data were analyzed by the PatentGuilder software. The UCINET network analysis software by Borgatti, et al. [63] Borgatti, Everett and Freeman was performed on the affiliation network.

4 Analyses
We were interested in determining to what extent technological networks facilitate or influence the formation of the technological position of patents. In other words, how patents might be “connected” to one another through shared USPC, which could highlight which patents are most important, and if there are particular structural configurations of a network which yield core technology, or the most active technological event, or a core/periphery regime of technologies.

As Moody, et al. [64] states: “small networks can focus on detailed elements of the graph structure while larger networks can mainly capture gross topology. Visualizing networks of tens of thousands of nodes requires further abstraction yet”. In order to explain in detail the network’s effects, we abstracted the large scale network of 779 patents and calculated network measures for each patent.

4.1 Data screening
Moody, et al. [64] suggested abstracting the nodes into less than 100 and that usually less than 50 nodes appropriate for the purpose of analysis. Therefore, we need some criteria to abstract the scale of patents.

Hall, et al. [59] argues that forward citations of a patent are a good indicator of the economic value of the invention. In other words, the frequency of citations by subsequent patents of a given patent is one of the main indicators of the scientific and technological significance or quality of the invention. A number of studies have also confirmed the value of forward and backward citations in measuring patent value [14, 15, 59, 65]. The greater the number of citing firms, the better will be the market valuation of the firm that owns the patent [59]. Only a firm with successful innovational paths will dominate a niche position and with high prestige in the market [2]. This study tried to explore the technological position from technological networks. Therefore, the patents used for further investigation were based on the following criteria:

1. The number of forward citations of a patent is over 10 [14].
The patent had been co-cited by at least 5 firms [15].

(3) The patent right is under the name of a firm not an individual inventor[59].

(4) At least two inventors[16].

(5) The number of USPCs of a patent is greater than 4 (the average USPC of patents is 3.92).

In total, out of the 779 patents, only 91 patents fulfilled the above criteria and were selected. This number also met the criteria set by Moody [66]. The resulting sample of 91 patents consists of 16 technological classifications. These patents are with well representative and validity, and sufficiently allow us to address the institutional or structural context issues in a large scale sample.

4.2 Technological classification

The primary purpose of technological classification for a patent is to facilitate the searching and retrieving of patent documents by patent offices and users. Clearly defined patent classifications allow people to identify the classification areas in which the subject technology can be. International Patent Classifications (IPC) and U.S. Patent Classifications (USPC) are currently the major patent classification systems. There are two approaches of classification, one based on a function-oriented principle and another on an application-oriented principle. IPC is a combined function/application classification system in which the function theoretically takes precedence. USPC is principally a function-orientated classification[67], such as a classification according to the action carried out by the invention. Besides, the classification 705 was originally created by the USPTO for fulfilling demands of E-Commerce. Therefore, the classification approach of USPC matched the purpose of our analysis. In this paper the patents are the “actors” in our network analysis, and technological classifications are further assumed to be the “events” according to the R&D activities in the technology context based on the USPC system. 91 patents contained 16 USPCs, and we treated these USPCs as 16 distinct technological events. Table 2 shows the 16 technological events.

4.3 Affiliation matrix

The network data was built from the patent and USPC affiliations of the patents. USPCs are treated as the relational evidences of a patent participating in different technological events. These affiliations were then arranged as a 2-mode patent-by-USPC matrix $A$, where $a_{ij} > 0$ if patent i has a classification from USPC j and, and the value of $a_{ij}$ will become either 1 or 0. The co-affiliation matrix of patents and events are also calculated, respectively. In the co-membership matrix, for example, patent 05930759 and 04491725 joined the same technological event up to 3 times, and with patent 05873066 only 1 time. From event overlap matrix as table 3,
USPC705/002 is the most active technological event with 38 participations and USPC705/035 is the next. 13 patents co-participated in USPC705/002 and USPC705/003. 6 patents co-participated in USPC705/002 and USPC705/040, etc. Because the co-membership matrix is too huge to displace, we only displaced the event overlap matrix as table 3.

4.4 Measures of centrality
The properties of the network structure are needed to explore the affiliation of technologies to technological classifications and to identify the relative position of technologies located within the structure of the technological network. We were interested in exploring the connections of patents to USPC and further identifying those patents most centrally located within the technological structure. Four kinds of centralities were measured which were degree, betweenness, closeness, and eigenvector. Table 4 shows the affiliation centrality measures for each patent. And Table 5 shows the affiliation centrality measures for each technological event.

Observing table 4, the 7 patents with the higher centrality of degree, betweenness and closeness are patent 05956687, 04491725, 05890129, 05966693, 06304859, 06647374 and 06937990. This indicates that these patents had a direct association with many other patents, and these patents should be recognized by others as a major channel of relational information. Especially patent 05956687 had the highest value among all. But which patent is the most important one?

Eigenvector centrality is a measure of the influence of an actor in a whole network. Therefore, the score of eigenvector centrality had to be checked. The patent 04491725 had highest score (0.224) of eigenvector and was located in the centre of the technological network connecting 4 events (USPC705/002, USPC705/040, USPC705/003 and USPC705/039) as well. We may reasonably assume that this patent might be the most important patent among all the patents for insurance industry. Patent 06076066 with score 0.2 of eigenvector is worthy to pay attention. Patent 06076066 was granted in 2000/06/13 and with 25 citations by others. Patent 05884274 and 06128598 were with high degree and closeness but low eigenvector. While we inspected the network structure, these two patents connected 3 events and located at the same position in the network.

Looking at Table 5, Event USPC705/002 has 0.478 of betweenness centrality and 0.890 of eigenvector centrality. This implicates that event USPC705/002 is the most active technological event with respect to most of patents related to this area. The next is event USPC705/003. On the contrary, event USPC705/36R is unpopular.

Table 3: The event overlap matrix of 16 USPCs

<table>
<thead>
<tr>
<th>USPC</th>
<th>705/002</th>
<th>705/035</th>
<th>705/003</th>
<th>705/038</th>
<th>705/36T</th>
<th>705/001</th>
<th>705/007</th>
<th>705/037</th>
<th>705/039</th>
<th>705/040</th>
<th>705/008</th>
<th>705/039</th>
<th>705/101</th>
<th>705/104</th>
<th>709/217</th>
</tr>
</thead>
<tbody>
<tr>
<td>705/002</td>
<td>38</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>705/035</td>
<td>0</td>
<td>16</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>705/003</td>
<td>13</td>
<td>1</td>
<td>14</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>705/038</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/001</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/36T</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/007</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>705/037</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/36R</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/040</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>705/400</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No.</td>
<td>Patent</td>
<td>Degree</td>
<td>Closeness</td>
<td>Betweenne</td>
<td>Eigenvect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>05930759</td>
<td>0.188</td>
<td>0.536</td>
<td>0.006</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>05873066</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>06208973</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>05950169</td>
<td>0.125</td>
<td>0.458</td>
<td>0.028</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>05301105</td>
<td>0.125</td>
<td>0.492</td>
<td>0.002</td>
<td>0.155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>05797134</td>
<td>0.063</td>
<td>0.350</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>04491725</td>
<td>0.250</td>
<td>0.576</td>
<td>0.041</td>
<td>0.224</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>04858121</td>
<td>0.188</td>
<td>0.570</td>
<td>0.031</td>
<td>0.165</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>05809478</td>
<td>0.063</td>
<td>0.397</td>
<td>0</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>06119093</td>
<td>0.125</td>
<td>0.460</td>
<td>0.010</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>04642768</td>
<td>0.063</td>
<td>0.428</td>
<td>0</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>06343271</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>04722055</td>
<td>0.063</td>
<td>0.320</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>04839804</td>
<td>0.063</td>
<td>0.320</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>05907848</td>
<td>0.063</td>
<td>0.380</td>
<td>0</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>05870711</td>
<td>0.063</td>
<td>0.443</td>
<td>0</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>05970464</td>
<td>0.063</td>
<td>0.430</td>
<td>0</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>06336096</td>
<td>0.063</td>
<td>0.450</td>
<td>0</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>05613072</td>
<td>0.125</td>
<td>0.527</td>
<td>0.022</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>05704044</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>05956687</td>
<td>0.313</td>
<td>0.645</td>
<td>0.177</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>06064983</td>
<td>0.063</td>
<td>0.373</td>
<td>0</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>06009402</td>
<td>0.063</td>
<td>0.350</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>05070452</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>05752236</td>
<td>0.063</td>
<td>0.373</td>
<td>0</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>05915241</td>
<td>0.188</td>
<td>0.573</td>
<td>0.035</td>
<td>0.202</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>06344126</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>05550734</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>05225976</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>05235507</td>
<td>0.125</td>
<td>0.492</td>
<td>0.002</td>
<td>0.155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>05253164</td>
<td>0.063</td>
<td>0.490</td>
<td>0</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>05754980</td>
<td>0.063</td>
<td>0.428</td>
<td>0</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Affiliation centrality measures of 91 patents
If we consider whether a particular event might be able to pair with other patents/events in the network, betweenness centrality is a more useful measurement to be checked, because interactions between two nonadjacent actors might depend on the other actors in the set of actors, especially the actors who lie on the paths between the two. Event USPC705/002 has the highest betweenness centrality. This clearly indicates that event USPC705/002 (score 0.89) and USPC705/003 (score 0.394) play a critical role here. Besides, we turned to check the event overlap matrix (Table 3). Event USPC705/002 lied on the

<table>
<thead>
<tr>
<th>No.</th>
<th>Patent</th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweenness</th>
<th>Eigenvector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>705/002</td>
<td>0.063</td>
<td>0.450</td>
<td>0.012</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>05884274</td>
<td>0.188</td>
<td>0.533</td>
<td>0.026</td>
<td>0.007</td>
</tr>
<tr>
<td>3</td>
<td>05903873</td>
<td>0.125</td>
<td>0.533</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>4</td>
<td>04750121</td>
<td>0.063</td>
<td>0.428</td>
<td>0.006</td>
<td>0.010</td>
</tr>
<tr>
<td>5</td>
<td>06272528</td>
<td>0.125</td>
<td>0.400</td>
<td>0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>6</td>
<td>05664109</td>
<td>0.125</td>
<td>0.533</td>
<td>0.003</td>
<td>0.015</td>
</tr>
<tr>
<td>7</td>
<td>06003007</td>
<td>0.125</td>
<td>0.547</td>
<td>0.064</td>
<td>0.007</td>
</tr>
<tr>
<td>8</td>
<td>07343307</td>
<td>0.063</td>
<td>0.490</td>
<td>0.132</td>
<td>0.008</td>
</tr>
<tr>
<td>9</td>
<td>05970479</td>
<td>0.063</td>
<td>0.390</td>
<td>0.000</td>
<td>0.006</td>
</tr>
<tr>
<td>10</td>
<td>06330551</td>
<td>0.063</td>
<td>0.390</td>
<td>0.000</td>
<td>0.024</td>
</tr>
<tr>
<td>11</td>
<td>05926800</td>
<td>0.125</td>
<td>0.500</td>
<td>0.005</td>
<td>0.139</td>
</tr>
<tr>
<td>12</td>
<td>04987538</td>
<td>0.063</td>
<td>0.490</td>
<td>0.132</td>
<td>0.190</td>
</tr>
<tr>
<td>13</td>
<td>05202827</td>
<td>0.063</td>
<td>0.320</td>
<td>0.000</td>
<td>0.190</td>
</tr>
<tr>
<td>14</td>
<td>06077890</td>
<td>0.125</td>
<td>0.533</td>
<td>0.003</td>
<td>0.190</td>
</tr>
</tbody>
</table>

Table 5: Affiliation centrality measures of 16 technological events
paths among 38 patents, event USPC705/035 among 16 patents and USPC705/003 among 14 patents. Event USPC705/002 was adjacent to the majority of patents.

4.5 The affiliation network structure

The affiliation matrix can be presented as a bipartite graph. UCINET software was used to yield the diagram. However, it is important to notice that the affiliation network relies solely on the pattern of connections, rather than the spatial position of the nodes or the length of lines [26]. Figure 2 represents the network structure of connections between patents and technological events.

In the graph, patents are represented by round nodes and technological events by square nodes. This bipartite graph allowed us to identify the structure of technological relations among the patents and technological events. It is obvious that event USPC705/002 is located in the centre position of the network with the greatest number of patents connected; thus, it plays the main role within the whole network. In contrast, event USPC705/400 and event USPC705/36R apparently are in the periphery position in network. Event USPC705/035 (betweenness 0.243), USPC705/001 (betweenness 0.124) and USPC705/003 (betweenness 0.121) act as intermediaries and bridge the different patents and technological events. The configuration of the network’s structure matched the result of centrality measurement.

5. Core/periphery structure

Besides the bipartite graph, the core/periphery structure of the affiliation network was also demonstrated. We ran UCINET to yield the core/periphery diagram of the patents and technological events (see Figure 3 below). Figure 3 obviously demonstrates that 45 patent are located in the core position of the network. Event USPC705/002, USPC705/035, USPC705/003 were the most active events.

The density of each block is shown in Table 6. The density of the core is 0.367; in contrast the density of the periphery is 0.061. The starting fitness of structure is 0.280 and the final fitness of structure is $\rho =0.366$. This means that the correlation coefficient of Figure 3 to the ideal structure is 0.366 ($\rho$ is sufficiently high). This indicates that this affiliation network of patents and technological events has strong structure differentiation according to their participation.

<table>
<thead>
<tr>
<th>Position</th>
<th>Core</th>
<th>Periphery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>0.367</td>
<td>0.091</td>
</tr>
<tr>
<td>Periphery</td>
<td>0.153</td>
<td>0.061</td>
</tr>
<tr>
<td>Structure fitness $\rho =0.366$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6 Results

The social network analysis was conducted to study the interaction between patents (R&D activities) and technological classification (technological areas). 91 patents were selected out of 779 patents to demonstrate the structural
effects of the network and elicit a detailed explanation of how the network’s effects shape the positions. These 91 patents totally included 16 USPC. In this paper, we treated patents as “actors”, and we further assumed that these 16 USPC were as the 16 different “technological events”.

The affiliation matrix was then arranged as a 2-mode patent-by-USPC matrix. The affiliation matrix was then converted into co-membership matrix and event overlap matrix. The event overlap matrix showed that technological event USPC705/002 was the most active, and the next were USPC705/003 and USPC705/035. From co-membership matrix, we might be able to check the technological similarity of patents for how many time of two patents joined the same technological event. For example, 05930759 and 04491725 were somewhere in comment, because these two patents joined the same technological event for up to 3 times. Moreover, the properties of the network structure were measured to identify the relative position of technologies located within the technological network. 4 kinds of centralities were calculated. Each indicator represents different implications for positions in the network. For the structural properties of the network, an actor which located in the central position has the maximum possible degree, and all other actors are maximally close to it. Besides, it falls on the geodesic connections between the largest possible numbers of others. Table 4 and table 5 showed that every patent and technological event had its own position in network according to the network measurement of centrality.

In addition, to visualize the affiliation matrix, a bipartite graph and core/periphery structure of the 2-mode network was produced. The visualization of the affiliation matrix is consistent with the centrality measurements.

7 Discussion and conclusion
Social network analysis can be a useful tool in determining the relationship between the R&D activities and technological areas. Technological affiliation data may provide important information on technology network and the spread of innovations. Analysis of the technological affiliation network may further provide information on how to target key innovation to maximize coverage of technologies for the purpose of learning and invention. The concepts of affiliation networks used here are to provide the dual perspectives of which patents are linked to each other as members of a given technological activity, and activities are linked to each other through shared patents. In this paper, we treat a patent as an “actor” and USPCs technological classification as the technological “event” in which R&D activities are involved. Through an analysis of the affiliation network, we inquired into the technological position in the technological network. We obtained a clear and detailed interpretation from the result of the affiliation network analysis and intuitive interpretations from the structure of the bipartite graph.

Instead of thinking of the bipartite structure in terms of patents and technological events in which interactions are concentrated, or as the core/periphery structure of patents and technological events who engage in similar R&D activities with others, positions identified from affiliation networks might lead us to focus on how technological events gather patents together and how patents bring technological events together. More importantly, positions are not defined solely by technological area but are also occupied by patents. We investigated the detail by means of the social network approach. In our results, 16 distinct technological events are mutually composed by a subset of patents. A detailed investigation showed technological interactions between patents and their attending technological events. For example, a patent/event brings several different technological events/patents together. Patent 05930759 brings USPC705/002, USPC705/003 and USPC705/040 together. Patent 04491725 brings USPC705/002, USPC705/003, USPC705/040 and USPC705/039 together. Patent 05999917 brings USPC705/36R, USPC705/037 and USPC705/035 together. In other word, patents interact while attending technological events, and technological events reversely contribute to
R&D activity through the patents that commonly participate in them. This phenomenon implies a spillover of knowledge and skills.

The results of the study suggest that the number of USPC is indeed an important factor that contributes to technological importance of R&D. Indeed, degree centrality is defined as the number of links incident upon a node. Therefore, a patent with high degree can be interpreted as a major channel of technological information recognized by others. Betweenness may be defined as the number of geodesic paths that pass through a given node [26]. This indicates a patent with higher betweenness functions as a bridge from patent to patent, from patent to technological event and from technological event to technological event. Therefore, the strategy of having R&D activities function as bridges between distinct clusters of technological areas apparently has a benefit for technological development. Eigenvector centrality is a measure of the influence of a node in a network. This implies that a patent with higher eigenvector is in a position to influence others in the technological network. This patent could be the most popular and play a key role in technological development. Finally, our perspectives and empirical results can provide suggestions for making R&D decisions. And identifying technological positions from patent-by-USPC matrix here represents an example of technology studies to explore similar issues from two-mode data. Of course, the ultimate usefulness of our method depends on the extent to which the positions we identify constitutes a significant technological phenomena.

References


[46] D. Knoke and S. Yang, Social network analysis: Quantitative applications in the social sciences, 2


