Advanced Engineering Informatics 27 (2013) 366–377

Contents lists available at SciVerse ScienceDirect

Advanced Engineering Informatics

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

Exploring technology evolution and transition characteristics of leading countries: A case of fuel cell field

Ssu-Han Chen [a], Mu-Hsuan Huang [b], Dar-Zen Chen [a,*]

[a] Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, 10617 Taipei, Taiwan, ROC
[b] Department of Library and Information Science, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, 10617 Taipei, Taiwan, ROC

ARTICLE INFO

Article history:
Available online 5 March 2013

Keywords:
Patent analysis
Technology evolution
Leading countries
Transition patterns
Fuel cell technology

ABSTRACT

In the course of the technology evolution, the status of leading countries changes with the times. Leading countries may persist, appear, or disappear, i.e. the change of which is named "transitions" which may be effected by some characteristics. An integrated solution is proposed in this study which first identifying technology topics as well as the leading countries over time and then evaluating the relative importance of characteristics and the zero, positive, or negative significant effects of characteristics on different transitions. Analyzing a set of patents related to fuel cell, we found that this technology consisted of several communities in which the segmentation was achieved mainly based on the types of fuel cell. Among those communities, the management utilities were invested mostly, followed by components of polymer electrolyte membrane fuel cell and direct methanol fuel cell, followed by solid oxide fuel cell. The United States and Japan always dominated each sub-domain and Germany and Korea emerged recently. On the other hand, we found that the science linkage (SL), technology cycle time (TCT) and pending duration (PD) were the top three important discriminators for the transitions, whereas the originality index (OI) was the least important one. The stable countries presented increasing manners of the SL, technology dependence (TD), and eigenvector centrality (EC), but a decreasing manner of the TCT. The appearing countries presented increasing manners of the SL and PD, but decreasing manners of the TCT and clustering coefficient (CC). The exiting countries presented an increasing manner of the TCT, but decreasing manners of the SL and degree centrality (DC).

1. Introduction

In the constantly changing global environment, it is valuable to examine technology trends in order to look deeply into the technology structure and innovation development. Previous researches have attempted to use dynamics to explore how technologies had evolved over time [1–4]. The results of analyzing a bulk of documents are usually presented by a series of longitudinal maps where the technology structure of each snapshot is composed of patents which are tied by interdependence. Interests are especially highlighted in understanding underlying community-level structures where patents share a passion for a particular topic and interact to expand their field expertise [5]. In general, a technology topic is a coherent set of subject-related problems, concepts, apparatuses, or methods to which attention is paid by inventors. A field explanation, on the other hand, can be presented by a subsequent topic with novelty, enhancement, improvement, or additional functions.

Furthermore, in the study of technology development and innovation, it has long been recognized that certain participants make an unusually larger contribution than their peers [6]. We sought to introduce a rational and quantifiable method to identify such leading roles at a national level in which the leading countries varied according to the evolving topics over time. So-called leading countries are characterized by a large number of high-quality contributions to the inventive performance of their nations in each snapshot. They play prominent roles worldwide in the design, expansion, and integration of prior arts within their countries as there are abounding invention teams with different technological and scientific specialization.

We put forward here a framework that presented the evolving topics and the corresponding leading countries. The fuel cell field was served as an example in this study because it was worthy to explore the historical trends of such a developing field. Fuel cells are widely considered to be efficient and non-polluting power sources that offer much higher energy density and energy efficiency compared to other current or conventional systems [7]. In so doing, technology communities of a full-time dataset were detected using a community detection algorithm. Then communities' technology topics as well as leading countries over successive

* Corresponding author. Tel.: +886 2 3366 2723; fax: +886 2 2369 2178.
E-mail addresses: ssuhanchen@ntu.edu.tw (S.-H. Chen), mhhuang@ntu.edu.tw (M.-H. Huang), dzchen@ntu.edu.tw (D.-Z. Chen).

1474-0346/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aei.2013.02.001

snapshots were identified using the natural language processing and a two-dimensional scheme of Pilkington et al. [6] respectively. It is commonly believed that monitoring the technology evolution can be useful for effectively categorizing and tracking the changes of the technology dynamics. Such an understanding helps lead to a fuller comprehension of the innovation and technology development process and also the promotion of techniques to improve the efficiency of innovation activities at national levels [8].

In addition to identifying the evolution of technology topics and leading countries, another main objective of this study was to test an impact of several characteristics on countries' sustained behaviors of productivity and qualities. In terms of the importance of leading countries it was surprising that this issue had not yet been empirically examined. Therefore, this exploratory study attempted to offer some preliminary insights into this issue. In order to achieve this objective, we first transform all the identified results of leading countries of each snapshot into three types of transition patterns, including stable, appearing, and exiting in terms of their status of existence or non-existence during two successive snapshots. Characteristics including technology cycle time (TCT), science linkage (SL), pending duration (PD), originality index (OI), technology dependence (TD), degree centrality (DC), clustering coefficient (CC), and eigenvector centrality (EC) of the leading countries that are irrelevant with the quantity or quality are extracted over time and are then converted into the rate of changes during the two successive snapshots. With a set of input (i.e. the rate of change of characteristics) and output (i.e. transition patterns) pairs available, we suggest using a supervised learning model, back-propagation neural network (BPNN) that had been widely used in the patent analysis [9–13], to investigate such non-linear and complex input–output relationships. Finally, we partition the learned synaptic weights of the BPNN to determine the relative importance of those characteristics and show typically different rate of changes at each transition pattern by the means of statistical testing. Such an understanding provides an exploratory perspective that enables us to examine the possible reasons for characteristic fluctuation on transitions of leading countries.

The rest of this paper is organized as follows. In Section 2, we describe and explain the research methodology. In Section 3, we depict the experimental environment and present the results. The concluding remarks and further suggestions are discussed in Section 4.

2. Methodology

Patent documents related to a specific technology are collected and are arranged in increasing order by issue date. The similarities among all of patent pairs are calculated to construct a patent network which is clustered into communities. Patents in each community are divided into non-overlapping buckets. A series of technology topics and leading countries are identified longitudinally in terms of key terms and quantity–quality measurements respectively. Then we switch the focus on realizing the characteristics that help explain the transition patterns of leading countries. The analysis can be done through BPNN, relative importance measurement, and significant tests. The detailed process of the proposed method is illustrated in Fig. 1 and has been explained as follows.

2.1. Identification of technology topics and leading countries over time

The first research objective is to develop a generic patent analysis framework for knowledge discovery of the evolution of technology topics as well as their corresponding leading countries over time. The detailed process is explained below.

2.1.1. Community detection of the patent network

The citation-based patent analysis that covers bibliographic coupling, co-citation, and direct citation is one of the commonly used approaches for measuring the patent similarity. Among them, bibliographic coupling is chosen in this study because it provides current and immediate information about patent relationships [14] and reinforces regions of dense citation [15]. In general, bibliographic coupling measures the similarity between patents by the number of references two patents share in common. However, the raw coupling strength is too rough to serve as a measure to represent similarities, so there is a need to consider the coupling strength as well as the strength of each patent [16]. Hence, the coupling strength of the document pairs should be normalized based on Salton's cosine [17]. After the coupling strengths are normalized, there is also a need to select the relatively strong Salton's cosine since partial ones are extremely weak. This study chose employing a relative threshold introduced by Chen et al. [18] to select strong Salton's cosine since the counterpart, absolute threshold, e.g. top 25% rule [19] or user-specified coupling strength [20,21], has the disciplinary bias in the average normalized coupling strength frequencies. The relative threshold is adopted to select the strong bibliographic couplings with at least above the average and standard deviation of the Salton's cosine. With the set of nodes and ties of a given full-time dataset, the patent citation network can be composed. In the network analysis, technology communities are detected using the weighted Girvan–Newman algorithm [22] because it does not involve human judgment to set a parameter of number of clusters and it is suitable for detecting community structure in weighted networks. The weighted Girvan–Newman algorithm divides a network in terms of betweenness and modularity. The clustering result has the best split structure where there are many within-cluster ties and minimal between-cluster ties [22].

2.1.2. Identification of technology topics over time

After community structure is obtained, patents in a given community are divided into several non-overlapping buckets and then the corresponding technology topics over time are identified using key terms which are extracted using the natural language processing. The steps of key term extraction are as follows. Firstly, the patent titles and abstracts are collected as a corpus since they represent the most salient contribution of a document in condensed forms [23]. Secondly, the corpus carries out a purge of cleaning process by lower case conversion, punctuation and number removal, multiple whitespace stripping, and singularizing. Thirdly, each word is tagged a part of speech based on its context in the text [24]. Fourthly, three linguistic filters as shown in Eq. (1) are applied since most meaningful terms consist of nouns, adjectives, and sometimes prepositions [25]. With such filters, most of undesirable stop words are not included.

$$\begin{aligned}&\mathrm{Noun}+\mathrm{Noun}\\&(\mathrm{Adj}|\mathrm{Noun})+\mathrm{Noun}\\&((\mathrm{Adj}|\mathrm{Noun})+|((\mathrm{Adj}|\mathrm{Noun})*(\mathrm{NounPrep})?)(\mathrm{Adj}|\mathrm{Noun})*)\mathrm{Noun}\end{aligned} \tag{1}$$

Finally, these terms are scored by the term frequency-inverse community frequency scheme [4] to measure the frequency and uniqueness of terms in a certain community compared to the other communities. Such an automatic procedure assists the analysts to have a better interpretation in the process of interpreting communities' contexts and enables the analysts to realize the technological development and trends in a specialty.

2.1.3. Identification of leading countries over time

The leading roles of a few patentees seem to be a law of nature [26] and should be the leaders in underlying developing fields [8].

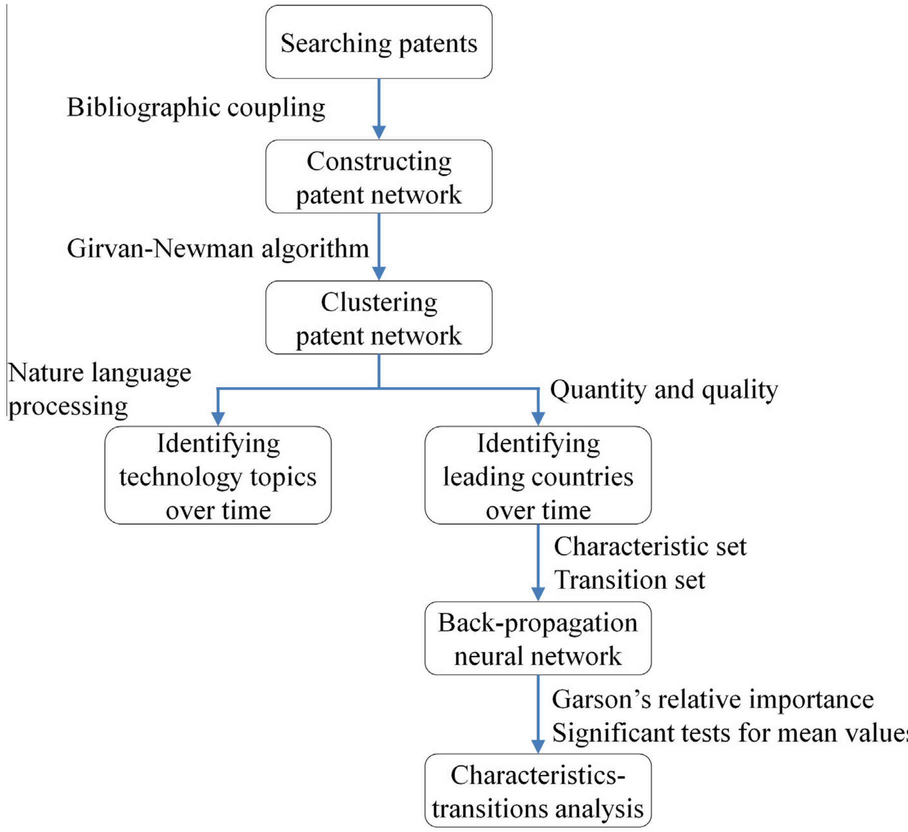

Fig. 1. Framework of the proposed methodology.

The analytical unit of the leading roles could be applied at various levels, from individual to country. Ernst has identified key inventors as those who have high patenting activity and high patent quality rating [27]. Later, Pilkington et al. extended Ernst's approach to help identify key firms which are considered to be both highly productive firms and widely cited ones [6]. In order to identify leading roles, this study further explores the framework of Pilkington et al. [6] which is applied at a country level. We detected the leading countries as the ones with higher than the average and standard deviation productivity (i.e. the number of patents granted) as well as quality (i.e. the number of citations received) compared to those of their peers. A two-dimensional scheme enables us to identify the most prolific countries whose patents are also highly cited by other countries. In our application, the leading countries could be considered subsequent to the procedure of natural language processing described in Section 2.1.2 for each bucket in a community. It is worth noting that the number of leading countries of a bucket could be zero, one, or more than one.

2.2. Characteristics related to transitions of leading countries

It is also our objective to deepen our knowledge about the main characteristics that help explain the transition patterns of leading countries. Three types of transition patterns are firstly defined. Characteristic sets are calculated and then converted into the rate of changes between two successive snapshots. Taking into account the input (i.e. the rate of change of characteristics) and output (i.e. transition patterns) information, we suggest using a BPNN model to investigate the input–output relationship. Finally, we dig out the relative importance of the various characteristics using Garson's method [28] and evaluate the significant zero, positive or negative effect of characteristics on different transitions using a statistical testing procedure. The detailed process is explained below.

2.2.1. The transition set

A leading country's transition at a given time is a change experienced by a leading country that has or has not been discovered at an earlier time. Within a community, leading countries may survive, new ones may emerge and old ones may die. Those transitions are named stable, appearing, and exiting patterns respectively in this study and the corresponding definitions are given as below:

- *Stable pattern*: leading countries that existed in two successive snapshots.
- *Appearing pattern*: leading countries in the 2nd snapshot that contained no unit from the 1st snapshot.
- *Exiting pattern*: leading countries that existed in the 1st snapshot but did not continue in the 2nd snapshot.

2.2.2. The characteristic set

A knowledge base is the foundation of a county's innovation [29]. To explore the phenomena in a deeper dimension of transitions, we investigated the changes in the nature of the knowledge base. Some characteristics, including attributive characteristics and topological ones, of knowledge base are adopted from the literature and industrial practice. The criterion of selecting characteristics is time-invariant property. Since training of the BPNN is achieved through static back-propagation, learning is restricted to time-invariant structures. In this study, the patent characteristics of an analytical unit are examined from different aspects covering backward citation, non-patent reference, application and issued date, USPC as well as self-citation because those metadata are available soon and are fixed after a patent is granted. We do not choose the time-variant characteristics such as generality, prosecution length, or patent age that take time to be triggered or accumulated. Young patents often show extremely low values or zeros due to their short lifetimes.

- Technology cycle time (TCT) is essentially the cycle time between generations of technology. TCT uses metadata, date of citing and citing patents, within an analytical unit to compute the median time lags, in years, of patent prior-art references of a set of patents. The time lag is computed from the granted date of a cited patent of an analytical unit to that of each citing patent [30]. TCT is considered to be the speed of invention. The smaller the TCT value is, the faster the technological progress will be.
- Science linkage (SL) reveals the strength linkage of knowledge resources from research had impact on technology development. SL needs references of all the granted patents from an analytical unit which should be further distinguished and count the number of non-patent references. Then SL is represented by the average number of scientific papers referenced in an analytical unit's patents [31]. The higher the SL value is, the more effects that science can exert on technology.
- Pending duration (PD) represents the time duration of the ultimately successful patents that have been in the application-grant process [32]. PD uses metadata such as publication date and the application date of an analytical unit's patents to compute the average time elapsed between those dates. PD with low value means that the examination process is quick and such patents are granted soon.
- Originality index (OI) measures the extent to which the patent is based on broad technological roots, because the patent is more likely to synthesize knowledge across a wide variety of disciplines [33]. The index is based on the technology categories of the inventions where the references and the corresponding United States Patent Classification (USPC) categories of an analytical unit's patents need to be prepared. A histogram of the USPC categories of the citing patents is constructed to indicate its technological roots distribution. OI is then built as a Herfindahl formula. A higher value means that the citations come from a more diverse set of technologies, otherwise they relate to monopolistic technologies.
- Technology dependence (TD) is considered the dependence of an analytical unit's technology development. Technology dependence (TD) can be measured by the proportion of self-citations [33], so the metadata and the first assignee country of citing and citing patents from an analytical unit are required. TD is the capability to protect one's innovations from being copied by others, thus monopolizing any profits from the innovations. The lower the value, the more vulnerable designs are to be copied, and thus the smaller the profits can be reaped from the innovations.

The selected attribute characteristics in this study are available soon after a patent is granted. We do not choose the characteristics that take time to be triggered or accumulated. It is obvious that young patents often show extremely low values or zeros because their lifetimes are as yet short.

Following Wang et al. [34], we also extracted a wider assortment of undirected node-level network measures. We examined the first assignee country of each patent as well as the corresponding references' assignee countries. Each country citation network is then derived from the entire set of patents in a snapshot where nodes represent individual countries and weights of ties between two nodes represent the number of direct citation involvement. For example, a link with the form, 'Country A–Country B' means that country A's patents have been cited by country B's patents and vice verse, and the weight of the tie represents the total number of these citations. The calculations range from local, semi-local to global levels that correspond to the following indices:

- Degree centrality (DC) refers to the number of ties attached to a given node [35]. In order to attain the standardized score, each score should be divided by the number of nodes minus one. Nodes that have more ties to/from others may be advantaged positions. This measure is considered local because its calculation only involves characteristics of the focal node itself. It is very simple, but often very effective to measure a node's centrality.
- Clustering coefficient (CC), sometimes called local cluster coefficient, measures the probability that the neighborhoods of a node are connected. It is the ratio of ties between the nodes within its neighborhood divided by the number of ties that could possibly exist between them [36]. This measure is considered semi-local because its calculation only involves the activity of a node's neighbors. A node with higher CC means the neighbors of a node are closer and may share knowledge sources.
- Eigenvector centrality (EC) is one of the node metrics that characterize the global prominence of a node because it is calculated using properties of the entire network. The gist is to compute the centrality of a node as a function of the centralities of its own neighbors [37]. If we rank nodes by their EC, we can see which one is important in the network.

2.2.3. The back-propagation neural network model

We suggest using a BPNN model for mapping relationship of the characteristics related to the transitions of those leading countries. In proposed model, the input is represented by the normalized rate of changes of eight characteristics, TCT, SL, PD, OI, TD, DC, CC, and EC while the output is represented by three types of transition pattern, stable, appearing, and exiting, of leading countries. An artificial neural network is a computing model whose structure mimics the knowledge acquisition and organization of the human brain in essence. This neural network is suitable for our study because of its capacity for handling characteristics which requires no assumption about data statistical distribution. It can also estimate the non-linear effects between inputs and outputs and approximate any continuous function if there are sufficient hidden neurons available [11,12]. The neural network is composed of an input layer, one or more hidden layer(s), and an output layer. The benefit of hidden layer neurons is their ability to develop internal representation of the input–output mapping. The complex internal representation capability allows a network to learn any mapping and not just the linearly separable ones. Each layer consists of multiple neurons that are connected to neurons in the next layer via synaptic weights. While a neural network does not have to be adaptive itself, its practical use comes with the back-propagation learning algorithm to alter the weight strength. When working with the back-propagation algorithm, it is important to include a bias term in the input layer and the hidden layers. The bias term lies in one layer, which connected to all the neurons in the next layer, but none is in the previous layer and it always emits 1. Without a bias term, back-propagation is restricted to find a hyper-plane that intersects the origin, so a good output pattern may not be found. Learning of the neural network is carried out basically through the presentation of a series of vectors associated with input and desired output pairs. Each neuron in a hidden and output layer processes its inputs by multiplying the weight, and summing the products, followed by passing the sum through an S-shaped sigmoid transfer function to generate a result. The learning procedure is activated by adjusting the weights using the gradient descent method to minimize the average sum squared error between actual and desired outputs. Because a thorough treatment of a BPNN is beyond the scope of this study, the basic mathematical concepts can be found in the literature [38]. The important parameters of BPNN are introduced as follows [39–41]:

- *Transfer function*: The purpose of redefining the transfer function was to permit intermediate network layers to produce visible influence at the output layer and to provide nonlinearities between in the input and output signals. The non-linear differentiable functions such as sigmoid and hyperbolic tangent are commonly used.
- *Learning rate*: The back-propagation algorithm requires that the synaptic weight changes be proportional to the derivative of the error. The larger the learning rate is, the larger the weight will be changed on each iteration, and the quicker the network learns. However, the size of the learning rate could also affect whether the network achieves a stable solution.
- *Momentum*: The concept of momentum is that previous changes in the synaptic weights should affect the current direction of movement in weight space. With momentum, once the weights start moving in a particular direction in weight space, they tend to continue moving in that direction.
- *Iteration*: In the employment of the back-propagation algorithm, each iteration of learning involves the following steps: feed forward training inputs, calculate errors, compute differences, propagate error backwards, and update synaptic weights. The changed weights are then implemented throughout the network, the next iteration begins, and the entire procedure is repeated using the next input.

2.2.4. Characteristics-transitions analysis

After learning the final synaptic weights, weights could be partitioned to determine Garson's relative importance of the characteristics of each transition pattern [28]. Furthermore, for each characteristic, the one-sample right-tailed and left-tailed *t*-tests are employed to investigate the mean value of each transition pattern. Those tests realize the directionality of mean values which is either not significantly different from zero, more than zero, or less than zero. If more than two transition patterns of a characteristic indicate similar directionality, a series of two-sample right-tailed or left-tailed *t*-tests are employed to detect whether one mean is larger or smaller than the other. Levene test is applied to assess variance homogeneity. If the *p*-value is less than 0.05, it concludes that the variances are unequal and should use the *t*-test based on unequal variance, otherwise it uses the *t*-test based on a pooled estimate of the variances. The procedure of the tests is described by pseudo code as below:

```
FOR each characteristic
  r = 0;
  l = 0;
  FOR each transition pattern
    Imply one-sample right-tailed t-test
    IF p-value ⩽ 0.05
      r = r + 1;
    ENDIF
    Imply one-sample left-tailed t-test
    IF p-value ⩽ 0.05
      l = l + 1;
    ENDIF
  ENDFOR
  IF r ⩾ 2 OR l ⩾ 2
    Imply Levene test
    IF p-value ⩽ 0.05
      Imply two-samples right- or left-tailed t-tests based on
      unequal variance for mutual comparison
    ELSE
      Imply two-samples right- or left-tailed t-tests based on
      equal variance for mutual comparison
    ENDIF
  ENDIF
ENDFOR
```

3. Results and discussion

3.1. Case profile

To demonstrate the feasibility of the research methodology, fuel cell field was chosen as a case. In addition to electricity, fuel cells only produce water, heat, and a very small amount of nitrogen dioxide and other emissions, so such clear energy has been selected as one of the most promising energy alternatives. This has raised great expectations concerning environmental protection, energy saving, and reduced carbon emissions. The rapid development of fuel cell field is attracting attention from both inventors and funding bodies. Patents were treated as the document source for analysis in this study because companies could monitor the development of technology and evaluate the position of peers in the market through the patent analysis [42]. The search rules of collecting patents in this study are based on USPCs. Our focused technology was represented by parts of the current USPC classes 48/61 (Generators), 252/182.1 (Having utility as a reactive material in an electrochemical cell), 429/400 (Fuel cell, subcombination thereof, or method of making or operating), 502/101 (Making catalytic electrode), and 524/435 (Transition metal atom). 5,799 patent documents were retrieved from the database of the United States Patent and Trademark Office (USPTO) from 1996 to 2010. In addition, both attribute data (i.e. application date, issue date, title, abstract, assignee nationality, and USPC class) and relational data (i.e. patent or literature citation and reference's assignee nationality) for the selected patents were recorded. As shown in Fig. 2, the number of yearly patents on this subject began with a linear growing pattern till 2004. Then the growing trend stopped and there has been approximately 500 patents annually.

3.2. Technology topics and leading countries over time in fuel cell field

After collecting the patents, data was analyzed using a self-programming toolkit under the 'R' environment with the igraph, tm, RWeka, stringr, openNLP, and wordnet packages (see http://cran.r-project.org/). All patents as well as bibliographic couplings were aggregated into a patent citation network where the citations with relatively strong Salton's cosine were preserved. Related patents were assembled as 130 communities through a Girvan–Newman clustering operation and the dominant communities for subsequent analysis were then identified. The problem of identifying dominant communities is transformed to generate a threshold for filtering out the communities with relatively small size. We stacked the normalized size of communities of the networks and produced a histogram. According to our experience, the shape of histogram usually follows an exponential distribution as shown in Fig. 3. The shape of the distribution is primarily related to the Girvan–Newman algorithm which potentially provides many extremely weak communities as well as few of strong ones [4]. - Rosin's threshold was used in our method to automatically determine a corner point of the exponential histogram for preserving dominant communities [43]. In this study, the corner point of

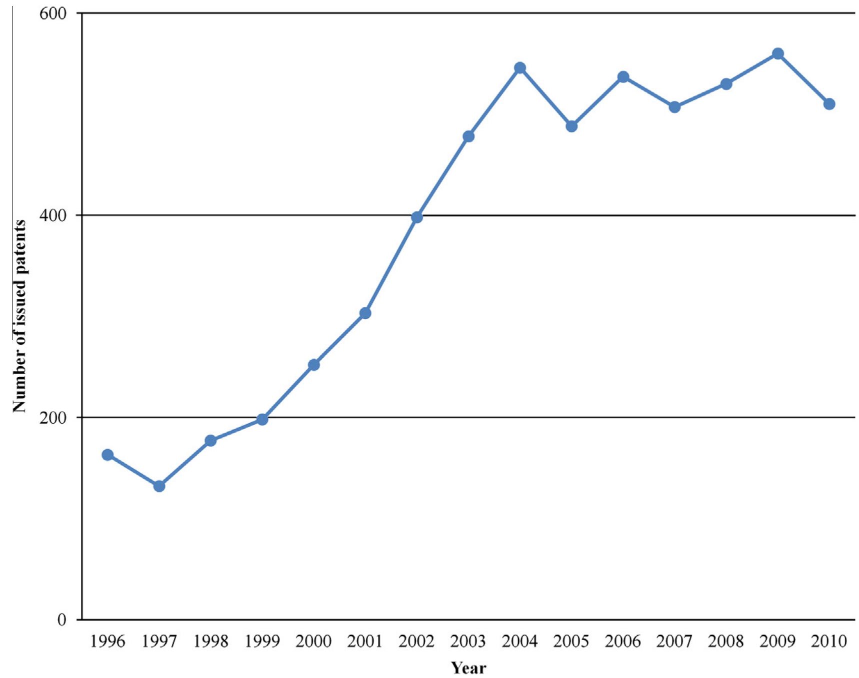

Fig. 2. Chronological development of the number of granted patents concerning fuel cell field.

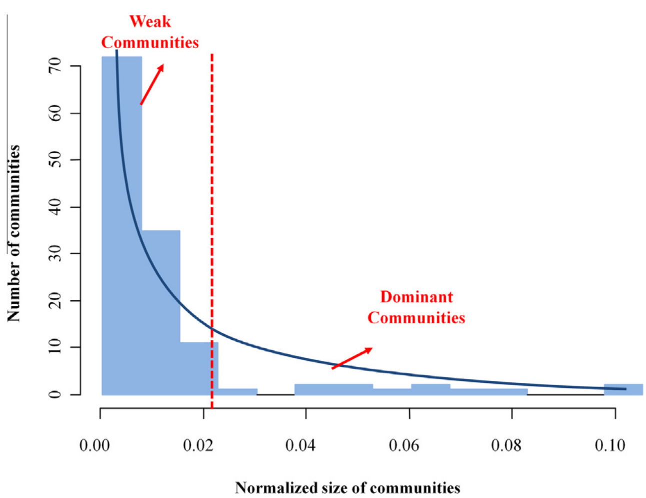

Fig. 3. Histogram of normalized size of communities.

histogram is marked by a dotted line as shown in Fig. 3, which means that the threshold was set to 0.023. Only the first twelve communities in normalized size as well as with larger size than the threshold would be discussed. After community structure was obtained, patents in a given community were divided into 3-year non-overlapping snapshots and then a post-assignment of concise and descriptive names were required in order to help analysts to interpret the results. This kind of problem was often solved by human experts where community names were given manually. However, due to the fact that it is often time-consuming in the current information-flooded era, it would be desirable to further suggest generic topic terms for ease of naming a process. As shown in Section 2.1.2, the automatic method of extracting the key terms for each cluster by natural language processing would be of a great help. Since multiple key terms are more capable of representing the precise view of technology topics than a single one, a set of key terms were aggregated into subjects manually over time as shown in Table 1. Basically, topics in each community interact all the time and do not appear definitely following the sequence of Table 1. We only engaged with pointing out the typical and focused topic of each snapshot. In addition to technology topics, leading countries of each snapshot were also identified. A two-dimensional scheme which was described in the Section 2.1.3 allowed us to identify the most prolific countries whose patents are also highly cited by their peers. They are also appended in Table 1. In order to avoid verbosity, the next section focuses on, in great detail, the evolutions of the top six communities.

Community C01 is composed of various universal management utilities in a great diversity of fuel cells, which help create a more efficient environment in extracting energy from fuel. From 1996 to 1998, maintaining water balance in the cells had been the most popular issue which required maintaining optimal conditions in the anode and cathode side. Since then, thermal management has received a lot of attentions because the heat of the fuel cell is non-reversible and nearly half of the energy dissipation is heat. The heat load is so high that it demands more technologies on the design of a temperature management system. Then the effects of changing the temperature and water content on the humidity of a water stream within a fuel cell are interesting. Balancing humidity in a fuel cell is critical to the efficient operation. Since 2005, gas management has come to light. Inefficient removal of exhausts may block anode channels and decrease efficiency of the fuel cells due to the confined mass transport and reactant maldistribution [44]. Recently, a set of hydrogen, coolant or gas purification systems have been designed and developed. Among those topics, the leading country of the United States or Japan is still intact; Germany and Canada have been popular for several periods but now they lose their positions; and the United Kingdom has emerged recently.

Community C02 is related to the flow field plate in a polymer electrolyte membrane fuel cell (PEMFC) that incorporates the functions of flow fields and bipolar plates into one unit. Flow fields are used to supply and distribute the fuel and the oxidant to the anode and the cathode electrocatalyst respectively. A bipolar plate not only makes an electrical connection with stacks but also separates

Table 1
Chronological development of thematic topics and the corresponding leading countries concerning fuel cell field.

Community: category (size, age)	Items	1996–1998	1999–2001	2002–2004	2005–2007	2008–2010
C01: Management, (500, 5.49)	Topics	Water management	Thermal management	Humidity management	Gas management	Purification methods
	Leading countries	US, DE, JP	US, JP, DE, CA, CH	US, JP, DE, CA	JP, US, CA	JP, US, UK
C02: PEMFC, (497, 5.66)	Topics	Assembly of manifold and bipolar plant	Fluid management	Flow field channels	FFP geometrics	Bipolar plant enhancements
	Leading countries	CA, US, DE, IT	US, JP	US, JP, CA,	US, JP	US, JP, DE
C03: DMFC/ PEMFC, (372, 6.02)	Topics	Methods for producing membrane electrode assembly	Electrode structures or materials	Electrode structures or materials	Gas diffusion electrode	Electrode catalyst layer
	Leading countries	US, JP, UK	US, JP, DE, CA	US, JP	US, JP	JP, US
C04: SOFC, (328, 5.37)	Topics	Producing and structure of components	Interconnector	Improved electrolyte	Different cell types	Appended functions
	Leading countries	US, JP	US, DE	US	US, JP, CA	US, JP, DE, KR
C05: PEMFC, (315, 4.86)	Topics	Processes for generating electrical energy	Carbon dioxide concentration	Sealing assembly	Stating, operating, or protecting a fuel cell under freezing conditions	Stating, operating, or protecting a fuel cell under freezing conditions
	Leading countries	US	US, JP, DE	US, CA	US, JP	US, JP
C06 PEMFC/DMFC, (290, 4.22)	Topics	Polymer membrane materials	Polymer membrane materials and processes	Polymer membrane forms	Polymer membrane materials and appended functions	Thin-film composite membranes
	Leading countries	US, JP, CA	US	JP, US	JP, US	JP, KR, US, DE
C07 DMFC, (282, 4.57)	Topics	Catalyst materials	Catalyst materials	Methanol concentration	Fuel cartridge and fuel container	Fuel cartridge
	Leading countries	US	US	US	US, JP	US, JP, KR
C08 MCFC, (233, 5.56)	Topics	Catalyst for high temperature	Power generation	Vapor interception	Gas leakage prevention	Gas leakage prevention
	Leading countries	US, UK	US	US, JP	US, JP	JP, US
C09 PEMFC, (223, 4.87)	Topics	Electrically interconnecting	Gasket	Separator materials	Seal methods, materials, and structures	Seal methods, materials, and structures
	Leading countries	None	US, CA	JP, US	US, JP, CA	JP
C10 AFC (204, 8.59)	Topics	Positive/negative electrode active materials	Non-sintered nickel electrode	Storage capacity improvement	Electrode plate	Positive electrode active materials
	Leading countries	JP, US	JP	JP	JP	JP
C11 SOFC, (183, 5.91)	Topics	Air electrode composition	Manifold reformer	Exhaust side	Internal geometry	Electrochemical optimization
	Leading countries	US, JP	US, DE	US	US	US, JP
C12 Fuel cell stacks, (140, 5.40)	Topics	Compression assembly	Manifold arrangement	Fastening structure	Compression and cooling system	Multi-stack
	Leading countries	US	CA	US, JP	US, JP	JP, US

the anode and cathode gases [45]. This community began with the assembly of external, internal, or edge manifolds and the assembly of bipolar plates. Then the topics that fuel cells employ the integrated fluid management receive attention from inventors. Next, various designs such as mirrored serpentine, serially-link serpentine, mirco-structured, converging/diverging, non-uniform, interdigitated are proposed to enhance the flow channels. From 2005 to 2008, flow field geometries are interesting in terms of improving water management. Recently, several bipolar plant enhancements have been devised like spring seals, offset, corrosion resistant, adhesive bonds, and filled-in fine scale porosities, etc. Among those topics, the United States is continuously active; Japan has been active since 1999; Canada and Italy have been popular for several periods but now they lose their positions; and Germany has emerged again recently.

Community C03 is associated with the membrane electrode assembly of direct methanol fuel cell (DMFC) and PEMFC. In general, membrane electrode assembly consists of a proton exchange membrane, catalyst layers, diffusion layers, and sealing gaskets; some of them are disclosed in this community. This community is initialized by processes for fabricating membrane electrode assembly. Then the topics of electrode materials and structures are widely spread over two periods. Electrode materials include graphite, polymer, composite and nano-composite. Electrode structures cover layered, porous, sputter-deposited, nickel positive, film, etc. Next the preparation and performance of high effective gas diffusion electrode is everybody's concerns, because a unique gas diffusion electrode technique results in little to no leftover, therefore increasing the overall effectiveness and performance of the fuel cell. Recently, electrode catalyst layer has extensively attracted interests in this field. Inventors engaged with reducing the amount of catalyst contained in the layers, while avoiding unexpected drops in voltage of fuel cells. Among those topics, the United States and Japan have been dominant all the time; the United Kingdom, Germany, and Canada were popular in the early period only.

Community C04 is concerned with the components, functions, or stacks of solid oxide fuel cell (SOFC). This community started with the producing method as well as structure designs of electrolyte, electrode, and cells. Then in order to obtain the desired amount of electrical power, individual fuel cells are combined to form a fuel cell stack, in which case the multi-unit constructions and electrical interconnectors are the most important components. Next the inventions intend to provide a SOFC having a supported electrolyte film, roughened electrolyte interface layer, textured or ceramic electrolyte which show reliability and yield a high output power density. Since 2005, various types of SOFC have been devised such as hybrid thin film/thick film, fused zirconia-based and anode-supported flat-tubular. Lately, distinct functions are appended on cells, for instance, intrinsic energy storage, protective coating, portable power source, and temperature swing reforming, etc. to meet demands. Among those topics, the United States has been dominant all the time; and Japan, Korea and Germany emerge recently.

Community C05 concerns the internal environment of PEMFC in the process of generating electrical energy. Processes or methods for generating electrical energy in fuel cells are first involved. Because PEMFC is highly sensitive to carbon dioxide poisoning, this leads to a performance drop and ultimately limits the lifetime of fuel cells. Knowing the carbon dioxide concentration in the gas supplied to the fuel cell can potentially prevent permanent damage, reduce electrical efficiency loss, and monitor performance of the fuel preparation process. Topic then turned to the sealing assembly whose design always affects the thermal stresses. Lately, issues about operating, starting up, or shutting down fuel cells during freezing conditions or issues about protecting fuel cells from freezing are becoming popular. Among those topics, the United States has dominated all the time; Japan tends to catch up; and Germany and Canada were popular in the early stage only.

Community C06 mainly refers to the membrane materials of PEMFC and DMFC that not only separate the positive and negative electrodes but provide a conduit for ion movement between the electrodes. Consequently, there is considerable development around the world to try new membrane materials and their corresponding manufacturing methods, especially for polymer- or copolymer-based, with tailored physical and transport properties [46]. In addition to the new materials, some remarkable forms of membranes such as graft polymer, bridged polymer, block copolymer, multi-layer, and cross-linked have been proposed since 2002 and some appended functions such as low humidification, high duration, high water permeability, and impregnation have been disclosed since 2005; they all engaged with improving electrochemical fuel cell performance. Recently, thin-film composite membranes have started to show its domination in the applications of fuel cells. Among those topics, the United States has dominated all the time; Japan died away just in the second snapshot; and Korea and Germany are springing up late.

The remaining communities include components of DMFC (C07), components of molten carbonate fuel cell (MCFC) (C08), separator plants of PEMFC (C09), electrodes of alkaline fuel cell (AFC) (C10), flow fields of SOFC (C11), and the assembly of fuel cell stacks (C12). In addition to showing topic and leading country evolution overtime in Table 1, we summarized the differences of transition patterns among the leading countries in Table 2. We found that the United States and Japan are the countries which always remain stable. Except for the community of AFC which is mainly monopolized by Japan, the United States and Japan dominate over other communities. The leading role of Canada and Germany is oscillatory; their appearing and exiting patterns frequently occur in the first six communities. Note that the oscillations of appearing and exiting patterns of Germany are more recent while those of Canada are in the past. Being an emerging country, Korea has three appearing patterns in recent period. It is noteworthy that Korea and Germany emerge on the membrane design and Korea has seized the position of developing components of DMFC recently. Finally, the leading role of the remaining countries is transient.

3.3. Characteristics related to transitions of the leading countries in fuel cell field

We have explained a methodology to help answer the question about who the leading countries of a specific technology over time are. That provides a historical overview about dominators in the process of technology development. Take community C01 in Table 1 as an example. The research and development strength of the United States or Japan acted as a leader continuously, the United Kingdom emerged during 2008–2010, and Germany and Canada resigned their leading positions during 2005–2007. The fact indicates that the leading status of a country may maintain, appear, or disappear respectively. This observation inspired us to further deepen the knowledge about some of the countries' main characteristics that could explain such transition patterns. In so doing, attributive characteristics as well as topological ones of all countries in each snapshot were calculated and normalized. Three transition patterns were then identified based on the criterion of the existence or non-existence of the leading countries between successive snapshots. As a result, we had 70 stable patterns, 32 appearing patterns, and 25 exiting patterns. Each of the pattern was accompanied with a 1×9 input vector (including a bias term) and a 1×3 output vector because it was represented by the rate of change normalized values of eight characteristics and three types of transition.

Having explained how we built up the dataset in the previous paragraph, we can now focus our analysis on evaluating the impact of different characteristics on those transition patterns. Theoretically, a BPNN with a single hidden layer has been shown to be capable of providing an accurate approximation of any continuous functions if there are sufficient hidden neurons [47]. Thus the structure was composed of only an input layer, a hidden layer, and an output layer in this study. Among layers, each neuron in the former layer was fully connected to every neuron in the latter. The number of neurons in the input and output layers was set according to the actual number of characteristics used plus one (due to the bias term), i.e. nine, and the number of transition patterns, i.e. three, respectively. The number of neurons, including a bias term, in the hidden layer was set to four which was selected

Table 2
The results of leading countries' transition patterns.

Transitions	Countries							
	US	JP	CA	DE	KR	UK	CH	IT
Stable	41	25	2	2	0	0	0	0
Appearing	2	10	8	7	3	1	1	0
Exiting	3	4	9	6	0	2	1	1

by replicating the learning process several times, starting with two neurons and then increasing the number while monitoring the average sum squared error. Learning was carried out until there was no significant decrease in the error [48]. Thus, the network structure was chosen as $9 \times 4 \times 3$. The activation functions used were hyperbolic tangent for the hidden layer and sigmoid for the output layer. After the synaptic weights were randomized, the BPNN learned to classify input vectors according to given targets. The gradient descent with momentum and adaptive learning rate was used as a back-propagation training algorithm, since it provided reasonably good performance and more consistent results [49,50]. The learning stopped whenever either the maximum allowable number of training epochs had been met or the error measurement did not improve. The associated learned weights of the neurons were stored in Table 3 where the weights of the hidden-input layer and hidden-output layer connections were shown. Basically, the synaptic weights between neurons are the links between the problem and the solution. The relative contributions of the inputs to the outputs depend primarily on the magnitude and direction of the synaptic weights. Negative synaptic weights represent inhibitory effects on neurons and decrease the value of the predicted response, whereas positive ones represent excitatory effects on neurons and increase the value of the predicted response [51].

To determine the contributions of the independent variables and the way they act on the dependent ones, the relative importance index can be assessed by examining these weights of the neurons. This involves partitioning the hidden-output weights into components connected with each input neuron. The results are summarized in Fig. 4. They indicate the contribution of characteristics in the network in which SL, TCT, and PD are the three most important input characteristics, followed by EC and TD, with the remaining of less importance, especial for OI. Note that the importance of the three topological characteristics is ranked in terms of their globalization of calculation involvement in a network. As shown in Table 4, the one-sample t-test is employed to the characteristics of different transition patterns to evaluate their mean values which are either not significantly different from, more than, or less than zero. It is found that the mean values of TCT's three transition patterns are smaller, smaller, and larger than zero respectively, and those of SL's three transitions are by contraries. The remaining results are: the mean values of OI's three transition patterns are not significantly different from zero; the mean values of TD's and EC's stable pattern are larger than zero, the mean values of PD's and CC's appearing pattern are larger and smaller than zero respectively, and the mean values of DC's exiting pattern is smaller than zero. Because there are two coincident decisions of the one-sample t-tests in stable and appearing patterns of TCT and SL, the two-sample t-test is proceeded to check whether or not the means of groups are equal. With the results of variance homogeneity derived from the Levene test, the two-sample t-test based on equal variance is used to compare the mean values between stable and appearing patterns of TCT and SL. The results show that both TCT and SL's mean values of stable pattern are significantly smaller than these of appearing pattern. The results of those tests are summarized in Table 5, where columns represent transitions and rows represent the characteristics that are ordered based on the relative important index in Fig. 4.

As we can observe, SL is the most important characteristic for discriminating the transitions. The appearing countries increased the activity of referring on science knowledge mostly. Science is the critical source of knowledge when innovations new to the market are developed. The literature citations thus increase at the emerging stage because at this stage there is still much fundamental or radical novel knowledge to cite which is published predominantly in scientific papers. SL of stable countries is also in an increasing fashion due to the fact that the public science is amongst critical sources of technical knowledge for maintaining innovative activities [52]. The experiment shows that the SL level of the appearing countries is higher than that of the stable countries. On the contrary, SL of exiting countries is reverse. It probably causes countries to withdraw from a competition because they get little inspiration from the science.

TCT is the second-most important characteristic for distinguishing the types of transitions where the stable countries decreased the time lags between the grant of citing patents and that of the cited ones. In order to maintain competitive advantages and react to others' own innovations as quickly as possible, stable countries as well as appearing ones relatively strive for gaining current knowledge. Although some of the older basic patents are still cited, the median function is able to smooth away parts of long time lags. The result shows that the stable countries became more fast-moving than appearing countries because their TCT decreased more. Nevertheless, TCT of exiting countries is reverse. This is probably caused by the reason that exiting countries cannot catch up with the core techniques of the latest prior arts.

The characteristic of PD comes third. This characteristic only relates to the appearing countries which present an increasing manner. In fact, at the beginning of the technology development the applicants tend to formulate a broad claim in order to reduce the chances for future applicants, which causes a long duration of the examination process [53].

The global, semi-local, and local topological characteristics are ranked fourth, sixth, and seventh respectively. The rank of their relative importance indices are based on the globalization of calculation involvement in a network and each of them is related to the stable, appearing, and exiting transitions. The value of EC increases so that stable countries maintain the global leadership position among peers. The value of CC goes down due to the fact that the appearing countries suddenly emerge in a certain position of the network where their citing/cited behaviors among their neighbors are still sparse. Besides, the decreasing tendency of the DC would be one of the clues that countries may secede from the leading groups.

Back to the fifth characteristic, TD, that is concerned with the stable countries only. In general, TD shows the capability of preventing one's previous knowledge from being easily imitated by others. The stable countries show an increasing trend of the proportion of self-cited patents. This implies that the stable countries have been successfully accumulating capabilities to keep reaping the benefit of their own preceding innovations [29].

Finally, OI is almost the last place among eight participating characteristics. The rate of change of OI of each transition almost

Table 3
Summary of final synaptic weights.

	Hidden layer			
	Hidden 1	Hidden 2	Hidden 3	Bias
Input layer				
TCT	0.71	−6.13	−1.73	
SL	0.51	3.72	5.70	
PD	4.93	0.62	2.76	
OI	−0.13	−2.11	0.04	
TD	−4.17	−0.94	−0.71	
DC	−2.03	0.73	1.97	
CC	2.20	0.03	−3.78	
EC	−2.35	1.73	−3.61	
Bias	−2.94	8.80	−8.62	
Output layer				
Stable	−0.35	0.70	−0.60	−1.07
Appearing	0.19	0.23	0.97	0.57
Exiting	0.23	−1.12	−0.29	0.74

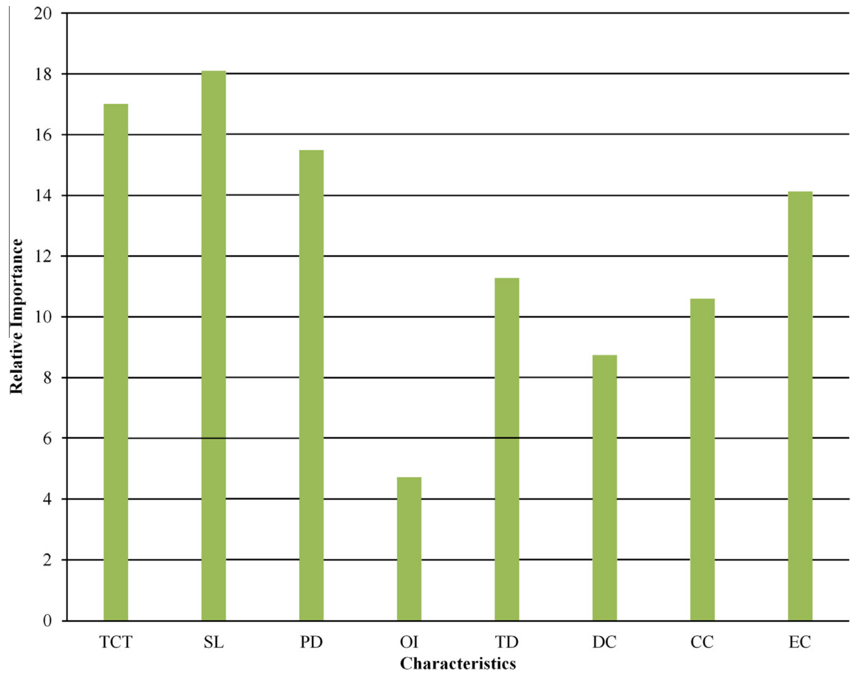

Fig. 4. Relative importance of each characteristic.

makes no difference, so we argue that we could not discriminate the transitions from the perspective of the discipline variety.

4. Conclusions and suggestions for future work

To study the evolution of the technology topics and the corresponding leading countries, we have borrowed from the social network theory and employed a bibliometric analysis to explore the issued patents of fuel cell technology in USPTO longitudinally. The basic idea was to divide a technology domain into strongly connected communities and to track their successive technology topics and leading roles in terms of non-overlapping snapshots. As described above, fuel cell field covers twelve main evolving communities in Table 1. As we observed, the combination of the individual community can be regarded as an epitome of the fuel cell family. Not only are the components of various well-known types of fuel cells such as PEMFC, DMFC, SOFC, MCFC, and AFC extensively discussed but the universal concepts of the fuel cell operation management and assembly of cell stacks are also mentioned. (In fact, because only the dominant communities with normalized sizes larger than 0.023 would be discussed, several famous topics related to phosphoric acid fuel cells and biological fuel cells are discharged.) Among these 12 communities, development of various universal management utilities is one of the dominant subjects because fuel cells thirsty for a more efficient environment in extracting energy from a fuel. In this community, the United States and Japan always lead the world while the United Kingdom rouses herself to catch up recently. Among a wide variety of fuel cells, PEMFC and DMFC have been extensively focused over the last two decades. These have emerged as one of the potential systems, which not only provide clean energy but also offer good commercial viability [7]. The components fabricated in PEMFC and DMFC are widely enclosed because fuel cell performance is essentially governed by those items. The former includes flow field, electrodes, membrane materials, separators, and components for internal environment whereas the latter contains electrodes, and membranes and catalyst materials. In these communities, the United States and Japan are always the leaders, and Germany and Korea latterly catch up from behind in some topics. Germany engages with flow fields and membrane materials while Korea is devoted to the membranes and catalyst materials. Then SOFC is also a kind of popular fuel cell. Advantages of this class of fuel cells include high efficiency, long-term stability, fuel flexibility, and low emissions. Components, functions, or stacks of SOFC are widely enclosed to pursue better performance. In these communities, the United States and Japan are always being the leaders, and Germany and Korea emerge recently. Finally, we can also observe the components of MCFC are occupied by Japan and the United States. The electrodes of AFC are monopolized by Japan. As governments increase the R&D investments to strengthen their national competitive position, they need information about the international comparison. Bibliometric analysis performed at such macro-level yield at best general assessments of fields as a whole [54]. For instance, how good a country's performance is, where a country's advantage is and when a country's emergence is in a given field. The proposed methodology could potentially aid policy makers in gaining a clearer understanding of worldwide inventions, technology trends and the innovativeness of countries. Such an understanding is required to keep abreast of current trends, to select appropriate sub-domains, and to make strategic timing of development and deployment.

In accordance with our objectives we concentrated on characteristics for which significant mean value differences between transitions could be expected. We found that: SL is the most important characteristic in which the SL levels of the appearing countries increase more than those of the stable ones while those of the exiting countries decrease. In the field of fuel cells, technologies are entered or maintained where and when scientists are publishing academic articles otherwise countries could secede from the top. TCT of both stable and appearing countries is in a decreasing manner and the decrease level of the former is less than that of the latter, while exiting countries is in an increasing manner. In order to occupy a leading position in fuel cells, the stable and appearing countries fight their ways out of the fierce competition environment by imitating, improving, or enhancing the recently popular devices, methods, or apparatuses. Besides, experiments also reveal some results in fuel cell field: a long dura-

Table 4
Logs for testing of various characteristics in different transitions.

Characteristics	Tests					
	One-sample right- or left-tailed t-test			Similar direction	Levene test	Two-sample right- or left-tailed t-test based on equal variance
	Stable	Appearing	Exiting			
TCT	.997 .003***	.985 .015**	.001*** .999	Stable vs. appearing	.083*	.000*** 1.00
SL	.000*** 1.00	.000*** 1.00	.992 .008***	Stable vs. appearing	.078*	.000*** 1.00
PD	.253 .747	.000*** 1.00	.511 .489	–	–	–
OI	.488 .512	.219 .781	.362 .638	–	–	–
TD	.000*** 1.00	.677 .323	.466 .534	–	–	–
EC	.000*** 1.00	.168 .832	.898 .102	–	–	–
CC	.718 .282	.997 .003***	.728 .282	–	–	–
DC	.444 .556	.399 .601	1.00 .000***	–	–	–

A hyphen represents that there is no need to employ the two-sample t-test.
* Significant at $p < 0.1$.
** Significant at $p < 0.05$.
*** Significant at $p < 0.01$.

Table 5
Testing results of the mean of various characteristics in different transitions.

Characteristics	Transitions		
	Stable	Appearing	Exiting
SL	↑	↑↑	↓
TCT	↓↓	↓	↑
PD	–	↑	–
EC	↑	–	–
TD	↑	–	–
CC	–	↓	–
DC	–	–	↓
OI	–	–	–

A hyphen represents that the mean of the characteristic in the transition is not significantly different from zero.
A single upward arrow represents that the mean is significantly more than zero.
A single downward arrow represents that the mean is significantly less than zero.
Double upward arrows represent that the mean is more than zero and further significantly larger than the mean of the characteristic that is signed by a single upward arrow.
Double downward arrows represent that the mean is less than zero and further significantly less than the mean of the characteristic that is signed by a single downward arrow.

tion of the examination process for newcomers causes an increasing manner of the appearing countries. The defense mentality helps create an increasing manner of the stable countries so as to accumulate capabilities to ensure the benefit of preceding innovations. The global, semi-local, and local topological characteristic is ranked based on the globalization of calculation involvement in a network and each of them is related to the stable, appearing, and exiting transitions in an increasing, decreasing, and decreasing manner respectively. Finally, OI is insensitive to the transitions.

Further suggestions are as below: The tie between nodes in this study indicated a BC between patents. Our suggestion for further research is to explore whether such a relationship can be replaced by other types of citations, term similarities or their combinations. Using alternative natural clustering algorithms (e.g. Walktrap, spinglass, or label propagation algorithm) rather than the Girvan–Newman method could be explored. Using alternative neural network models (e.g. radial basis neural network, general regression neural network, or support vector regression) rather than the BPNN could also be investigated. These approaches may enable researchers to excavate other latent intelligence from different perspectives. Finally, the proposed method could support different analysis levels (country, assignee, and inventor, etc.) for customizable subjects of debate. It is also suggested that future research explore (more) characteristics related to transitions of the leading roles in other technology fields.

Acknowledgements

This research is partially supported by the National Science Council, Taiwan, under Contract No. NSC101-3113-P-002-016.

References

[1] H. Small, Tracking and predicting growth areas in science, Scientometrics 68 (3) (2006) 595–610.
[2] N. Shibata, Y. Kajikawa, Y. Takeda, K. Matsushima, Detecting emerging research fronts based on topological measures in citation networks of scientific publications, Technovation 28 (11) (2008) 758–775.
[3] V. Kandylas, S.P. Upham, L.H. Ungar, Analyzing knowledge communities using foreground and background clusters, ACM Trans. Knowl. Discov. Data 4 (2) (2010) (Article 7).
[4] S.H. Chen, M.H. Huang, D.Z. Chen, Identifying and visualizing technology evolution: a case study of smart grid technology, Technol. Forecast. Soc. Change 79 (6) (2012) 1099–1110.
[5] M.E.J. Newman, Detecting community structure in networks, Eur. Phys. J. B 38 (2) (2004) 321–330.
[6] A. Pilkington, L.L. Lee, C.K. Chan, S. Ramakrishna, Defining key inventors: a comparison of fuel cell and nanotechnology industries, Technol. Forecast. Soc. Change 76 (1) (2009) 118–127.
[7] S. Sharma, B.G. Pollet, Support materials for PEMFC and DMFC electrocatalysts—a review, J. Power Sources 208 (15) (2012) 96–119.
[8] C. Le Bas, R. Bouklia-Hassane, A. Cabagnols, Prolific Inventors: Who are they and Where do they Locate? Evidence from a Five Countries US Patenting Data Set, International Centre for Economic Research Working Paper No. 14/2010, 2010. <http://ssrn.com/abstract=1625743> (retrieved 20.03.12).
[9] A.J.C. Trappey, F.C. Hsu, C.V. Trappey, C.I. Lin, Development of a patent document classification and search platform using a back-propagation network, Expert Syst. Appl. 31 (4) (2006) 755–765.
[10] Y.G. Kim, J.H. Suh, S.C. Park, Visualization of patent analysis for emerging technology, Expert Syst. Appl. 34 (3) (2008) 1804–1812.

[11] Y.S. Chen, K.C. Chang, Using neural network to analyze the influence of the patent performance upon the market value of the US pharmaceutical companies, Scientometrics 80 (3) (2009) 637–655.
[12] Y.S. Chen, K.C. Chang, Exploring the nonlinear effects of patent citations, patent share and relative patent position on market value in the US pharmaceutical industry, Technol. Anal. Strateg. 22 (2) (2010) 153–169.
[13] A.J.C. Trappey, C.V. Trappey, C.Y. Wu, C.W. Lin, A patent quality analysis for innovative technology and product development, Adv. Eng. Inf. 26 (1) (2012) 26–34.
[14] D.Z. Chen, M.H. Huang, H.C. Hsieh, C.P. Lin, Identifying missing relevant patent citation links by using bibliographical coupling in LED illuminating technology, J. Informetr. 5 (3) (2011) 400–412.
[15] H. Small, Update on science mapping: creating large document spaces, Scientometrics 38 (2) (1997) 275–293.
[16] O. Persson, The intellectual base and research fronts of JASIS 1986–1990, J. Am. Soc. Inf. Sci. 45 (1) (1994) 31–38.
[17] G. Salton, J.M. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, New York, 1983.
[18] S.H. Chen, M.H. Huang, D.Z. Chen, S.Z. Lin, Detecting the temporal gaps of technology fronts: a case study of smart grid field, Technol. Forecast. Soc. Change 79 (9) (2012) 1705–1719.
[19] B. Jarneving, A comparison of two bibliometric methods for mapping of the research front, Scientometrics 65 (2) (2005) 245–263.
[20] W. Glänzel, H.J. Czerwon, A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level, Scientometrics 37 (2) (1996) 195–221.
[21] B. Jarneving, Bibliographic coupling and its application to research-front and other core documents, J. Informetr. 1 (4) (2007) 287–307.
[22] M.E.J. Newman, Analysis of weighted networks, Phys. Rev. E 70 (5) (2004) 056131.
[23] F. Ibekwe-SanJuan, Information science in the web era: a term-based approach to domain mapping, Proc. Am. Soc. Inf. Sci. Technol. 46 (1) (2009) 1–23.
[24] M. Mitchell, B. Santorini, M.A. Marcinkiewicz, Building a large annotated corpus of English: the Penn Treebank, Comput. Linguist. 19 (2) (1993) 313–330.
[25] K.T. Frantzi, S. Ananiadou, H. Mima, Automatic recognition of multi-word terms: the C-value/NC-value method, Int. J. Digit. Libr. 3 (2) (2000) 115–130.
[26] F. Narin, A. Breitzman, Inventive productivity, Res. Policy 24 (4) (1995) 507–519.
[27] H. Ernst, Key inventors: implications for human resource management in R&D, in: Portland Int. Conf. Manage. Eng. Technol., vol. 2, 1999, pp. 420–427.
[28] G.D. Garson, Interpreting neural-network connection weights, AI Expert 6 (4) (1991) 46–51.
[29] S.H. Joo, K. Lee, Samsung's catch-up with Sony: an analysis using US patent data, J. Asia Pac. Econ. 15 (3) (2010) 271–287.
[30] F. Narin, Patent bibliometrics, Scientometrics 30 (1) (1994) 147–155.
[31] M.P. Carpenter, F. Narin, Validation study: patent citations as indicators of science and foreign dependence, World Pat. Inf. 5 (3) (1983) 180–185.
[32] Y. Xie, D.E. Giles, A survival analysis of the approval of U.S. patent applications, Appl. Econ. 43 (11) (2011) 1375–1384.
[33] M. Trajtenberg, A. Jaffe, R. Henderson, University versus corporate patents: a window on the basicness of invention, Econ. Innovat. New Technol. 5 (2) (1997) 19–50.
[34] D.J. Wang, X. Shi, D.A. McFarland, J. Leskovec, Measurement error in network data: a re-classification, Soc. Networks (2012), http://dx.doi.org/10.1016/j.socnet.2012.01.003.
[35] L.C. Freeman, Centrality in social networks: conceptual clarification, Soc. Networks 1 (3) (1979) 215–239.
[36] D. Watts, S. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (6684) (1998) 440–442.
[37] P. Bonacich, Power and centrality: a family of measures, Am. J. Sociol. 92 (5) (1987) 1170–1182.
[38] D.E. Rumelhart, J.L. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, Cambridge, MA, 1986.
[39] J.G. Carney, P. Cunningham, The Epoch Interpretation of Learning, Technical Report, 1998. <https://www.cs.tcd.ie/publications/tech-reports/reports.98/TCD-CS-1998-09.pdf> (retrieved 05.11.12).
[40] D. McAuley, The Backpropagation Network: Learning by Example, Technical Report, 1999. <http://itee.uq.edu.au/~cogs2010/cmc/chapters/BackProp/> (retrieved 02.11.12).
[41] D. Leverington, A Basic Introduction to Feedforward Backpropagation Neural Networks, Technical Report, 2009. <http://www.webpages.ttu.edu/dleverin/neural_network/neural_networks.html> (retrieved 01.11.12).
[42] C.V. Trappey, H.Y. Wu, F. Taghaboni-Dutta, A.J.C. Trappey, Using patent data for technology forecasting: China RFID patent analysis, Adv. Eng. Inf. 25 (1) (2011) 53–64.
[43] D.B. Perng, S.H. Chen, Directional textures auto-inspection using discrete cosine transform, Int. J. Prod. Res. 49 (23) (2011) 7171–7187.
[44] V.A. Danilov, J. Lim, I. Moon, K.H. Choi, Gas management in flow field design using 3D direct methanol fuel cell model under high stoichiometric feed, Korean J. Chem. Eng. 23 (5) (2006) 753–760.
[45] K. Scott, Mass transfer in flow fields, in: W. Vielstich, H.A. Gasteiger, A. Lamm, H. Yokokawa (Eds.), Handbook of Fuel Cells-Fundamentals, Technology and Applications, vol. 1, John Wiley & Sons Ltd., England, 2010, pp. 1–26.
[46] P.N. Pintauro, R. Wycisk, Polymeric Membranes for Fuel Cells: Overview and Future Outlook, Workshop on Membrane Science, Advanced Photon Source, Argonne, Illinois, 2004. <http://www.aps.anl.gov/Future/Workshops/Membrane_Science/Presentations/Pintauro.pdf> (retrieved 25.03.12).
[47] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks 4 (2) (1991) 251–257.
[48] A.T.C. Goh, Back-propagation neural networks for modeling complex systems, Artif. Intell. Eng. 9 (3) (1995) 143–151.
[49] A. Hassan, M.S. Nabi Baksh, A.M. Shaharoun, H. Jamaluddin, Improved SPC chart pattern recognition using statistical features, Int. J. Prod. Res. 41 (7) (2003) 1587–1603.
[50] S.K. Gauri, Control chart pattern recognition using feature-based learning vector quantization, Int. J. Adv. Manuf. Technol. 48 (9–12) (2010) 1061–1073.
[51] J.D. Olden, D.A. Jackson, Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks, Ecol. Model. 154 (2002) 135–150.
[52] A. Arundel, A. Geuna, Proximity and the use of public science by innovative European firms, Econ. Innovat. New Technol. 13 (6) (2004) 559–580.
[53] R. Haupt, M. Kloyer, M. Lange, Patent indicators for the technology life cycle development, Res. Policy 36 (3) (2007) 387–398.
[54] E.C.M. Noyons, R.K. Buter, A.F.J. van Raan, U. Schmoch, T. Heinze, S. Hinze, R. Rangnow, Mapping Excellence in Science and technology across Europe: Nanoscience and Nanotechnology, Centre for Science and Technology Studies, Leiden University, The Netherlands, 2004.