Technological Forecasting & Social Change 78 (2011) 953–967

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Technology clustering based on evolutionary patterns: The case of information and communications technologies ☆

Hyoung-joo Lee [a], Sungjoo Lee [b,]*, Byungun Yoon [c]

[a] Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom
[b] Department of Industrial & Information Systems Engineering, Ajou University, San 5, Woncheon-dong, Youngtong-gu, Suwon-si, Gyeonggi-do 443-749, Republic of Korea
[c] Department of Industrial & Systems Engineering, Dongguk University-Seoul, Pil-dong, 3 ga, Joong-gu, Seoul 100-715, Republic of Korea

ARTICLE INFO

Article history:
Received 2 February 2010
Received in revised form 26 January 2011
Accepted 6 February 2011
Available online 10 March 2011

Keywords:
Technology clustering
Patent analysis
Trend analysis
Evolutionary patterns
Information and communications technology

ABSTRACT

Technology trend analysis anticipates the direction and rate of technology changes, and thus supports strategic decision-making for innovation. As technological convergence and diversification are regarded as emerging trends, it is important to compare the growth patterns of various technologies in a particular industry to help understand the industry characteristics and analyse the technology innovation process. However, despite the potential value of this approach, conventional approaches have focused on individual technologies and paid little attention to synthesising and comparing multiple technologies. We therefore propose a new approach for clustering technologies based on their growth patterns. After technologies with similar patterns are identified, the underlying factors that lead to the patterns can be analysed. For that purpose, we analysed patent data using a Hidden Markov model, followed by clustering analysis, and tested the validity of the proposed approach by applying it to the ICT industry. Our approach provides insights into the basic nature of technologies in an industry, and facilitates the analysis and forecasting of their evolution.

1. Introduction

Technology forecasting anticipates the direction and rate of technology change, so helping inform managers' decision-making processes for such issues as setting the priorities setting, allocating the resources, and reducing the risks inherent in developing technology, at both national and private levels [1]. Governments need technology forecasting to advance public agendas in the face of both increasing rates of technology change and budgetary constraints [2], while companies, facing intensive business sector competition, employ it to prioritise their R&D projects and inform their strategic alliance thinking [3].

Among the various factors that can affect an industry's future, technology development is becoming increasing important [4]. Amidst the ever-faster pace of technological innovation, the processes of technology development have undergone several changes [5], at the centre of which is the merging and overlapping of technologies [6,7], which have meant that as well as diversification, technological convergence or fusion have come to play increasingly important roles in technology development in almost every industry over the past decade [8,9]. The increased interest in cross-disciplinary technologies has resulted in the promotion of collaboration between different scientific and technological fields in the anticipation that such activities would generate breakthroughs at higher rates [10]. This focus has intensified efforts to analyse similarities, differences, and/or relationships between technologies, to increase understanding about an industry and its associated technologies, and so ultimately drive strategic technology development and efficient R&D investment.

☆ It is confirmed that this item has not been previously published nor is it currently being submitted.
* Corresponding author at: Industrial & Information Systems Engineering, Ajou University, San 5, Woncheon-dong, Paldal-gu, Suwon-si, Kyunggi-do, Republic of Korea. Tel.: +82 31 219 2419; fax: +82 31 219 1610.
E-mail addresses: imhjlee@gmail.com (H. Lee), sungjoo@ajou.ac.kr (S. Lee), postman3@dongguk.edu (B. Yoon).

0040-1625/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.techfore.2011.02.002

Several researchers have investigated the technology innovation process in industry settings, especially in high-tech industries, viewing an industry as a collection of its relevant technologies, and many have used patent analysis to provide useful data [11]. As a rich source of technical and commercial knowledge about advances in technology and innovative activities, a patent is widely seen as a good proxy measure of technology [12,13]. Its publicly availability means patent data is easily accessible for various analyses, such as longitudinal research [14], so patent analysis has a long history of application — in almost every field of technology — and numerous studies have focused on developing intelligent analysis methods [15–17].

The goal of patent analysis is, in essence, to identify the characteristics of technologies or industries. Citation analysis, one of the most mainstream techniques, measures knowledge flows between technologies and sometimes between clusters of similar technologies [18–21], while other popular approaches include the cross-impact analysis of technologies [22,23], competition analysis of their co-evolutionary processes [4], and how they can be effectively visualised to discover emerging technologies [24]. Patent information has also been used to measure the progress of technologies and to project future technological trends [25–29]. The characteristics of an industry are greatly affected by the technologies it utilises (albeit to varying degrees in different industries), so it is meaningful to investigate the growth patterns of various technologies in a particular industry to help gain insights into that industry's future and to facilitate analysis of its technology innovation processes. However, few previous patent analyses studies have tried to incorporate the growth patterns of technologies, despite the potential value of this approach. And as the focus of most existing trend analysis studies has been on individual technologies, little attention has yet been paid to synthesising and simultaneously comparing multiple technologies. Analysing the growth patterns of multiple technologies could help elucidate what industries core technologies are, so promoting further examination of their possible inter-relationships (e.g. substitutive or complementary).

This research attempts to cluster technologies according to their growth patterns, based on the assumption that the trend of patent applications in a particular field reflects the trend of technological growth in that field. First, a Hidden Markov model (HMMs have been widely used to infer sequential structures of data) is used to model technology growth patterns based on patent data [30], and technologies with similar growth patterns identified by cluster analysis from the HMM results. Using HMMs for trend analysis has several advantages: the technique models raw data to capture true dynamic behaviour, is less sensitive to noise than other techniques, and also shows representative growth patterns of technologies clustered in a group. Finally, we used our approach to investigate technological growth trends in the Information and Communications Technology (ICT) industry, and our results confirmed the suitability of the proposed approach for technology trend analysis. Once technologies with similar growth patterns are clustered, it is possible to identify technologies that are in the same positions in their life cycles, allowing their effects and their future trajectories to be better predicted. Clustering can also be used as a tool to analyse inter-technology relationships by identifying those with similar growth patterns and distinguishing emerging from declining technologies.

This paper is organised as follows: HMMs are briefly explained in Section 2; Section 3 describes our overall research process, the application of an HMM to patent data the interpretation of our results, after which Section 4 presents our case study of the ICT industry and briefly discuss the contributions and policy implications of this research. Finally, we draw conclusions in Section 5.

2. Background

2.1. Theoretical background: taxonomy of technologies

In general, two kinds of approaches can be used to group technologies — classification or clustering. In a *classification* approach, a set of criteria to compare technologies is created, and technologies are then classified based on this set, thus creating a taxonomy of technologies [30,31]. Another popular method is to develop a technology portfolio or indicators for classifying various technologies into several groups (typically four groups when two criteria are used to create a 2 by 2 matrix) [32–34]. The key to this method is to design criteria that both describe technology characteristics well, and are suitable for classification purposes. The classification approach is usually used for strategic purposes based on experts' judgement and thus differs from the approach used in this study, which is based on data analysis. In a *clustering* approach, data that describe technological characteristics are used to cluster technologies with similar characteristics. A popular data source is a patent database, given the extensive technological knowledge included in patent documents, which makes this data useful for examining how technologies behave over time, for identifying technology breakthroughs, and for analysing cross-fertilisation between technologies.

There are two main techniques that can be applied to classify technologies in the clustering approach. The first (which is one of the most frequently used techniques in innovation studies) is based on the *technological knowledge-flows* between different technologies, and usually uses patent citation and network analyses. Citation analysis measures the degree of knowledge flows, allowing the relationships between technologies to be analysed, after which these relationships are visualised by network analysis, which also identifies technologies of particular importance (e.g. in terms of knowledge providers, absorbers or intermediaries). Technologies are assigned to the same group if active knowledge flows are observed between them [35] or where they play similar roles in knowledge flows [35–38]. While this technique is quite useful — because it analyses the direct relationships between technologies — it has the limitation that its analyses are mostly static, so the results provide little insight into the dynamics of technological changes.

The second technique is based on *growth curves* [39,40]. Time-series data for each technology in a database are plotted, and the growth curve model that represents the data characteristics selected, after which the data are fitted to the growth curve to obtain parameter values. This process is then repeated for all technologies, and the parameter values used to group technologies with similar growth patterns. Though this approach is easy to perform and effective, it cannot be used if data do not fit well to a growth curve. Another drawback is that this approach needs a growth curve model to be set up in advance, while the nonparametric method suggested in this paper does not require a predefined model. In our suggested method, once data are provided, the most appropriate number of stages to model the data is estimated, and technologies are then clustered based on similarity values at each stage. Growth curve model assumptions and parameter estimation processes are therefore not necessary in this method, which means that it is possible to analyse data that are not well fitted to an existing model. In addition, as the suggested method uses a representative value for each stage, it can consider the absolute number of patent applications when analysing growth patterns. Finally, the suggested method enables users to set the number of technology clusters themselves (though an optimum number can be suggested), which may not be possible with a parameter-based approach. The approach presented in this paper is therefore expected to be especially useful when analysing an industry where a number of technologies co-exist, when a growth curve model is difficult to define in advance, and when there is insufficient information to describe growth patterns.

2.2. Methodological background: Hidden Markov models

2.2.1. Basic concept of HMMs

HMMs [41], which are special cases of a state space model with discrete states, where observations can be observed but states are *hidden*, have been widely used in areas such as bioinformatics and signal processing. Fig. 1 shows a graphical model of HMMs [42], where the blank nodes indicate *state* variables, s_t, $t=1,\dots,T$, and the filled nodes *observation* variables, x_t, $t=1,\dots,T$. Each state takes a discrete value $j\in\{1,2,\dots,J\}$, while each observation can take on either a discrete or a continuous value. The edges and their directions of the nodes indicate dependencies between the variables. For example, a current state s_t is independent of any other variables but the previous state s_{t-1} — this is called a first-order *Markov* property:

$$p\left(s_t = j \mid s_{t-1} = i, \dots, s_1 = i'\right) = p(s_t = j \mid s_{t-1} = i). \tag{1}$$

Similarly, x_t depends only on s_t. These particular properties enable an HMM to model sequential data structures, so that an HMM can be broken down into two parts:

$$\text{state transition}: \ p(s_t = j \mid s_{t-1} = i) \tag{2}$$

$$\text{observation model}: \ p(x_t \mid s_t = j). \tag{3}$$

The former defines the probability of state j at time t given state i in the previous step, while the latter defines the probability of an observation at time t given state j. Since, in this research, observations are the annual numbers of patents, the probability density function for the observation model is assumed to have a Poisson distribution, which is commonly used for the number of events in specified intervals,

$$p(x_t \mid s_t = j) = \mathrm{Po}\left(x_t; \lambda_j\right) = \frac{\lambda_j^{x_t} e^{-\lambda_j}}{x_t!}, \tag{4}$$

where λ_j is determined by state j.

2.2.2. Model building

To build an HMM, the model parameters as well as the hidden states need to be estimated. In a typical HMM, the parameters include the initial state probability $\boldsymbol{\pi}=[\pi_j]$, where $\pi_j=p(s_1=j)$, the state transition matrix $\mathbf{M}=[M_{ij}]$ where $M_{ij}=p(s_t=j|s_{t-1}=i)$, and the parameters of the Poisson distributions $\boldsymbol{\lambda}=[\lambda_j]$. Therefore, the joint likelihood of an HMM over N sequences can be written as

$$p\left(\{s_{1:T}^n, x_{1:T}^n\}_{n=1}^N \mid \boldsymbol{\theta}\right) = \prod_{n=1}^{N} \left[p(s_1^n)p(x_1^n \mid s_1^n) \prod_{t=2}^{T} p(x_t^n \mid s_t^n)p(s_t^n \mid s_{t-1}^n)\right], \tag{5}$$

where $\boldsymbol{\theta} = \{\boldsymbol{\pi}, \mathbf{M}, \boldsymbol{\lambda}\}$ and the superscript n denotes the n^{th} sequence. An HMM is typically obtained by maximising the joint likelihood using a maximum likelihood method such as the *Baum-Welch* algorithm [43], a special case of the *Expectation–*

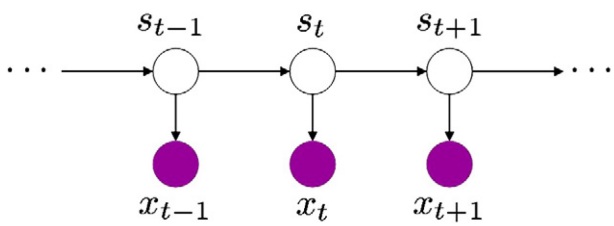

Fig. 1. AN HMM illustrated as a graphical model.

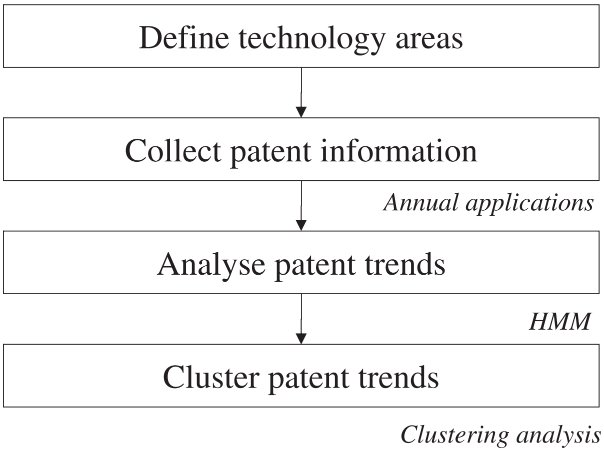

Fig. 2. The overall research process.

Maximisation (EM) algorithm [44]. In addition, the *Viterbi* algorithm [45] can be used to estimate the most probable state sequences.

Finally, we must determine the model order, i.e. the number of states. In this study, we fit a Poisson mixture model [46] to the given data set, assuming the probability distribution of the data is modelled by a linear combination of Poisson distributions:

$$p(y) = \sum_{k=1}^{K} P(k)p(y|k), \tag{6}$$

where $P(k)$ is a mixing parameter for the kth mixture component and $p(y|k) = \mathrm{Po}(y; \lambda_k)$ is the component-conditional probability density for the kth component. The Bayesian Information Criterion (BIC) [47] is then employed to find the optimal number of components, which will be used as the number of states in the HMM.

3. Methodology

3.1. Overall research process

This section describes the main idea of this paper — of clustering patent trends based on an HMM. Technology areas of interest should first be clearly defined for collecting relevant patent information, after which an HMM can be used to analyse patent trends by determining each technology's stage of progression at each time step. Finally, an agglomerative hierarchical clustering algorithm is employed to group technologies with similar growth patterns. The overall process is described in Fig. 2, and the processes of applying an HMM and analysing clusters among patent data are described in detail in the following section.

3.2. Clustering patent sequences

3.2.1. Data preparation

In this research, a sequence is defined as a series of observations over time, each of which is the annual number of patents issued in a technology field, while corresponding state variables are the stage of growth of that technology at that time. Based on the observations (the numbers of patents), an HMM is obtained with a set of estimated parameters, $\boldsymbol{\theta}$. Then we use the model to estimate the most probable state sequence, $\hat{s}^n_{1:T} = [\hat{s}^n_1, \hat{s}^n_2, \ldots, \hat{s}^n_T]$, $n = 1, \ldots, N$, i.e. which growth stage each technology is currently in. Using the HMM allows us to identify sequential technology growth over time.

3.2.2. Agglomerative hierarchical clustering

Clustering is the technique of partitioning a data set into subsets (clusters), so that the data within each are more similar to each other than are data between different clusters [48]. One of our goals in this paper is to cluster patent trends for a set of technologies. In particular, we treat each sequence of states, $\hat{s}^n_{1:T} = [\hat{s}^n_1, \hat{s}^n_2, \ldots, \hat{s}^n_T]$ as an object that represents a technology's growth pattern — based on which, those with similar growth trends are grouped into a cluster. A clustering algorithm is generally determined by three factors [49]: (1) how to measure distance or dissimilarity (or, inversely, similarity), (2) how to agglomerate (or to partition) clusters, and (3) how to determine the number of clusters.

To specify the distance between patent sequences, the usual L_2 norm or Euclidean distance measure cannot be used because a sequence consists of ordinal values. Hence, the L_1 norm or Manhattan distance measure is used [50]:

$$L_1\left(q^n_{1:T}, q^m_{1:T}\right) = \sum_{t=1}^{T} q^n_t - q^m_t. \quad (7)$$

For instance, L_1 — the distance between two sequences, 1–2–3–4–5 and 1–3–5–4–2, is $6(=|1-1|+|2-3|+|3-5|+|4-4|+|5-2|)$. With N sequences, an $(N\times N)$ distance matrix $\mathbf{D}=[D_{nm}]$ is obtained.

The data for clustering analysis in this research are not defined in Euclidean space — only the distances between them are available. Thus, rather than use the popular K-means clustering method, we use the agglomerative hierarchical clustering (AHC) algorithm. Each sequence initially forms a cluster. Pairs of smaller clusters are then recursively merged into larger ones until the number of clusters reaches a pre-specified number. The average distance of all possible combinations between elements of two clusters is calculated, and the pair of clusters whose average linkage is the smallest is selected for this merging process. (Although there are other linkage functions, they make little difference on the resultant clusters.) It is a difficult and subjective matter to determine how many clusters are optimal. The AHC algorithm produces a dendrogram (see Fig. 5), based on which we choose an appropriate number of clusters.

4. Case study: clustering patents trends for ICTs

4.1. Data

We performed a case study to illustrate how our proposed approach could be applied in practise, selecting the ICT sector for four reasons. First, it is a fast growing sector that has been and is at the forefront of industrial globalisation [51]. Second, the sector involves many competing, substitute and complementary technologies which have short life cycles and wide ranging applications across several industries. ICT technologies have also been observed as being convergent [7], all of which make the ICT sector ideally suited to our aim of studying the growth trends of different technologies. Third, ICT inventions accounts for nearly one-fifth of all

Table 1
USPC classes and ICT-related patents.

Class	Description
235	Registers
318	Electricity: motive power systems
340	Communications: electrical
341	Coded data generation or conversion
342	Communications: directive radio wave systems and devices (e.g., radar, radio navigation)
343	Communications: radio wave antennas
345	Computer graphics processing and selective visual display systems
348	Television
349	Liquid crystal cells, elements and systems
353	Optics: image projectors
361	Electricity: electrical systems and devices
365	Static information storage and retrieval
367	Communications, electrical: acoustic wave systems and devices
370	Multiplex communications
375	Pulse or digital communications
379	Telephonic communications
381	Electrical audio signal processing systems and devices
382	Image analysis
386	Television signal processing for dynamic recording or reproducing
438	Semiconductor device manufacturing: process
455	Telecommunications
700	Data processing: generic control systems or specific applications
701	Data processing: vehicles, navigation, and relative location
705	Data processing: financial, business practise, management, or cost/price determination
706	Data processing: artificial intelligence
707	Data processing: database and file management, data structures, or document processing
708	Electrical computers: arithmetic processing and calculating
710	Electrical computers and digital data processing systems: input/output
711	Electrical computers and digital processing systems: memory
712	Electrical computers and digital processing systems: processing architectures and instruction processing (e.g. processors)
713	Electrical computers and digital processing systems: support
714	Error detection/correction and fault detection/recovery
715	Data processing: presentation processing of document, operator interface processing, and screen saver display processing
716	Data processing: design and analysis of circuit or semiconductor mask
717	Data processing: software development, installation, and management
719	Electrical computers and digital processing systems: interprogramme communication or interprocess communication (ipc)

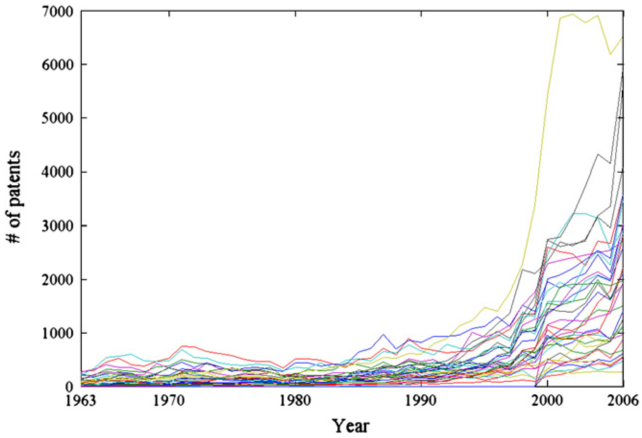

Fig. 3. The numbers of patents registered from 1963 to 2006 as broken up into 36 technology families.

patents, a fraction that is increasing [52], hence data availability is guaranteed. Fourth, ICT is a key area of research and a wide range of studies have examined its importance in the economy [53,54], its spill-over effects [14,55,56], its investment results [57–59], and the effects of national differences and/or policies [35,60], giving a wealth of other studies with which we could compare our analysis results.

We selected the United States Trademarks and Patent Office (USTPO) online database, which contains a large number of patents from all over the world, as our primary data source, a database that has not only been used for trend, diffusion, and interaction analyses of individual technologies [22,35], but also for comparative analysis between nations [61]. Before analysing the ICT sector patent data, we defined the sector itself and ICT-related patents using the United States Patent Classification (USPC) system, a widely used patent classification method: based on previous studies [4,22,35], we were able to identify 36 USPC classes of ICT-related patents, as listed in Table 1. In this research, technologies were defined by the classes and thus 36 technologies were identified.

The annual numbers of patents registered in each class between 1963 and 2007 were collected. Fig. 3 shows that the number of patents increased regardless of the technology type. This increase indicates that technology development in the ICT sector expanded over that period, although the increased interest in intellectual property rights and drive to legally protect technologies by registering them at the USPTO may also have contributed to this rise in the numbers of patents. Although the general tendency was one of growth, the absolute numbers of patents and the degrees of growth varied according to the technology: some have been growing dramatically, while others appear to have reached their peak and have stabilised or even be declining. We argue that our suggested approach is useful given the difficult of classifying technologies by their growth patterns.

4.2. Building an HMM

We developed an HMM to analyse the growth patterns of ICT technologies, using the following patent data characteristics:

- The numbers of patents registered each year in 36 (N) classes of technologies
- During a 44-year (T) period from 1963 to 2006
- 1-dimensional sequential data for each technology

A Poisson mixture model was fitted to the data so as to determine the optimum number of states for our HMM (corresponding to the number of technology development stages), and the BIC indicated that technology growth in the ICT sector could be best explained by eight steps of growth. So we set the number of sates in out HMM to eight, meaning that the growths of the 36 technologies were presented by combinations of the eight steps. Note that the states were sorted so that each successive state

Table 2
The estimated initial state probabilities.

$s_1 = j$	$p(s_1)$
1	0.5833
2	0.3344
3	0.0823
4	0
5	0
6	0
7	0
8	0

Table 3
The estimated transition matrix.

$S_{t+1}=j$ / $S_t=i$	1	2	3	4	5	6	7	8
1	0.9124	0.0767	**0.0082**	0	**<u>0.0027</u>**	0	0	0
2	0.0223	0.8670	0.1107	0	0	0	0	0
3	0	0.0724	0.8201	0.1074	0	0	0	0
4	0	0	0.0433	0.7429	0.2138	0	0	0
5	0	0	0	0.0617	0.6546	0.2590	**0.0247**	0
6	0	0	0	0	0.0542	0.7091	0.2367	0
7	0	0	0	0	0	0.0396	0.8449	0.1155
8	0	0	0	0	0	0	0	1

The figures in bold: relatively rapid growth transitions.
The figure underlined: the most rapid growth transitions.

included increasing numbers of patents (more in state 2 than in state 1, more in state 3 than in state 2, and so on) The order of the states can be considered to correspond to different stages of technology progression.

Tables 2–4 illustrate the estimated initial state probabilities — the percentage of technologies starting at each state, the transition matrix — the percentage of technologies moving from state i to state j, and the Poisson parameters — the expected number of patents in each state. The great majority of technologies start at State 1 or 2 at the initial time step, although 8.34% of them show rapid early growth, starting at State 3, but none at State 4 or higher. While ICT industries are characterised as rapidly developing sectors, and almost half the technologies involved recorded many patents in their early development stages, all were in their early stages in 1963 (State 1), producing only small numbers of patents (see Table 2).

State transitions, which are the advance to the next stage, are mostly one-way, with technologies moving from lower to higher stages, as shown in Table 3. As shown in Fig. 1, the growth rates for most technologies have increased, and are either still increasing or have stabilised: technological growth tends to progress rather than to retrogress. Such upward transitions usually occur one step at a time (from one state to the next): technologies only very unusually to return to earlier states. Thus 65.46% of technologies at Stage 5 maintained their growth rate (i.e. remained at Stage 5), while 25.9% exhibited steady growth, moving from Stage 5 to Stage 6. More unusual patterns were rapid growth (from Stage 5 to Stage 7 — a probability of 2.47%) and decrease in growth rate (probability 6.17%). (Such transits to lower states may indicate technology maturity and/or market saturation.) Other relatively rapid growth transitions were observed from State 1 to State 3 (probability 0.82%), but really big jumps were recorded only from State 1 to State 5 (probability 0.27%). Nevertheless, the overall trends suggest that ICT technologies progress gradually rather than abruptly. It is important to identify the maturity levels of technologies to prepare for next-generation developments. It is also necessary to identify which technologies exhibit rapid growth, so that firms can keep pace with them and examine them to find new opportunities.

Table 4 presents the estimated parameters for the Poisson distributions of the eight states. Because the expected value $E[x|\lambda]$ of a Poisson distribution $Po(x;\lambda)$ is λ, each parameter is equal to the expected number of patents in each state. Thus patent numbers in a year (the eight representative states of patent applications based on HMM analysis) are expected to be 14 at State 1, 132 at State 2, followed by values of 343, 728, 1276, 2097, 3062 and 6335. While only a few patents are observed at the initial state, the growth to the next stage (14–132) is remarkable, and (on average) patents numbers almost double annually thereafter. The patterns of growth in the 36 technologies indicate that the number of patent applications varies by year and by technology type.

Fig. 4 shows an example of an observation sequence and its most probable state sequence for Technology 345 (Computer graphics processing and selective visual display systems). In Fig. 4a, the thin curve indicates the actual observation sequence (the annual number of patents) while the thick curve is the expected number of patents estimated by the HMM. Basically, the two figures are the same, except that the growth pattern on the left is defined by the number of patents (showing the actual numbers of Technology 345 patents increasing from 12 in 1963 to 4077 in 2006) while that on the right (Fig. 4b) defines the changes in the technology state (showing that it moved from State 1 to State 7 over the period), and that the expected number of annual patents increased from 14 (at State 1) to 3062 (at State 7). State 8 is not observed in this case. We can also see that the technology's growth rate increased continuously, transitioning towards higher stages without retrogression. The technology has an S-shaped growth

Table 4
The estimated observation parameter for each state.

$s_t=j$	λ_j
1	14
2	132
3	343
4	728
5	1276
6	2097
7	3062
8	6335

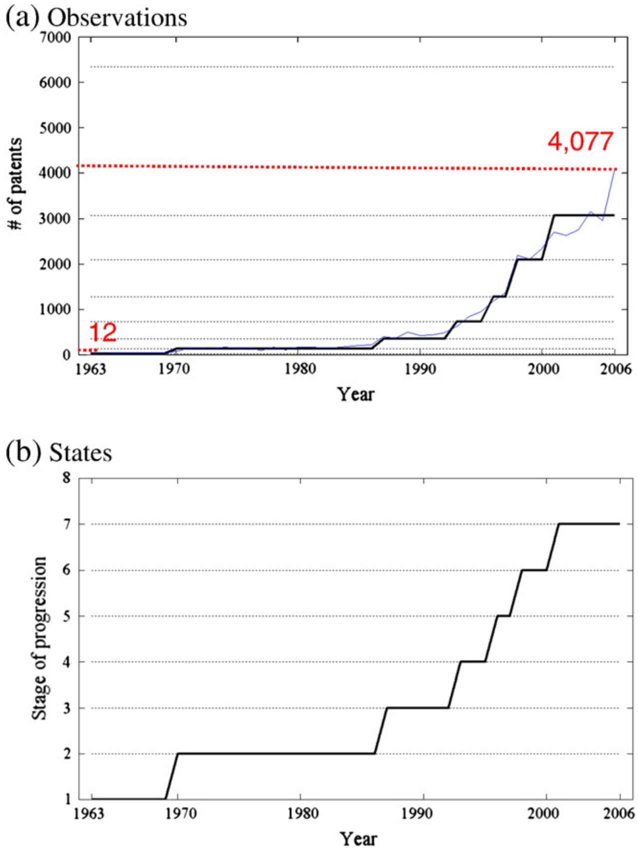

Fig. 4. The HMM result for Technology 345.

curve (before the point of inflexion in the rising phase), staying constant at Stages 1 and 2 for a relatively long time and then moving quickly from Stages 3 to 7. State 8 is not involved in the development of this technology, as the annual number of patent applications has not reached 6335, the representative value at State 8. So — from the technology life cycle perspective, which assumes that technologies evolve from emerging to pacing, to key and finally to base technologies — it seems that Technology 345 is still a pacing or becoming a key technology in its relevant industries.

4.3. Clustering patent trends

We can summarise the complicated growth patterns of 36 technologies using eight stages into a vector with 44 elements, with the value of the elements set to whichever one of the eight states the technology had reached in a particular year. Assuming that technology A started with State 1 in 1963, remained at that stage for 10 years, and then moved to successive stages every 5 years from 1973 to 2003 (from State 2 to State 7) and finally ended up at State 8, the vector that characterises the technology A's growth patterns would be as follows:

$$[1\ldots1, 2\ldots2, 3\ldots3, 4\ldots4, 5\ldots5, 6\ldots6, 7\ldots7, 8, 8, 8] \tag{8}$$

The next step is to measure the distance between two technologies by calculating the distance between their vectors, which is then used as the basis of clustering. The most probable state sequences for the 36 technologies were estimated, and a 36×36 distance matrix generated using the L_1 norm.

The distance matrix was analysed using the AHC algorithm, producing a dendrogram (as is shown in Fig. 5) which links together technologies with similar trends. For example, technologies 716 (Data processing: design and analysis of circuits or semiconductor masks) and 717 (Data processing: software development, installation, and management) have quite similar growth patterns. They both belong to the data processing areas, and are closely related to each other, possibly in a complementary way: Technology 716 is for hardware development while Technology 717 is more applicable to software development. Technology 719 (Electrical computers and digital processing systems: inter-programme communication or inter-process communication (ipc)) has a similar growth trend to technologies 716 and 717, but with a lower degree of similarity. Again, Technology 719 — which facilitates linking software systems, hardware systems or the two types together, can be seen as being complementary to

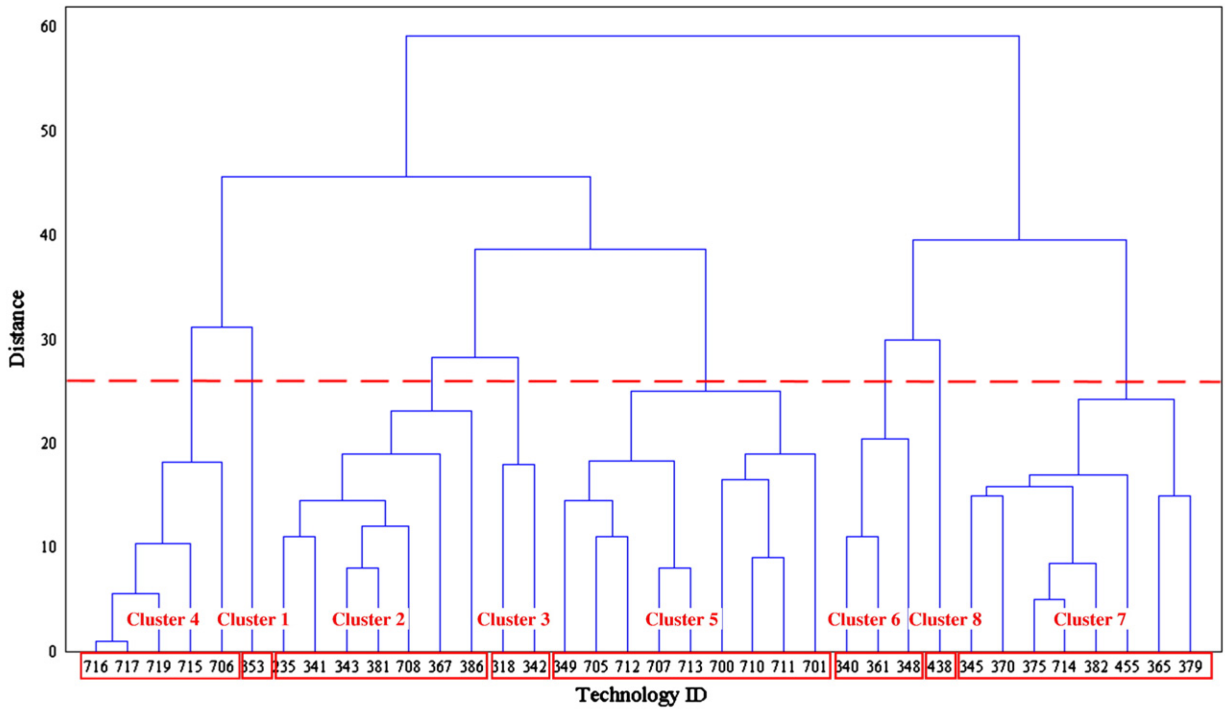

Fig. 5. The dendrogram generated by the average-linkage AHC algorithm.

technologies 716 and 717. By incorporating this diagram and our qualitative assessment, eight clusters were chosen, as indicated by the horizontal dashed line.

Fig. 6 shows the distance matrix sorted by the clusters, which are outlined in thick lines. Note that the rows and columns are arranged so that technology 'families' in the same cluster are grouped together. The lighter a cell is, the smaller the distance between its technologies. Generally, technologies in the same cluster are close to each other, and further away from technologies in different clusters. For example (reading from top-left to bottom-right), the first and last clusters, which only have one technology each, are coloured white, because the distance value is zero, and the fourth cluster has five technologies that are similar (and thus near) to each other, so the cells are all fairly light. In contrast, the fifth cluster, which involves nine technologies, is relatively heterogeneous, and has several light and dark cells.

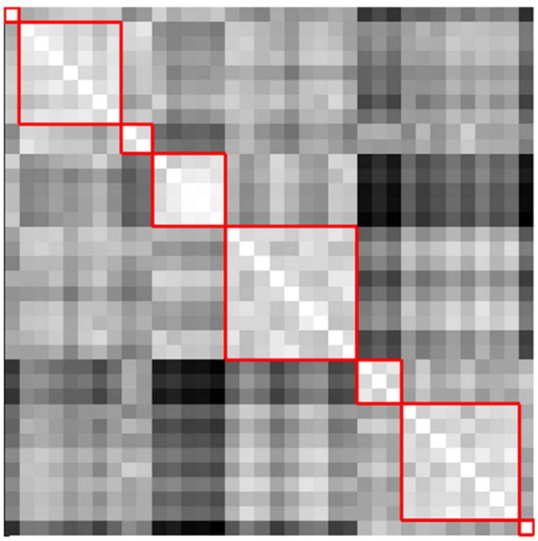

Fig. 6. The distance matrix of the data according to the eight designated clusters.

Table 5 indicates the technologies and their characteristics according to their clusters. The smallest clusters are Clusters 1 and 8 — with only one technology each — whereas the largest is Cluster 5 with nine. The characteristics of the clusters are discussed in the following section.

4.4. Clustering patents trends

Fig. 7 shows the general trends in each cluster as represented by the expected annual numbers of patents, with the graphs showing the most probable sequence of states and observations for each cluster. A thick curve denotes the expected number of patents associated with each cluster while the thin curves indicate the actual numbers of patents expected to be observed in the cluster. Based on this figure, the clusters have the following growth pattern characteristics.

Cluster 1: The only technology in Cluster 1, Optics: image projectors (353), was marginally popular in the 70s, advanced little much in the 1980s, as shown by its decreased growth rate during this period, but has been progressing since the 1990s, although only very slowly. Image projector technology seems to have developed independently from other ICT technologies, resulting in distinguished patterns of innovation, which has been incremental rather than radical, with the result that implementation techniques and application areas have changed little.

Cluster 2: The technologies in Cluster 2, Registers (235), Coded data generation or conversion (341), Communications: radio wave antennas (343), Communications, electrical: acoustic wave systems and devices (367), Electrical audio signal processing systems and devices (381), Television signal processing for dynamic recording or reproducing (386) and Electrical computers: arithmetic processing and calculating (708), have never been popular: in general, they are seen as base ICT technologies and are mainly used for data generation, storage, operation and delivery. Annual patent numbers did not change until the early 90s, since when only very slow progress has been evident. Among those with similar growth patterns, Registers (235) and Electrical computers: arithmetic processing and calculating (708) seem to be complementary, both being related to computer processing technologies. Another possible set includes the signal processing technologies Electrical audio signal processing systems and devices (381) and Television signal processing for dynamic recording or reproducing (386). Most technologies in this category were dominated by several key technologies until the 1990s, after which a small increase in patent numbers is observed. The stable growth driven by a few important technologies indicates that technology life cycle in this category is relatively long, so acquiring an original or core technology can be essential to be a category leader.

Cluster 3: Although the two technologies in Cluster 3 — Electricity: motive power systems (318) and Communications: directive radio wave systems and devices (342), both associated with motive power systems, TVs, and radios — were quite popular throughout the 70s and 80s, they have not been much sought after since then, compared to the technologies in Clusters 5, 6, 7, and 8. They have been used for many years without any significant changes in their underlying principles and (as with the Cluster 2 technologies), the acquisition of base technology in advance may be important if you want to secure long-term market presence.

Cluster 4: Cluster 4 contains relatively new technologies, for which patents only started appearing in the late 90s or early 00s: being young, they have not yet fully prospered. The five technologies in this cluster are: Data processing: artificial intelligence (706), Data processing: presentation processing of document, operator interface processing, and screen saver display processing (715), Data processing: design and analysis of circuits or semiconductor masks (716), Data processing: software development, installation, and management (717) and Electrical computers and digital processing systems: inter-programme communication or inter-process communication (719). As these technologies are related to software technologies, the trend we can observe towards these emerging technologies may be explained by a shift in R&D investments from hardware into software in the 2000s. All technologies except 719 are related to Data processing and can be complementary to each other — with more investment, they will soon reach maturity level.

Cluster 5: Little effort was spent on the technologies in Cluster 5 until the early 1990s, but they have subsequently gained popularity, meaning lots of patents, as emerging technologies. Cluster 5 has the greatest number of technologies (nine in all)

Table 5
Technologies in each cluster.

Cluster	Technology									Characteristics
1	353									Meagre
2	235	341	343	367	381	386	708			Standstill
3	318	342								Gently fluctuating
4	706	715	716	717	719					Newly emerging
5	349	700	701	705	707	710	711	712	713	Gently growing
6	340	348	361							Early popular
7	345	365	370	375	379	382	455	714		Continuously growing
8	438									Rapidly growing

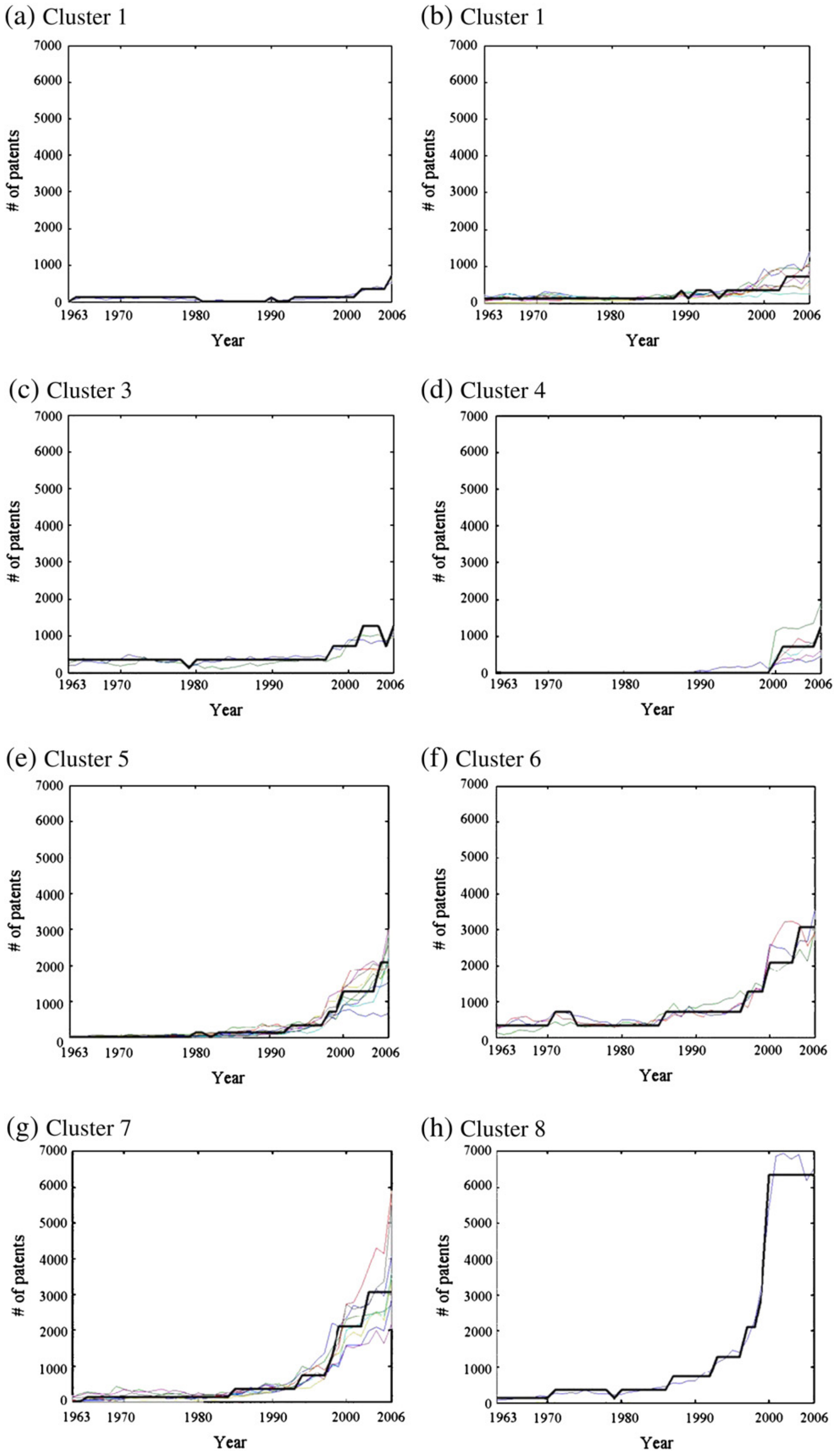

Fig. 7. The most probable sequence of states and the actual observations for each cluster.

including Liquid crystal cells, elements, and systems (349), Data processing: generic control systems or specific applications (700), Data processing: vehicles, navigation, and relative location (701), Data processing: financial, business practise, management, or cost/price determination (705), Data processing: database and file management, data structures, or document processing (707), Electrical computers and digital data processing systems: input/output (710), Electrical computers and digital processing systems: memory (711), Electrical computers and digital processing systems: processing architectures and instruction processing (712) and Electrical computers and digital processing systems: support (713). Thus, the patterns of growth of these technologies can be said to represent the general growth of the ICT sector. The numbers of patents for these technologies have grown gradually, step by step, from State 1 to State 6 since the early 1990s. Technologies such as LCDs, data processing, electronic computers and digital processes have played a leading role in the growth of the ICT sector, and are closely related to each other. Thus the performance of each affects that of the others, and investing in one technology may facilitate the development of other technologies, though their precise relationships should be examined by other methods such as citation analysis or cross-impact analysis.

Cluster 6: consists of three technologies — Communications: electrical (340), Television (348) and Electricity: electrical systems and devices (361) — that have always been popular. There was a sudden flood of patents in the early 1970s and since then numbers have increased consistently. The demand for electrical systems and televisions has been continuous with the rapid progress of technological innovation from monochrome displays to liquid crystal display and plasma display panels. Although these technologies have become more mature, they are still being developed to enhance their efficiency and performance on different electrical systems, with an increasing focus on digital and mobile televisions. In addition, early demands for these types of systems have made them available for some time. Technology life cycles in this category are relatively short with many minor innovations, so continuous investment is required to stay up to speed in the various associated industries.

Cluster 7: These technologies were in little demand until the mid 1990s, but since then the numbers of patents have increased rapidly. Cluster 7 (the second biggest) covers eight technologies: Computer graphics processing and selective visual display systems (345), Static information storage and retrieval (365), Multiplex communications (370), Pulse or digital communications (375), Telephonic communications (379), Image analysis (382), Telecommunications (455) and Error detection/correction and fault detection/recovery (714). These technologies (which can complement each other) are mainly linked to images, graphics, multiplexing, and telecommunications, and are thus becoming extremely important to meet customers' ICT device requirements, which explains their tremendous growth

Cluster 8: Cluster 8's only technology — Semiconductor device manufacturing: process (438) — was modestly popular throughout the early 1990s, but has been exploding ever since. The devices are core components of many IT devices and the expansion in semiconductor applications since the 1990s has driven remarkable growth in the manufacturing processes with. It is likely that this stand-alone technology has developed independently from the other technologies. Industries entering this sector are likely to require huge investments to keep up with such a fast growing technology.

4.5. Discussion

4.5.1. Data

The USPTO database is the main data source for this research, from which we collected the numbers of patent applications in 36 technologies over 44 years: two issues exist with respect to this process. Firstly, there is no limitation on time span in adopting the HMM approach; as in most other trend analysis research, what is more important is the amount of data collected, which can vary by data type (yearly, quarterly or monthly data) across the same time span. This research used yearly data over the 44 year span because we were interested in the long term growth patterns — but the same analysis would be possible using, say, quarterly data over 11 years. In circumstances when technological change is more rapid, collection over shorter time periods might be preferable. Second, it is worth discussing what amount of data would be appropriate. In HMM, the number of states suitable for explaining the patterns of data is determined by fitting the data to a Poisson mixture model — in this study, eight states were identified as describing the growth patterns of 36 technologies. Transforming raw data to standardised values (states values) would counter any problems of over-fitting or noise with a large dataset. On the other hand, attempting to cluster many technologies when there are only a few data available for each is likely to mean the data cannot be fitted well to a Poisson mixture model so that clustering results may not be meaningful. Therefore a dataset of a suitable size is needed to apply an HMM.

4.5.2. Contributions

This study has clustered technologies in the ICT sector based on their growth patterns. Although the methods suggested in this research may seem complex, they have been used widely — via automated processes — for a variety of other purposes. We focused on only 36 technologies, but, systematic clustering of technologies with similar growth patterns using a logical process could be more meaningful with greater numbers of technologies. As well as the possibilities of automating the process, there are other benefits of applying HMMs. The results of trying to cluster technologies from raw data may be unreliable, as it is usually noisy, making it difficult to see systematic patterns of technological growth. An HMM can represent the temporal characteristics of the data as it models them on the basis of sequential dependency, so it both captures true dynamic behaviour well and is less sensitive to noise. Our approach displays the representative growth patterns for each technology group, after clustering has been completed, and we anticipate this approach will be very useful when many technologies are grouped together, when it can be difficult to analyse and visualise common growth patterns simply by plotting the data.

We predict our approach will support the discovery of an industry's core technologies by discriminating emerging from declining technologies, and also identify the industry's growth drivers by analysing which common factors have affected technological growth. The research output will provide insights into the characteristics of industries in terms of technology growth, particularly in those where technologies play a significant role, as well as enhancing our understanding of the technology substitution process by identifying emerging and declining technologies. Of course, emerging technologies are not necessarily substitutes for declining technologies, but the analysis results can illustrate changes in the growth of key technologies at the industry level, which will help understand the evolution of industries. Again, technologies with similar growth patterns could be complementary: although our approach cannot indicate substitutive or complementary relationships between technologies, it can provide basic information for further investigation, and where such relationships are expected, data could be analysed further using Lotka–Volterra equations model or citation analysis.

4.5.3. Extensions

For our approach to provide meaningful results, well-defined industries and technologies are a prerequisite, because patent trends are affected by two factors: the rate of change in the number of patents and the absolute number of patents in a year — this latter, in particular, is determined by the definition of technologies. Careful consideration should be given to patent measurements, especially when keywords are used to define technologies instead of the USP classifications.

Although the trend clustering we performed was simple, our approach could be extended in the following ways. First, the data generated could be used to compare global and national technology trends. The factors that lead to the observed differences in technologies can be studied in terms of technology policies or strategies, which can have important implications. In a similar vein, nations or firms can be the object of analysis, for which it should be possible to cluster nations or firms by trends in their R&D activities or performances for a particular technology, enabling researchers to identify groups of nations or firms with similar patterns. Finally, using the proposed approach for marketing research would involve deriving clusters of products/services with similar demand patterns, and clustering results can reveal factors that have affected a particular demand patterns.

5. Conclusions

The convergence and divergence of technologies are regarded as emerging trends, and have therefore received particular attention in technology forecasting and trend analysis. Considering the significance of the latter — especially in terms of possible differences in technological trends within an industry — we have suggested a new approach to modelling technology trends and to clustering technologies with similar trends which we expect will enhance our understanding of technologies in an industry, particularly in terms of changes in key technologies, and provide useful information to identify relationships between technologies. Patent information was subjected to HMM analysis and clustering, and patent data from the ICT sector used to explore and validate the applicability and utility of our approach. Our study makes a meaningful contribution to the growing literature on trend analysis: in terms of methodology, it introduces the HMM as a valuable pattern recognition tool to the field of technology management, while in practical terms, it explores trends in different ICT segments, an industrial sector that is increasing rapidly in importance.

Despite these contributions, this research represents an early attempt to apply an HMM to patents, and thus has several limitations which indicate the need for further research. First, we used the USPTO database for trend analysis and ICT-related technologies were defined based on the USPC system. However, the USPC system may change based on the policy of the US Patent Office and on technological necessity, which may make keyword-based patent retrieval more appropriate to investigate technology trends. Second, because the purpose of this research was to propose a new approach, in-depth analysis of the case results was omitted from its scope. Following the trends and clustering analyses of ICT, ad-hoc analysis to identify the reasons behind each specific trend would be extremely helpful for those in charge of technology management in the ICT sector.

Acknowledgement

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (no. 2009-0089674).

References

[1] B. Yoon, Y. Park, Development of new technology forecasting algorithm: hybrid approach for morphology analysis and conjoint analysis of patent information, IEEE T. Eng. Manage. 54 (3) (2007) 588–599.
[2] D. Zhu, A.L. Porter, Automated extraction and visualization of information for technological intelligence and forecasting, Technol. Forecast. Soc. Change 69 (2002) 495–506.
[3] D. Barker, J.H.D. Smith, Technology foresight using roadmaps, Long Range Plann. 28 (2) (1995) 495–506.
[4] S. Lee, M. Kim, Y. Park, ICT co-evolution and Korean ICT strategy — an analysis based on patent data, Telecommun. Policy 33 (2009) 253–271.
[5] OECD, Technology Fusion: A Path to Innovation, The Case of Optoelectronics, Paris, 1993, pp. 7–11.
[6] F. Harianto, J. Pennings, Technological convergence and scope of organizational innovation, Res. Policy 23 (1994) 293–304.

[7] F. Hacklin, C. Marxt, F. Fahrni, Coevolutionary cycles of convergence: an extrapolation from the ICT industry, Technol. Forecast. Soc. Change 76 (2009) 723–736.
[8] H.J. No, Y. Park, Trajectory patterns of technology fusion: trend analysis and taxonomical grouping in nanobiotechnology, Technol. Forecast. Soc. Change 77 (2010) 63–75.
[9] F. Kodama, Technology Fusion and the New R&D, Harvard Bus. Rev., July–August 1992, pp. 70–78.
[10] I. Rafols, M. Meyer, How cross-disciplinary is bionanotechnology? Explorations in the specialty of molecular motors, Scientometrics 70 (3) (2007) 633–650.
[11] T.U. Daim, G. Rueda, H. Martin, P. Gerdsri, Forecasting emerging technologies: use of bibliometrics and patent analysis, Technol. Forecast. Soc. Change 73 (2006) 981–1012.
[12] B.L. Basberg, Patents and the measurement of technological change: a survey of literature, Res. Policy 16 (131) (1987) 141.
[13] H. Ernst, Patent information for strategic technology management, World Pat. Inf. 25 (2003) 233–242.
[14] J. Park, J.S.K. Shin, G.L. Sanders, Impact of international information technology transfer on national productivity, Inform. Syst. Res. 18 (2007) 86–102.
[15] J.C. Lamirel, S.A. Shehabi, M. Hoffmann, C. Francois, Intelligent Patent Analysis Through the Use of a Neural Network: Experiment of Multi-view Point Analysis with the MultiSom Model, Proceedings of the ACL-2003 Workshop on Patent Corpus Processing, 2002.
[16] Y.H. Tseng, Y.M. Wang, D.W. Juang, C.J. Lin, Text Mining for Patent Map Analysis, IACIS Pacific 2005 Conference Proceedings, 2005, pp. 755–765.
[17] B. Yoon, Y. Park, A text-mining-based patent network: analytic tool for high-technology trend, J. High Tech. Manage. Res. 15 (1) (2004) 37–50.
[18] L. Fleming, O. Sorenson, Technology as a complex adaptive system: evidence from patent data, Res. Policy 30 (2001) 1019–1039.
[19] H. Lee, Retrospective technology assessment: the impacts of computer communications on the structure of the public telephone network, Telecommun. Policy 21 (9/10) (1997) 30–51.
[20] L. Rosenkopf, A. Nekar, Beyond local search: boundary-spanning, exploration and impact in the optical disc industry, Strategic Manage. J. 22 (2001) 287–306.
[21] M. Trajtenberg, A penny for your quotes: patent citations and the value of inventions, RAND J. Econ. 21 (1990) 172–187.
[22] C. Choi, S. Kim, Y. Park, A patent-based cross impact analysis for quantitative estimation of technological impact: the case of information and communication technology, Technol. Forecast. Soc. Change 74 (2007) 1296–1314.
[23] G.H. Jeong, S.H. Kim, A qualitative cross-impact approach to find the key technology, Technol. Forecast. Soc. Change 55 (3) (1997) 203–214.
[24] Y. Kim, J. Suh, S. Park, Visualization of patent analysis for emerging technology, Expert Syst. Appl. 34 (2008) (1804) 1812.
[25] D. Archibugi, M. Pianta, Measuring technological change through patents and innovation surveys, Technovation 16 (1996) 451–458.
[26] B.L. Basberg, Patent statistics and the measurement of technological change: an assessment of the Norwegian patent data, 1840–1980, World Pat. Inf. 6 (1984) 158–164.
[27] D.-Z. Chen, H.-W. Chang, M.-H. Huang, F.-C. Fu, Core technologies and key industries in Taiwan from 1978 to 2002: a perspective from patent analysis, Scientometrics 64 (2005) 31–53.
[28] P. Gangulli, Patents and patent information in 1979 and 2004: a perspective from India, World Pat. Inf. 26 (2004) 61–62.
[29] H. Grupp, D. Lacasa, U. Schmoch, Tracing technological change over long periods in Germany in chemicals using patent statistics, Scientometrics 57 (2003) 175–195.
[30] R. Aunger, Types of technology, Technol. Forecast. Soc. Change 77 (2010) 762–782.
[31] J.M. Maqueira-Marín, S. Bruque-Camera, J. Moyáno-Fuentes, What does grid information technology really mean? Definitions, taxonomy and implications in the organizational field, Technol. Anal. Strateg. 21 (4) (2009) 491–513.
[32] T.Y. Chen, O.S. Yu, G.J.Y. Hsu, W.N. Sung, Renewable energy technology portfolio planning with scenario analysis: a case study for Taiwan, Energ. Policy 37 (8) (2009) 2900–2906.
[33] B.W. Lin, C.J. Chen, H.L. Wu, Patent portfolio diversity, technology strategy, and firm value, IEEE T. Eng. Manage. 53 (1) (2006) 17–26.
[34] S. Roth, S. Hirschberg, C. Bauer, P. Burgherr, R. Dones, T. Heck, W. Schenler, Sustainability of electricity supply technology portfolio, Ann. Nucl. Energy 36 (3) (2009) 409–416.
[35] J. Shin, Y. Park, Building the national ICT frontier: the case of Korea, Inf. Econ. Policy 19 (2007) 249–277.
[36] J.C. Guan, Y. He, Patent-bibliometric analysis on the Chinese science — technology linkages, Scientometrics 72 (3) (2007) 403–425.
[37] M. Meyer, Patent citations in a novel field of technology — what can they tell about interactions between emerging communities of science and technology? Scientometrics 48 (2) (2000) 151–178.
[38] G. Yu, M.Y. Wang, D.R. Yu, Characterizing knowledge diffusion of Nanoscience & Nanotechnology by citation analysis, Scientometrics 84 (1) (2010) 81–97.
[39] A. Sood, G.M. James, G.J. Tellis, Functional regression: a new model for predicting market penetration of new products, Market. Sci. 28 (1) (2009) 36–51.
[40] J.T.C. Teng, V. Grover, W. Guttler, Information technology innovations: general diffusion patterns and its relationships to innovation characteristics, IEEE T. Eng. Manage. 49 (1) (2002) 13–27.
[41] L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proceedings of the IEEE 77 (2) (1989) 257–285.
[42] M.I. Jordan, Learning in Graphical Models, MIT Press, 1999.
[43] L.E. Baum, An inequality and associated maximization technique in statistical estimation of probabilistic functions of Markov processes, Inequalities 3 (1972) 1–8.
[44] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc. B 39 (1) (1977) 1–38.
[45] A.J. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, IEEE T. Inform. Theory 13 (2) (1967) 260–269.
[46] G. MacLachlan, D. Peel, Finite Mixture Models, Wiley-Interscience, 2000.
[47] G.E. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (1978) 461–464.
[48] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM Comput. Surv. 31 (3) (1999) 264–323.
[49] P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Pearson Education, 2006.
[50] K. Jajuga, A clustering method based on the L1-norm, Computation. Stat. Data Anal. 5 (4) (1987) 357–371.
[51] OECD, Measuring Globalization: OECD Handbook on Economic Globalization Indicators, OECD, Paris, 2005.
[52] OECD, OECD Information Technology Outlook 2004, OECD, Paris, 2005.
[53] D.W. Jorgenson, V. Khuong, Information technology and the world economy, Scand. J. Econ. 107 (4) (2005) 631–650.
[54] C. Lee, Modeling the business value of information technology, Inform. Manage. 39 (2001) 191–210.
[55] J. Dedrick, V. Gurbaxani, K.L. Kraemer, Information technology and economic performance: a critical review of the empirical evidence, ACM Comput. Surv. (CSUR) 35 (1) (2003) 1–28.
[56] R. Gholami, X. Guo, M. Higón, S. Lee, Information and communications technology (ICT) international spillovers, IEEE T. Eng. Manage. 56 (2) (2009) 329–340.
[57] H. Kivijarvi, T. Saarinen, Investment in information systems and the financial performance of the firm, Inform. Manage. 28 (1995) 143–163.
[58] T. Mukhopadhyay, A. Barua, C.H. Kriebel, Information technologies and business value—an analytic and empirical investigation, Inform. Syst. Res. 6 (2) (1995) 3–23.
[59] K.Y. Tam, The impact of information technology investments on firm performance and evaluation: evidence from newly industrialized economies, Inform. Syst. Res. 9 (1998) 85–98.
[60] M. Pohjola, Information Technology and Economic Growth: A Cross-country Analysis in Information Technology and Economic Development, in: M. Pohjola (Ed.), Oxford University Press, New York, 2001, pp. 242–256.
[61] J. Shin, W. Lee, Y. Park, On the benchmarking method of patent-based knowledge flow structure: comparison of Korea and Taiwan with USA, Scientometrics 69 (3) (2006) 551–574.

Hyoung-joo Lee received the PhD in industrial engineering from Seoul National University, South Korea, in February 2007. He has since held research positions with University of Oxford, UK. His research interests include novelty detection, Bayesian inference, kernel learning machines and their applications to various areas.

Sungjoo Lee received the B.S. and Ph.D. degrees in technology management from Seoul National University, Seoul, Korea, in 2002 and 2007, respectively. After spending 6 months as a senior researcher at the Ubiquitous Computing Innovation Center, she moved to the UK to be working as a visiting scholar at the University of Cambridge. She is currently a faculty member at Ajou University, Suwon, Korea. Her research interests include patent analysis, technology roadmap, and service technology planning.

Byungun Yoon received the B.S., M.S., and Ph.D. degrees in technology management from Seoul National University, Seoul, Korea, in 1994, 2000, and 2005, respectively. Currently, he is an assistant professor in the Department of Industrial & Systems Engineering, Dongguk University, Seoul, Korea. Earlier, he was a visiting scholar in the Centre for Technology Management (CTM), University of Cambridge, Cambridge, UK, a senior researcher at Kangnung National University, Kangnung, Korea, and a senior consultant with LG, Korea. His research interests include patent analysis, new technology development methodology, and visualisation algorithms, enhancing technology road mapping, and product designing with data mining techniques.