

A patent time series processing component for technology intelligence by trend identification functionality

Hongshu Chen · Guangquan Zhang ·
Donghua Zhu · Jie Lu

Received: 11 December 2013 / Accepted: 13 May 2014 / Published online: 8 June 2014
© Springer-Verlag London 2014

Abstract Technology intelligence indicates the concept and applications that transform data hidden in patents or scientific literatures into technical insight for technology strategy-making support. The existing frameworks and applications of technology intelligence mainly focus on obtaining text-based knowledge with text mining components. However, what is the corresponding technological trend of the knowledge over time is seldom taken into consideration. In order to capture the hidden trend turning points and improve the framework of existing technology intelligence, this paper proposes a patent time series processing component with trend identification functionality. We use piecewise linear representation method to generate and quantify the trend of patent publication activities, then utilize the outcome to identify trend turning points and provide trend tags to the existing text mining component, thus making it possible to combine the text-based and time-based knowledge together to support technology strategy making more satisfactorily. A case study using Australia

patents (year 1983–2012) in Information and Communications Technology industry is presented to demonstrate the feasibility of the component when dealing with real-world tasks. The result shows that the new component identifies the trend reasonably well, at the same time learns valuable trend turning points in historical patent time series.

Keywords Technology intelligence · Patent time series · Piecewise linear representation · Patent analysis

1 Introduction

The increasingly rapid technological development of our society has dramatically accelerated the emergence and accumulation of data and information that represent technologies, such as patents, scientific literatures, national research and development (R&D) project records, gross domestic R&D expenditure and so forth. Among them, patent is a widely accepted and used technology indicator since 1980s, as patents contain explicit technical information and hold hidden knowledge indicating relations, status and trends of technologies themselves and related R&D activities [1, 2]. Thus, patent analysis can provide significant decision-making support for technological strategy making and planning in both public and private domains. With the purpose of obtaining more opportunities in the intense technological competition nowadays, the demand for efficient patent mining is becoming increasingly important for future technological development of organizations ranging from individual ones like companies to the multinational level like government unions [3].

As mentioned above, patents and other technology indicators are generating and accumulating in a dramatic

H. Chen · D. Zhu
School of Management and Economics, Beijing Institute of
Technology, Beijing 100081, China
e-mail: hongshu.chen@student.uts.edu.au

D. Zhu
e-mail: zhudh111@bit.edu.cn

H. Chen · G. Zhang · J. Lu (✉)
Decision Systems and e-Service Intelligence Lab, Centre for
Quantum Computation and Intelligent Systems, Faculty of
Engineering and Information Technology, University of
Technology Sydney (UTS), 15 Broadway, Ultimo,
PO Box 123, Sydney, NSW 2007, Australia
e-mail: jie.lu@uts.edu.au

G. Zhang
e-mail: guangquan.zhang@uts.edu.au

speed. In order to process and analyze a large volume of data automatically and intelligently, the concept and tools of technology intelligence are proposed to improve the traditional expert-experiences dominant decision process by using artificial intelligence techniques [4, 5]. Just like Business Intelligence to support business decision making, technology intelligence is promising to turn “data” found in patents or scientific literatures into “knowledge,” and help users survive data tsunamis and eventually, to succeed in strategy making [6]. Existing frameworks and applications of technology intelligence, such as *Techpioneer* [4], *TrendPerceptor* [7], *VantagePoint* [8] and *Aureka* [9], are mainly constructed on the fundamentals of semantic properties of patent documents or scientific literatures. Techniques such as information extraction, text clustering and classification are used to access knowledge hidden in massive documents efficiently, which makes a great progress in technology intelligence design and implementation. In summary, the theories and tools of existing technology intelligence are mainly text-based.

However, patents have both semantic properties and time-based attribute. The time-based attribute is embodied by publication/application activities of patents. These activities reflect the evolvement of the technologies in a certain technological area. By observing and analyzing the way these activities vary over time, we can learn trend of the corresponding technologies from historical data. As future technological impacts in a time period of interest can be directly assessed by using a future patent quantity count [10], the activities of patent publication/application can be represented by corresponding patent time series [11]. Nonetheless, the time-based attribute of patents is seldom taken into consideration while designing and constructing the technology intelligence. The existing design and tools, as mentioned, are mainly text-based, which have outcomes such as clusters, keywords or extractions without trend tags on them, or only with simple time labels such as year, month or season. Although methods such as bibliometrics analysis [12] and growth curves [13] can reveal the general trend of patent publication/application activities, trend turning points are still hard to get, which makes it difficult to analyze the text-based and time-based knowledge comprehensively.

In order to capture the hidden trend turning points of patent publication/application activities, at the same time improve the framework and applications of existing technology intelligence, this paper proposes a time series processing component with trend identification functionality. To emphasize the detailed modules of the new component, this research sees the text mining function as a black box. The main contributions of this study include (1) A comprehensive-analysis technology intelligence framework is proposed to process and analyze patent time series and text

data synthetically; (2) A time series processing component with piecewise linear representation (PLR) module is presented to process raw patent time series and help with producing a more reasonable time-based measure for future patent text mining; (3) Trend states and corresponding trend turning points of patent publication/application activities are first quantitatively identified while processing patent time series; (4) We perform a case study on Australia patents in Information and Communications Technology (ICT) industry, and the results show that the new component learns valuable trend turning points, when dealing with real-world tasks.

After the introduction, this paper is organized as follows: Sect. 2 introduces the background and related works of this research by discussing technology intelligence, patent time series and PLR. Section 3 presents a patent time series processing component and its detailed modules by illustrating the conceptual model and architecture of a comprehensive-analysis technology intelligence framework. Section 4 provides a case study of Australia ICT to demonstrate the feasibility of the new component. Conclusions and future studies are given in the last Section.

2 Related works and background

In this session, we illustrate the works related to our research. Previous study on technology intelligence, patent time series and PLR is reviewed.

2.1 Technology intelligence

The concept of technology intelligence was first systematically mentioned in supplier management research for understanding and monitoring of new technology worldwide [14]. Reviewing existing research, technology intelligence can be regarded as an “activity” to be conducted by a set of agents, or as a knowledge management “product” with consumers, that provides an organization with the capability to capture and deliver information in order to develop an awareness of technology threats and opportunities [5]. Compared to the traditional expert-based approaches, technology intelligence enables us to process massive data that cannot be analyzed by humans alone [7], and also, it is capable of generating knowledge by integrating resources from different sources to visualize the outcomes.

Since this concept was presented, research in technology intelligence gradually expanded from experience-based to data mining-based and became more intelligent in recent years. That is, an increasing number of researchers focused on the use of extremely powerful information technologies and a vast amount of available data that digitally provides us with technology intelligence [15]. *Techpioneer* [4] uses text

mining and morphology analysis to seek potential technology opportunities. *TrendPerceptor* [7] is designed to identify TRIZ (Russian Acronym of the Theory of Inventive Problem Solving) trends in invention concepts by using a property–function-based approach. *VantagePoint* and *Aureka* are another two intelligent systems that support users in analyzing trends or relationships by providing clustering, mapping and searching techniques [8, 9]. There are also some applications built based on bibliometric approaches [12], for instance, a visualization system called *Diva* was proposed to perform bibliometric analysis of scientific literature and patents for trend presentation [16]. In summary, the existing design and tools of technology intelligence are mainly constructed on the fundamentals of semantic properties of patent documents or scientific literatures. Time-related property of patents or scientific literature publication/application activities are seldom taken into consideration while processing data, the outcome trends/relations/changes are mainly text-based. More specifically, the results usually do not have any time tag or only with simple labels, such as year, month or season on them.

2.2 Patent time series

Patents are the ideal data source for technology intelligence study. Since a positive relationship between R&D activities and subsequent patenting activities has been found [17], the value of utilizing patent data for empirical research has been emphasized increasingly in recent years. Moreover, it is not difficult to obtain patent data from public patent offices of many different countries for academic study or commercial business purposes. This makes patent analysis a useful and convenient method to support technology R&D planning, competition analyses and analytic studies of how technologies emerge, mature and disappear [2, 18, 19]. Among patent databases from different countries, the United States Patent and Trademark Office (USPTO) database which contains all US patents from 1790 to today is mostly used for standardization reasons [20]. Here, in this research, we are focusing on the ICT industry of Australia, thus data from Australian government intellectual property department patent database is used.

As an indicator of technology changes, patent publication and application behavior can be seen as a nonlinear system that is affected by factors, such as the technology innovation and upgrading, political environment, economic situation, intellectual property rights infringement and protection and so forth. These activities can be represented by corresponding patent time series. More specifically, the quantity changes of published/applied patents within a particular industry over time, under a certain search statement such as relevant International Patent Classification (IPC), keywords, or their combination, can be presented as a vector. If time is here set at uniform

intervals, the vector we obtain from patent publication/application history can be seen as a time series. Let $P_i = \{p_{i1}, p_{i2}, \dots, p_{in}\}$ define one sequence, where i indicates the number of search statements in the target technology area, while n stands for the number of intervals, and p_{in} shows the number of patents which appear in each corresponding time interval.

2.3 Piecewise linear representation

In order to capture the development stages of patent publication activities, we need to explore the corresponding trend patterns hidden in time series. In this research, we utilized PLR approach to assist the extraction of the main trend of patent time series. PLR refers to a time series approximation which represents the original data in several compressed segments [21]. Owing to its ability to simplify a time series, PLR has been applied to time series mining in stock prediction [22, 23] and audio signal analysis [24] in recent studies. Generally, PLR refers to the approximation of a time series P , of length n , with k straight lines [25]. Given a time series $P = \{p_1, p_2, \dots, p_n\}$, the PLR of P can be described as follows [23]:

$$P_{\text{PLR}} = \{L_1(x_1, x_2, \dots, x_{t_1}), L_2(x_{t_1+1}, x_{t_1+2}, \dots, x_{t_2}), \dots, L_i(x_{t_{i-1}+1}, x_{t_{i-1}+2}, \dots, x_{t_i}), \dots, L_k(x_{t_{k-1}+1}, x_{t_{k-1}+2}, \dots, x_{t_n})\}, \quad (1)$$

Here, $L_i(x_{t_{i-1}+1}, x_{t_{i-1}+2}, \dots, x_{t_i})$ indicates the i th segment of P_{PLR} , which approximated $x_{t_{i-1}+1}, x_{t_{i-1}+2}, \dots, x_{t_i}$ to a straight line with the beginning time $t_{i-1} + 1$ and the end time t_i . There are several piecewise segmentation algorithms which appear under different names yet the implementation of theirs can be summarized into one of following three types [25]:

- *Top-down*: The time series is recursively partitioned until certain stopping criteria are met.
- *Bottom-up*: Starting from the finest possible approximation, segments are merged until certain stopping criteria are met.
- *Sliding windows*: A segment is grown until it exceeds an error bound. The process repeats with the next data point not included in the newly approximated segment.

In this paper, a bottom-up algorithm is used to approximate the patent time series into a number of straight lines. Patterns of the original data become easier to be captured after the segmentation, and segments produced by the PLR are then ready for the trend state transformation.

3 Patent time series processing component

Users of technology intelligence expect to perceive from patents not only the text-based knowledge hidden in text

data, but also what is the corresponding technological trend of this knowledge. In other words, the combination of trend in patent publication activities and knowledge hidden in corresponding patent documents, has the ability to provide decision makers with comprehensive awareness of technological advances in two different dimensions. This requirement needs to be satisfied on the fundamental of utilizing both text mining techniques and time series analysis in technology intelligence.

In this session, on the basis of previous text-based technology intelligence research, we design and construct a time series processing component, which enriches the existing framework of technology intelligence with trend identification functionality. The concept of the new framework, comprehensive-analysis technology intelligence is introduced. Then, the detailed modules of the component as well as the input and output of the system are presented and explained.

3.1 Comprehensive-analysis technology intelligence

The existing concept of technology intelligence is built fundamentally on an intersection of artificial intelligence and technology strategy making. In order to satisfy the demands of exploring knowledge from patent documents as much as possible, it then unites with text mining techniques, which help to improve the system into a text-based technology intelligence level. In this research, we apply time series analysis to the existing framework, which provides the new framework with the ability of processing patent time series data.

As shown in Fig. 1, a time series processing component, the shaded rectangle, is created to enrich functionalities of text-based technology intelligence. The new framework with the time series processing component is named as *comprehensive-analysis technology intelligence framework*, which indicates that the new system has the ability to present decision makers with a comprehensive awareness of technological advances. To sum up, comprehensive-analysis technology intelligence is an intelligent system which utilizes time series analysis and patent text mining techniques synthetically, for technology development planning and strategy-making support.

3.2 The conceptual model of comprehensive-analysis technology intelligence

After giving a brief introduction of comprehensive-analysis technology intelligence, this subsection presents and describes the conceptual model of the framework and its specific components.

The whole framework of comprehensive-analysis technology intelligence builds on the fundamental of

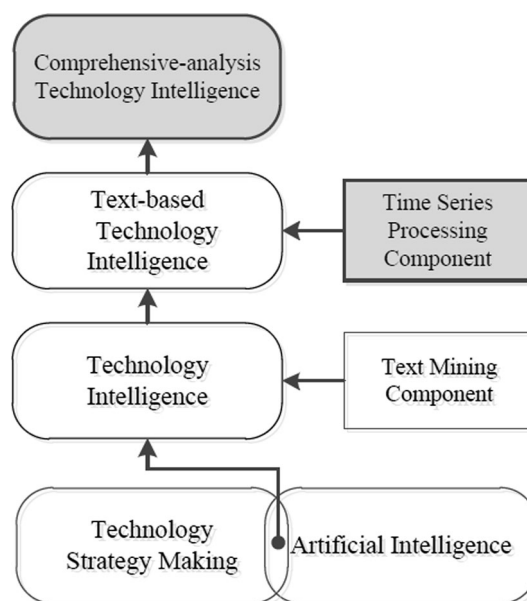


Fig. 1 Brief introduction of comprehensive-analysis technology intelligence

understanding, extracting and utilizing both time-related properties and semantic attributes of historical patent records. As shown in Fig. 2, to learn the current and historical development of target technologies, users need to first initiate their technology area of interests as system input. Here, users indicate technology R&D managers of companies and technology planning officers in government sectors. For future technical development of their organizations, they all need to assess external technological developments to determine how they can gain from technology changes, avoid potential risk and plan their future R&D activities [26]. The technology range determination of users will then be transformed into one or several patent query commands for the public patent database under expert supervision, that is, we can obtain a group of patents that conform to the query functions, such as IPC relates to the target area, selected keywords that restrict the topic of the patents and so forth. Although experts will still participate in system procedures, the effort they put is confined to supervise the input and provide appropriate advice on patent search commands. In the next step, the selected patents are processed into time series data and text data separately and transferred to time series processing component and text mining component correspondingly.

The outcome of the system integrates the results from both text mining component and time series analysis component. That is, the output includes identified trend turning points, trend states and their corresponding text-based knowledge, such as keyword-clusters, relations and topics. The patent knowledge will then be delivered to users for technology strategy support and future

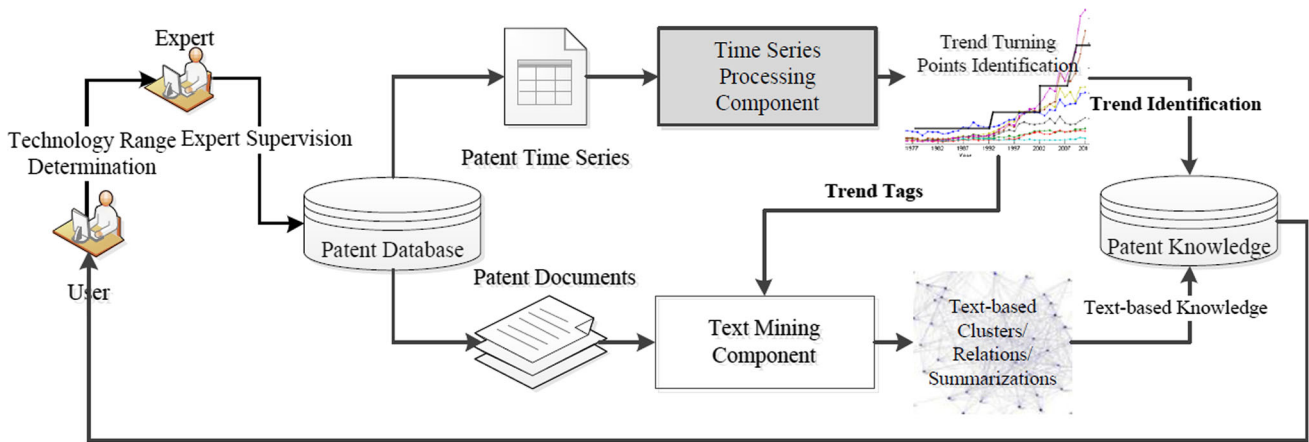


Fig. 2 The Conceptual model of comprehensive-analysis technology intelligence

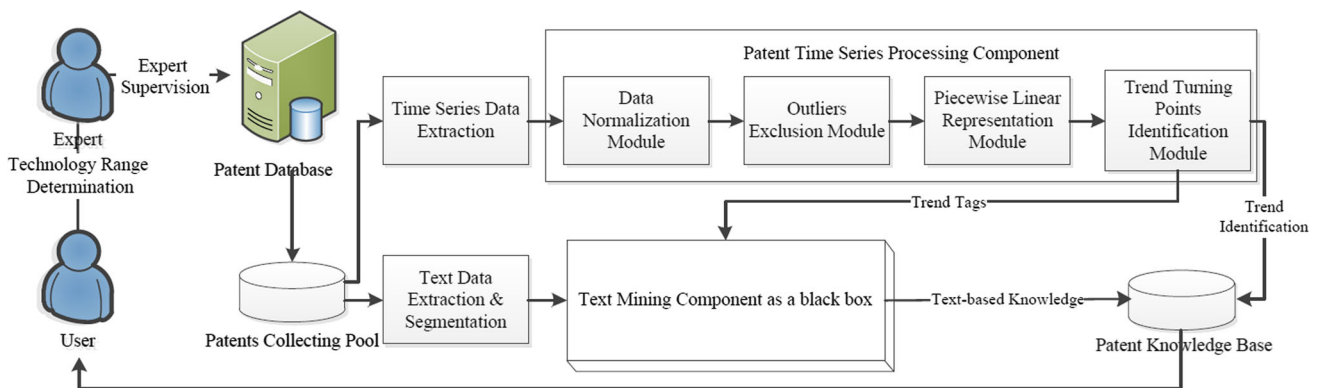


Fig. 3 The modules of patent time series processing component

developmental planning. By means of constructing the time series analysis component and utilizing it to interact with text mining component, we can finally learn the knowledge showing how the text-based clusters, relations or summarizations distributed while technological trend changing, thus provide users with more comprehensive decision assistance information.

3.3 Patent time series processing component

There are two main purposes of the patent time series analysis component: (1) identifying technology trend turning points; (2) interacting with existing text mining component. In order to emphasize the function and modules of the new component, we treat here the text mining component as a black box. Figure 3 shows the overall architecture of comprehensive-analysis technology intelligence framework; at the same time, it describes the detailed modules of the time series processing component. After the users define the technology range of their concern under expert supervision, all the patents that conform to the query statement are collected into a patent-collecting pool

and processed into patent time series data and text data separately.

The component of time series analysis receives the raw data from the processing pool and passes to the data normalization module. The normalized time series $P = \{p_1, p_2, \dots, p_n\}$ which has values between 0.0 and 1.0 is then transferred to outliers exclusion module to eliminate interference of identifying the main trend. Here, n denotes the number of uniform time intervals, and p_n indicates the patent quantity of each corresponding interval. If the data is normally distributed, then we utilize the three sigma rule on the difference of normalized data and its polynomial fit to check if there is any outlier, the pseudo code is presented in Fig. 4. After outliers removing, we obtain the prepared time series, $\tilde{P} = \{\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_n\}$. If the data does not follow a normal distribution, then the outliers exclusion module could be skipped.

The prepared time series $\tilde{P} = \{\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_n\}$ is then transferred to PLR Module (PLR module), where the data is simplified and decomposed to several segments showing the trend more obviously. The segments provided by PRL

Fig. 4 The pseudo code for outliers exclusion

```

Input: the normalized patent time series  $P = \{p_1, p_2, \dots, p_n\}$  (if normally distributed), time  $T$ 
Output: the prepared time series without outliers  $\tilde{P} = \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n\}$ 
 $P_1 = \text{polyfit}(T, P, 2)$ 
% fit the time series to a Polynomial curve
 $P_2 = \text{abs}(P_1 - P)$ 
% calculate the absolute difference between original values and the Polynomial fit
 $\text{sigma} = \text{std}(P_2)$ 
% calculate the standard deviation of the difference
 $\text{outliers} = \text{abs}(P_2 - \text{mean}(P_2)) > 3 * \text{sigma}$ 
% use the three sigma rule to identify the outliers
 $\text{position} = \text{find}(\text{outliers})$ 
% find the positions of outliers
 $P(\text{position}) = (P(\text{position} - 1) + P(\text{position} + 1))/2$ 
% use the mean value of the two data points in the front and back of the outlier to replace it
 $\tilde{P} = P$ 
end

```

module can be used to generate trend states for further assistance text-based knowledge tagging. In the segmentation, we have

$$\tilde{P}_{\text{PLR}} = \{L_1(\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_{t_1}), L_2(\tilde{P}_{t_1+1}, \tilde{P}_{t_1+2}, \dots, \tilde{P}_{t_2}), \dots, L_k(\tilde{P}_{t_{k-1}+1}, \tilde{P}_{t_{k-1}+2}, \dots, \tilde{P}_{t_k}), \dots, L_m(\tilde{P}_{t_{m-1}+1}, \tilde{P}_{t_{m-1}+2}, \dots, \tilde{P}_n)\} \quad (2)$$

where \tilde{P}_{PLR} denotes the combination of m segments and $L_k(\tilde{P}_{t_{k-1}+1}, \tilde{P}_{t_{k-1}+2}, \dots, \tilde{P}_{t_k})$ indicates the k th ($1 < k < m$) segment of \tilde{P}_{PLR} . Here, the number of segments PLR produces, m , is a threshold affecting the sensitivity of the trend identification. Users can give their preferred threshold for PLR, or it can be set as the value maintains the balance of producing the smallest number of pieces and lowest RSS (residual sum of squares), since we prefer fewer segments to show the trend more obviously. In the Trend States Identification Module, the PLR-processed data is transformed into a trend signal with mean value of each segment. This module will provide trend turning points to the text mining component and give a trend tag to the text data in each trend state. The trend tags are presented as matrix (3), where each row of the matrix indicates a start and an end of a trend state.

$$\text{Tag} = \begin{bmatrix} 1, & t_1 \\ t_1 + 1, & t_2 \\ \vdots & \vdots \\ t_{k-1} + 1, & t_k \\ \vdots & \vdots \\ t_{m-1} + 1, & n \end{bmatrix} \quad (3)$$

The outcome of the two components will be combined in a patent knowledge base, which provides users with patent knowledge that indicates: (1) what the main trend in the technological area of interest is; (2) where the states turning points of the technology developing trend are; (3) what the text-based knowledge analyzed by the text mining

component is, such as keyword-clusters/reasons/topics, on each trend state; and (4) how these text-based knowledge evolving from one trend state to another. This outcome will help users get better awareness of technology development in target area over time and provides a prospect of the future trend state in their technical areas of interest.

4 Case study

In this section, the proposed component is applied in a real patent analysis context. We collect and utilize patents from the Australian government intellectual property department patent database to demonstrate the validity of trend identification functionality, when dealing with real-world tasks. The result of the experiment shows that the new component learns valuable trend turning points in historical patent time series.

4.1 Data sets

The data we use in this research comes from the Australian government intellectual property department patent database [27]. The industry background of the patent data we choose is ICT. ICT-related technologies have attracted attention increasingly in the area of industrial globalization for their rapid growth in recent years [28, 29]. According to the Organization for Economic Cooperation and Development (OECD), how to seize benefits and opportunities of ICT for economic growth and development has become an important concern to OECD governments [30], including Australia.

The data we employ in the case study is created based on search statement for patents with IPC indicating ICT technologies, published by OECD [31], which splits the ICT sector into telecommunications, consumer electronics, computers and office machinery and other ICT. The time interval unit here is set as month. We collect the quantity of

Fig. 5 The outliers exclusion for ICT patent time series

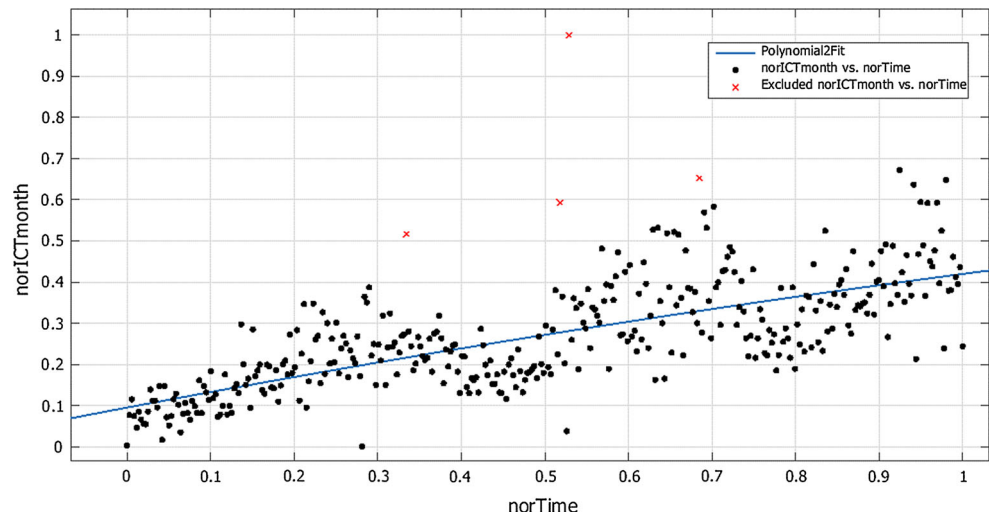


Table 1 The detailed outliers’ values of ICT patent time series

| Outliers position | Normalized time | Value of outliers | Value of outliers before normalization | Replacement value |
|-------------------|-----------------|-------------------|----------------------------------------|-------------------|
| 121 | 0.334262 | 0.516129 | 395 | 0.276999 |
| 187 | 0.518106 | 0.593268 | 450 | 0.29453 |
| 191 | 0.529248 | 1 | 740 | 0.148668 |
| 247 | 0.685237 | 0.652174 | 492 | 0.28892 |

issued patents in every month during 1983–2012 to create a patent time series, which makes 360 months in total. That is, n in $P_{case} = \{p_1, p_2, \dots, p_n\}$ equals to 360, p_i shows the number of issued patents in each corresponding month. The raw data we used in this case study follow a normal distribution.

4.2 Outliers exclusion

After data normalization, the original patent time series is converted to a new one without the outliers. As shown in Fig. 5, the ICT patent time series was fitted to a quadratic polynomial. By using the three sigma rule, we locate the position of outliers showing as red points and replace them with the mean value of the data points in the front and at the back of each outlier for trend maintaining. The detailed values of the outliers and replacement are showing in Table 1.

4.3 Trend states and trend turning points identification

After excluding the outliers, the prepared time series is processed and decomposed by PLR module. In this case study, we choose $m = 9$ as PLR threshold, as it maintains the relative balance of least segments and lowest RSS (RSS will reduce while m rising). At the same time, we use the mean value of each segment to generate a new series

showing trend states. As shown in Fig. 6, the original data is presented with blue line and represented as nine straight red lines by PLR to retain the main tendency. The final trend states are illustrated by the green line in the figure. We can observe that the trend changing points are July 1991, January 1998, March 1999, November 2011, September 2003, May 2006, September 2010 and May 2011. The detailed trend signal value and trend tags are showing in Table 2. The tags will be provided to text mining component, thus the text-based knowledge on each trend state can be learned. The corresponding text-based knowledge variation of the trend changing can be identified as well.

On the whole, the ICT development in Australia experienced fluctuant trend rise during the past 30 years. The growth trend descended slightly twice between year 1991–1998 and year 2003–2010. In the recent three years, the development of ICT is fluctuating and descending from 2011–2012, yet the main trend is escalating compares with previous years.

5 Conclusion and future work

With technological advances and the accumulation of intellectual property, technology intelligence design and construction will continue to be emphasized for its ability to assist decision makers with learning knowledge in massive data efficiently. In previous studies, the employment of text mining techniques makes a great progress in technology intelligence research and applications. However, users of technology intelligence expect not only to perceive the text-based knowledge, but also need to identify what is the corresponding technological trend of this knowledge over time. In order to provide decision maker with comprehensive awareness of technological advances in both text-based and time-based dimensions, this study

Fig. 6 Original data, PLR segmentation and trend signal of technologies in ICT industry

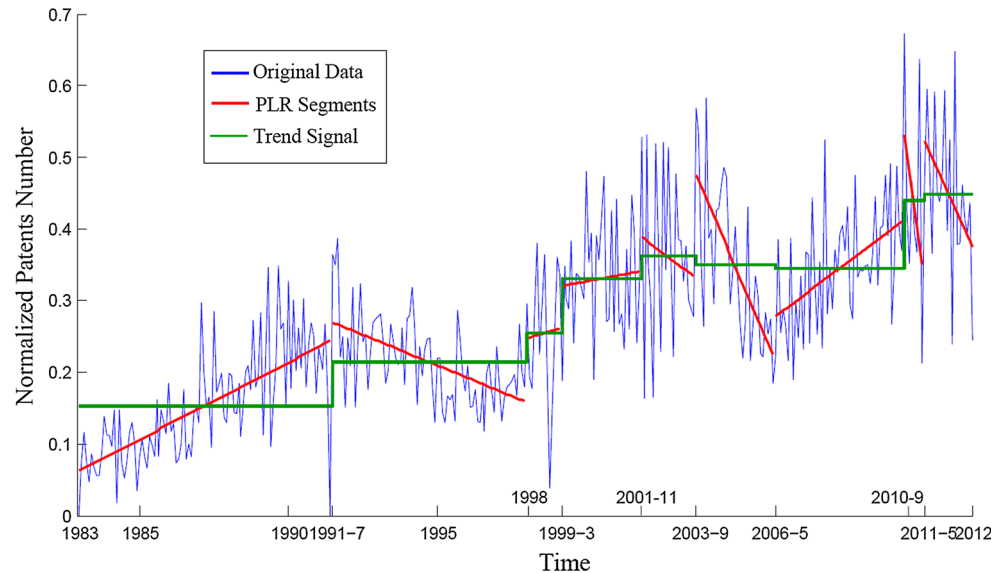


Table 2 Trend signal value and trend tags of ICT technologies in Australia

| Number of segments | Tag start | Tag end | Trend tags | Trend signal value |
|--------------------|-----------|---------|-------------------|--------------------|
| 1 | 1 | 102 | JAN 1983–JUN 1991 | 0.153081429 |
| 2 | 103 | 180 | JUL 1991–DEC 1997 | 0.214253605 |
| 3 | 181 | 194 | JAN 1998–FEB 1999 | 0.254157483 |
| 4 | 195 | 226 | MAR 1999–OCT 2001 | 0.330382188 |
| 5 | 227 | 248 | NOV 2001–AUG 2003 | 0.361596328 |
| 6 | 249 | 280 | SEP 2003–APR 2006 | 0.350017532 |
| 7 | 281 | 332 | MAY 2006–AUG 2010 | 0.344265832 |
| 8 | 333 | 340 | SEP 2010–APR 2011 | 0.440743338 |
| 9 | 341 | 360 | MAY 2011–DEC 2012 | 0.449228612 |

proposes a time series processing component with trend identification functionality. Patent time series processing is first taken into consideration while constructing a technology intelligence framework. PLR module is used in the framework to generate and capture the hidden trend of patent application/publication activities quantitatively. Compared with the methods which can only discover general trend, the usage of PLR overcomes the problem of identifying trend turning points. Finally, the component outcome provides trend tags to the existing text mining component that we see as a black box, thus making it possible to combine the text-based and time-based knowledge together to support technology strategy making more satisfactorily.

In the future work, we will further explore different approaches to model technological trend. In addition, the interaction experiments between patent time series processing component and text mining component will be

done practically. Hence, there will be a series of system implementation works addressed in future research.

Acknowledgments The author Hongshu Chen acknowledges the support from the China Scholarship Council (2011603004) and top-up scholarship from University of Technology, Sydney. In addition, this work was also supported in part by the Australian Research Council (ARC) under discovery Grant DP110103733.

References

- WIPO (2004) Technological and legal developments in intellectual property. In: WIPO Intellectual Property Handbook: Policy, Law and Use, 2nd edn. 489:17–40
- Campbell RS (1983) Patent trends as a technological forecasting tool. *World Patent Inf* 5(3):137–143
- Technology Futures Analysis Methods Working Group (2004) Technology futures analysis: toward integration of the field and new methods. *Technol Forecast Soc Chang* 71(3):287–303
- Yoon B (2008) On the development of a technology intelligence tool for identifying technology opportunity. *Expert Syst Appl* 35(1):124–135
- Kerr CIV, Mortara L, Phaal R, Probert DR (2006) A conceptual model for technology intelligence. *Int J Technol Intell Planning* 2(1):73–93
- Niu L, Lu J, Zhang G (2008) Cognitive orientation in business intelligence systems. *Intelligent decision and policy making support systems*. Springer, Berlin, pp 55–72
- Yoon J, Kim K (2012) TrendPerceptor: a property–function based technology intelligence system for identifying technology trends from patents. *Expert Syst Appl* 39(3):2927–2938
- Zhu D, Porter AL (2002) Automated extraction and visualization of information for technological intelligence and forecasting. *Technol Forecast Soc Chang* 69(5):495–506
- Trippe AJ (2003) Patinformatics: tasks to tools. *World Patent Inf* 25(3):211–221
- Lee C, Cho Y, Seol H, Park Y (2012) A stochastic patent citation analysis approach to assessing future technological impacts. *Technol Forecast Soc Chang* 79(1):16–29

11. Chen H, Zhang G, Lu J (2013) A time-series-based technology intelligence framework by trend prediction functionality. In: 2013 IEEE International Conference on System, Man and Cybernetics, 2013IEEE SMC. Manchester, UK
12. Daim TU, Rueda G, Martin H, Gerdtsri P (2006) Forecasting emerging technologies: use of bibliometrics and patent analysis. *Technol Forecast Soc Chang* 73(8):981–1012
13. Peg Y (1993) Technological growth curves: a competition of forecasting models. *Technol Forecast Soc Chang* 44:375–389
14. Shapiro RD (1985) *Toward effective supplier management: international comparisons: division of research*, Harvard Business School
15. Daim TU, Kocaoglu DF, Anderson TR (2011) Using technological intelligence for strategic decision making in high technology environments. *Technol Forecast Soc Chang* 78(2):197–198
16. Morris S, DeYong C, Wu Z, Salman S, Yemenu D (2002) DIVA: a visualization system for exploring document databases for technology forecasting. *Comput Ind Eng* 43(4):841–862
17. Griliches Z (1990) Patent statistics as economic indicators: a survey. *J Econ Lit* 28(4):1661–1707
18. Ernst H (1997) The use of patent data for technological forecasting: the diffusion of CNC-technology in the machine tool industry. *Small Bus Econ* 9(4):361–381
19. Faust K, Schedl H (1983) International patent data: their utilization for the analysis of technological developments. *World Patent Inf* 5(3):144–157
20. USPTO. Available from: <http://www.uspto.gov/patents/index.jsp>
21. Keogh E, Chu S, Hart D, Pazzani M (2004) Segmenting time series: a survey and novel approach. *Data Min Time Ser Databases* 57:1–22
22. Chang PC, Fan CY, Liu CH (2009) Integrating a piecewise linear representation method and a neural network model for stock trading points prediction. *IEEE Trans Syst Man Cybern Part C Appl Rev* 39(1):80–92
23. Luo L, Chen X (2013) Integrating piecewise linear representation and weighted support vector machine for stock trading signal prediction. *Appl Soft Comput* 13(2):806–816
24. Kimura A, Kashino K, Kurozumi T, Murase H (2008) A quick search method for audio signals based on a piecewise linear representation of feature trajectories. *IEEE Trans Audio Speech Lang Process* 16(2):396–407
25. Keogh E, Chu S, Hart D, Pazzani M (2001) An online algorithm for segmenting time series. In: *Proceedings IEEE ICDM 2001, IEEE International Conference on Data Mining* 289–296
26. Watts RJ, Porter AL (2003) R&D cluster quality measures and technology maturity. *Technol Forecast Soc Chang* 70(8):735–758
27. IP Australia, Australian Government IP Australia website. <http://www.ipaustralia.gov.au/>
28. OECD (2005) *OECD handbook on economic globalisation indicators*. OECD, France
29. Lee S, Kim MS, Park Y (2009) ICT Co-evolution and Korean ICT strategy-An analysis based on patent data. *Telecommun Policy* 33(5):253–271
30. OECD (2010) *Measuring Innovation: A New Perspective*. OECD Publication
31. OECD (2008) *Compendium of Patent Statistics*. www.oecd.org/sti/inno/37569377.pdf