Research on Technological Evolution in IPC Classification Revisions (Postprint)

Authors: Xu Jinglong, Yajuan Zhao

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] This study investigates technological evolution from the perspective of patent classification revisions, offering novel insights for technological evolution research. [Method/Process] First, based on the revision history of Section H of the IPC classification scheme from 2009 to 2018, four types of revision patterns are identified: newly added classifications, deleted classifications, intra-class transfers, and inter-class transfers. Second, to address the cross-filed documents generated by classification revisions, a reclassification method based on the Word2vec+TextCNN model is proposed, thereby bridging old and new versions of the classification scheme through reclassified patents. Finally, a preliminary exploration of technological evolution is conducted by integrating the revised classifications in Section H and the corresponding reclassified patents from 2009 to 2018. [Results/Conclusion] The patent reclassification model effectively resolves the issue of cross-filed documents, provides a reference for patent reclassification practices, and simultaneously bridges old and new versions of patent classification schemes. Integrating IPC classification revisions with reclassified patents enables the analysis of primary technological evolution directions embedded within classification revisions, thereby offering a new perspective for technological evolution research.

Full Text

Preamble

Authors: Xu Jinglong[1], Zhao Yajuan[2],[3]
[1] Library of Party School of Guangdong Provincial Committee of CPC, Guangzhou 510053
[2] National Science Library, Chinese Academy of Sciences, Beijing 100190
[3] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract: [Purpose/Significance] This study examines technological evolution from the perspective of patent classification revision, offering new insights for technology evolution research. [Method/Process] First, based on revisions to Section H of the IPC classification table from 2009 to 2018, we identified four revision types: new classification, deleted classification, intra-class transfer classification, and inter-class transfer classification. Second, to address the archived documents generated after classification revision, we proposed a Word2vec+TextCNN model-based reclassification method for archived documents, enabling connection between old and new classification tables through reclassified patents. Finally, we conducted preliminary exploration of technological evolution by analyzing revised classifications in Section H and their reclassified patents from 2009 to 2018. [Result/Conclusion] The patent reclassification model effectively solves the archived documents problem, provides reference for patent reclassification work, and bridges old and new patent classification tables. Combined with IPC classification revision and reclassified patents, this approach can analyze major technological evolution directions in classification revisions, offering a new perspective for technology evolution research.

Keywords: IPC classification; patent classification revision; patent reclassification; technology evolution
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2021.15.016

Introduction

Technological evolution refers to the description of technological development processes within a certain period, including technological development history, current status, and future trends. This can be achieved through patent bibliometric analysis, text mining, and trend forecasting to describe technological evolution and grasp technological development trajectories. Existing research methods for studying technological evolution based on patent analysis can generally be divided into two categories: (1) technology evolution research based on patent metadata, such as statistical ranking of patent quantities, application regions, and applicants, or analysis of patent classification co-occurrence relationships and patent citation relationships. These studies focus on analyzing trends or relationships in patent external bibliographic items over time. For example, Xie Shoufeng established a research framework for technological evolution from perspectives such as patent application volume and applicants[1]. Liu Yun et al. analyzed global carbon nanotube technology innovation characteristics by combining two-dimensional matrices of patent application volume and growth rate with country/region and institution data, along with patent quality analysis[2]. J. Suzuki et al. analyzed technological evolution trajectories using IPC main classifications, extended classifications, and JPO-specific "facet codes"[3]. (2) Technology evolution research based on patent content, often employing natural language processing and text mining techniques to analyze thematic content and thematic intensity evolution. For instance, Liao Liefa et al. combined LDA topic models with IPC classifications to analyze thematic intensity, thematic content, and technological topic intensity[4]. Chen Wei et al. combined LDA and HMM models to analyze technology theme evolution[5]. J. Yoon et al. analyzed technology trends based on SAO semantic structures[6]. H. Sasaki et al. studied technology convergence and identified technology opportunities through IPC co-occurrence over time series[7]. In these two types of research, some studies use patent classifications as metadata for technology evolution research, while others use them as technological content.

Technology evolution research using patent classifications as metadata often analyzes technology status and trends from statistical and co-occurrence perspectives of classification numbers, examining characteristics such as classification quantity, frequency, and co-occurrence network relationships. For example, J. Zheng et al. analyzed China's industrial and technological development from 2003 to 2008 by counting USPC numbers in six industries[8]. X. Zhou et al. tracked technological development paths through patent classification numbers[9]. S. Jun judged technology trends by mining association rules of patent classifications[10]. Huang Bin et al. identified technology correlation characteristics based on patent classification co-occurrence[11]. J.K.C. Chen et al. studied solar cell technology evolution in Japan, the United States, and other countries/regions through composite networks of patents and patent classifications[12]. Li Ruixi and J. Kraft et al. used centrality, structural holes, and broker methods in classification co-occurrence networks to identify core technologies, intermediary technologies, and emerging technologies in technology correlation networks[13–14]. Additionally, many studies use patent classifications as technological content for technology theme evolution. For example, Fang Shu et al. used classification numbers instead of keywords to cluster classification-document matrices based on semantics, thereby studying technology evolution[15]. Wu Hong et al. proposed a WI-LDA model using IPC classifications as technical context to improve the low topic identification and weak interpretability of traditional LDA models[16]. Lei Tao and Miao Hong et al. studied cross-domain technology correlation characteristics using main IPC and additional IPC flows as knowledge flows[17–18].

Currently, regardless of whether patent classifications are used as metadata or technological content for technology evolution research, most studies directly use the patent classification system as an established knowledge system, rarely considering that the patent classification system is also revised with technological evolution. Therefore, research exploring technological evolution from the perspective of patent classification revision is limited. F. Lafond et al. studied the long-term dynamics of USPC from three aspects: patent classification quantity, classification scale, and reclassification[19]. The RAND Corporation determined the leading or following position of patent holders in technology fields based on the timing of new patent classifications[20]. C.C. Wang studied technology evolution based on five types of U.S. patent reclassifications[21]. Wang Wenjing discussed the correlation between key technologies in the new energy vehicle industry and patent classification revisions[22]. Niu Li et al. described technology changes from IPC revisions in the air conditioning field[23], representing rare research focusing on technological evolution from the perspective of patent classification revision.

This article proposes the proposition of technology evolution research based on IPC classification revision, designs a patent reclassification model to achieve automatic patent reclassification, and uses reclassified patents to connect old and new IPC classification systems. The intention is to describe technology evolution by combining revised classifications and reclassified patents, representing technology evolution results through comparison of revision results between old and new classification tables, and representing the technology evolution process through the reclassification process of archived documents—where patents under revised classifications are transferred from old classifications to new ones. This study of technology evolution through IPC classification revision and patent reclassification can, on the one hand, enrich the methodology system of technology evolution research and provide new perspectives; on the other hand, the patent reclassification method proposed in this article provides implementation methods and pathways for technology evolution research based on IPC classification, offering ideas for WIPO's patent reclassification work and thereby improving IPC classification's function as a tool for patent literature retrieval and management.

1 IPC Classification and Its Revisions

The IPC classification was compiled according to the Strasbourg Agreement signed in 1971[24]. The World Intellectual Property Organization (WIPO) believes that "to maintain the IPC classification as an effective tool for patent literature retrieval and management, it must be dynamic and require continuous revision"[25]. Currently, the IPC classification has been updated to the eighth edition, which has been in use since 2006 and has been revised more than ten times. Since 2009, WIPO has uniformly published new versions on January 1 each year. The historical versions of IPC classification and their effective dates are shown in Table 1 .

IPC revision content includes classification numbers, classification titles (official notes), indexes, notes, and references. Among these, classification numbers and titles determine the scope of technology, while indexes, notes, and references are guiding information for using the IPC classification table. However, revisions to indexes, notes, and references do not involve technological changes. Therefore, this article only focuses on revisions to classification numbers and titles, summarizing four types of IPC classification revisions based on this:

(1) **New Classification**: This refers to directly subdividing existing classification numbers. When technological development enables a technology theme represented by an existing classification to be subdivided in a more refined manner, new classifications are needed. As shown in Figure 1 [Figure 1: see original paper], the revision of the H01L33/00 main group is an

example of new classification.

(2) **Deleted Classification**: This refers to deleting old classification numbers in the new classification table, typically migrating the deleted classification to other positions in the classification table. The migration results manifest in two situations: first, direct migration to other classifications (i.e., merging the classification with others); second, changing the classification number while the technology theme remains unchanged.

(3) **Intra-class Transfer Classification**: This refers to migrating classification numbers within a class, judged by the criterion that the original classification number is deleted and its technology theme is transferred to a new classification number with the same superior class as the original. As shown in Figure 2 [Figure 2: see original paper], the revision of the H04H subclass is an example of intra-class transfer.

(4) **Inter-class Transfer Classification**: This refers to migrating classification numbers between classes, judged by the criterion that the original classification number is deleted and transferred to a new classification with a different superior class. As shown in Figure 3 [Figure 3: see original paper], the revision of the H02S subclass is an example of inter-class transfer.

From the perspective of WIPO's current workflow and mechanisms, patents already classified are not reclassified after patent classification revision. This creates "archived documents"—patents that, due to IPC classification revision, do not have their classification numbers adjusted accordingly. Therefore, when using new classification numbers for patent retrieval, patent literature using old classification numbers cannot be returned as search results to users. These patents that are not reclassified with classification revision become archived documents. Archived documents cause discontinuity between different classification versions when studying technology evolution based on IPC classification revision, meaning old and new versions cannot be connected before and after revision. Consequently, relying solely on pre- and post-revision classification tables can only observe technology evolution results but cannot describe the technology evolution process. In view of this, this article designs a Word2vec+TextCNN model-based reclassification method for archived patent documents. By reclassifying archived documents, we establish mapping relationships between old and new patent classifications and use reclassified patents to describe patent flow information after classification revision, thereby describing the technology evolution process.

2 Data Sources and Research Methods

2.1 Data Sources

As mentioned above, the eighth edition of the IPC classification table is currently the latest version, and since 2009, new versions have been uniformly updated on January 1 each year. Therefore, this study selects 2009-2018 as the research time window. Additionally, comparing the revision frequencies of sections A-H in the IPC classification table from 2009 to 2018, Section H was revised 20 times—the most among all sections—indicating rapid technological innovation and development in the electrical field (Section H). Using Section H data for case study makes it easier to identify characteristics of technological evolution. Therefore, this article selects revised classifications in Section H of the eighth edition IPC classification table as the entry point, using patents published under Section H revised classifications during 2009-2018 as the data source.

2.2 Patent Reclassification Model

Patent reclassification is essentially automatic text classification of archived documents. This article proposes a Word2vec+TextCNN model-based patent reclassification method to solve the archived documents problem after patent classification revision through machine learning models.

Word2vec is a word embedding calculation tool open-sourced by Google in 2013. As a type of word embedding, it is an unsupervised learning model that embeds words into a mathematical space[26]. Simply put, Word2vec can be trained on a given corpus to obtain a pre-trained model that outputs vector representations of all words appearing in the corpus. Based on these vectors, relationships between words can be calculated, such as similarity and semantic associations.

The TextCNN model was proposed by Y. Kim in the paper "Convolutional Neural Networks for Sentence Classification" in 2014[27]. This algorithm uses Convolutional Neural Networks (CNN) for text classification tasks, employing multiple kernels of different sizes to extract key information from sentences, thereby better capturing local correlations.

This article proposes a Word2vec+TextCNN model-based patent reclassification method, which mainly includes data retrieval, data acquisition, data preprocessing, text segmentation, Word2vec model training, patent text vectorization, TextCNN model training, and patent reclassification. The complete technical implementation process is shown in Figure 4 [Figure 4: see original paper]:

(1) **Data Retrieval and Acquisition**: Using WIPO's published revised classifications as search criteria, we selected the Incopat patent database as the data source. Incopat's automatic translation function can translate foreign patents into Chinese, facilitating subsequent Chinese patent text processing. Therefore, we retrieved and obtained patent literature collections from the Incopat database.

(2) **Chinese Word Segmentation and Stop Word Removal**: We used the jieba segmentation tool for Chinese word segmentation, removing stop words such as punctuation marks, special characters, and modal particles. The segmented Chinese titles and abstracts after stop word removal serve as the training corpus for the Word2vec model.

(3) **Word2vec Model Training**: We used the gensim NLP package in Python to train Word2vec, with input corpus being the segmented patent titles and abstracts from step (2), saving the trained Word2vec model.

(4) **Dataset Division**: Using IPC classification revision time as the boundary, patents published before the revision time are archived patent documents (the task dataset for reclassification), while patents published after the revision time point are training data.

(5) **Training Data and Task Data Segmentation and Stop Word Removal**: Perform word segmentation and stop word removal operations separately on training data and task data.

(6) **Text Vectorization**: Call the trained Word2vec model from step (3) to calculate patent text vectors for both training and task data, serving as training corpus for the TextCNN model.

(7) **TextCNN Model Training**: Use tensorflow tools to build the TextCNN model, with the segmented training corpus from step (6) as model input to train the patent reclassification model. During training, randomly select 80% of training data as the training set for generating model parameters, and the remaining 20% as the test set for model effectiveness. We use this approach to calculate model accuracy because WIPO has not yet reclassified archived documents, making it impossible to obtain professionally labeled data (manually reclassified by classification examiners) for model performance validation.

(8) **Reclassification**: Use the trained reclassification model to reclassify the task dataset (archived documents) from step (4) and save the reclassification results. Table 2 shows examples.

2.3 Model Performance Evaluation

Reclassification model performance can be evaluated using common machine learning metrics such as accuracy and loss function. Accuracy represents the proportion of correctly classified samples, directly measuring text classification model effectiveness. Higher accuracy indicates better model performance.

The loss function calculates the distance between predicted and true values after classification. Larger gaps between predictions and true values result in larger loss function values, requiring model optimization to reduce loss. Smaller loss function values indicate better model performance. This article uses the L2 norm loss function, also known as Least Squares Error (LSE), obtained by calculating the minimum sum of squared differences between predicted and true values. The formula is as follows, where y_i and x_i represent true and predicted values respectively:

$$L_2 = \sum_{i}^{m} (y_i - x_i)^2$$

As shown in Figure 5 [Figure 5: see original paper], the horizontal axis represents model training steps, and the vertical axis represents model accuracy. As training steps increase, model accuracy continuously improves. At approximately 1,200 steps, model accuracy approaches 80%, with a maximum accuracy of 82%. Compared with current patent text automatic classification research, this accuracy is relatively high, indicating that the proposed patent reclassification model can effectively reclassify archived documents. As shown in Figure 6 [Figure 6: see original paper], the horizontal axis represents training steps, and the vertical axis represents loss function values. As training steps increase, the loss function value continuously decreases. After 1,000 steps, the loss function curve converges, with values below 0.5, indicating the trained model gradually approaches optimal state and achieves optimal accuracy. Considering both accuracy and loss function curves, the reclassification model in this study has achieved high accuracy—approximately 80% classification accuracy—providing data quality assurance for subsequent technology evolution analysis using reclassified patents.

3 Technology Evolution Analysis in IPC Classification Revisions: The Case of Section H

3.1 Application Trend Analysis

First, we analyze patent publication trends in Section H. The left panel of Figure 7 [Figure 7: see original paper] shows patent publication trends for Section H, revised classifications, and unrevised classifications from 2009 to 2018. Section H patent publications show a yearly upward trend. Unrevised classification publications in Section H grew slightly from 2011 to 2014. Revised classification publications grew steadily from 2009 to 2014, then grew rapidly after 2014. The trend line for revised classification patents roughly follows the same pattern as the overall Section H trend, indicating that Section H patent publication changes are significantly influenced by revised classification publication changes. Regarding the proportion of revised versus unrevised classification patents in Section H, the right panel of Figure 7 [Figure 7: see original paper] shows that before 2014, the proportions remained relatively stable, with unrevised classifications accounting for approximately 76% of all Section H publications. After 2014, the proportion of unrevised classifications in Section H 逐年下降, dropping to about 60% by 2018, while the proportion of revised classifications increased from 24% in 2009 to 40% in 2018, demonstrating rapid development in technologies represented by revised classifications.

3.2 Technology Evolution Analysis of New Classifications

New classifications can represent new developments in technology fields. When new technologies emerge that cannot be covered by existing classifications, new classifications are added to represent technological advances.

As shown in Table 3 , within the 10-year window of 2009-2018, Section H's new classifications concentrated in three areas: H01 (basic electrical elements), H02 (generation, conversion, or distribution of electric power), and H04 (electric communication technique), involving H01F (magnets; inductances; transformers; selection of materials for their magnetic properties), H01G (capacitors; rectifiers, detectors, switching devices, light-sensitive or temperature-sensitive devices), H01L (semiconductor devices), H01M (conversion of chemical energy into electrical energy, e.g., batteries), H01R (electrically-conductive connections; connecting arrangements of electric components), H01Q (aerials), H02K (electric motors), H02P (control or regulation of electric motors, generators, or dynamo-electric converters), H03K (pulse technique), H04B (transmission), H04L (transmission of digital information, e.g., telegraphic communication), and H04N (pictorial communication, e.g., television). Table 3 data shows that H01L (semiconductor devices) and H01M (conversion of chemical energy into electrical energy, e.g., batteries) accounted for the largest proportions of patent publications among Section H's new classifications from 2009 to 2018, at 24.78% and 24.83% respectively, indicating significant technological innovation in these fields. Next were H04L (transmission of digital information, e.g., telegraphic communication) and H04N (pictorial communication, e.g., television), accounting for 16.49% and 17.36% respectively. Patent publications for new classifications in remaining subclasses were all at or below 5%.

3.3 Technology Evolution Analysis of Deleted Classifications

In technological development, while new technologies emerge, old technologies are also eliminated and become obsolete. When old classification methods no longer suit new technological development, old classifications are deleted from the classification table. Therefore, deleted classifications can represent eliminated or obsolete technologies.

As shown in Table 4 , deleted classifications in Section H from 2009 to 2018 mainly involved four technology fields: H01 (basic electrical elements), H02 (generation, conversion, or distribution of electric power), H03 (basic electronic circuitry), and H04 (electric communication technique), including H01G (capacitors; rectifiers, detectors, switching devices, light-sensitive or temperature-sensitive devices), H01H (electric switches, relays, selectors, emergency protective devices), H01J (discharge tubes or discharge lamps), H01L (semiconductor devices), H01M (conversion of chemical energy into electrical energy, e.g., batteries), H01Q (aerials), H01R (electrically-conductive connections), H02K (electric motors), H02N (electric machines not otherwise provided for), H03F (amplifiers), H04B (transmission), H04J (multiplex communication), H04L (transmission of digital information, e.g., telegraphic communication), H04N (pictorial communication, e.g., television), H04Q (selecting), and H04W (wireless communication networks). Among these, H04N (pictorial communication, e.g., television), H04W (wireless communication networks), and H04L (transmission of digital information, e.g., telegraphic communication) accounted for 26.11%, 25.64%, and 21.38% respectively—the largest proportions—indicating more eliminated or obsolete technologies in these fields. In other subclasses, only H01L (semiconductor devices) accounted for 6% of patent publications, with remaining subclasses all below 5%.

3.4 Technology Evolution Analysis of Intra-class Transfer Classifications

Section H intra-class transfer classifications from 2009 to 2018 mainly involved classifications in H01 (basic electrical elements) and H04 (electric communication technique), including the H01R12/00 main group (electrically-conductive connections for printed circuits, flat or ribbon cables) revised in 2011, the H01R24/00 main group (two-part coupling devices) revised in 2011, the H01J11/00 main group (gas-filled discharge tubes with AC-induced discharge, e.g., AC-PDP) revised in 2012, and the H04N13/00 main group (stereoscopic video systems; multi-viewpoint video systems; parts thereof) revised in 2018.

The revision and technology evolution of the H01R12/00 main group (electrically-conductive connections for printed circuits, flat or ribbon cables) are shown in Figures 8 [Figure 8: see original paper] and 9 [Figure 9: see original paper]. The H01R12/00 revision adjusted its lower-level subgroups, adding more detailed technical branches. In Figure 9, the left side shows the pre-revision technology classification and patent volume, the right side shows the post-revision classification and patent volume, and the connecting lines represent connections between old and new versions through reclassification (same below). This shows that H01R12/71 (coupling devices for rigid printed circuits or similar structures) is the main technology evolution direction after classification revision, accounting for approximately 59.61% of patents in the H01R12/00 main group from 2009-2010. Next is H01R12/70 (coupling devices), accounting for about 25.9%.

The revision of the H01R24/00 main group (two-part coupling devices) deleted all subgroups in the old classification and redefined new subgroups. The pre- and post-revision technology structures are shown in Figure 10 [Figure 10: see original paper]. Technology evolution results in Figure 11 [Figure 11: see original paper] show that H01R24/38 (with coaxial or coaxial arrangement of contacts), H01R24/00 (two-part coupling devices), and H01R24/00 (two-part coupling devices) are the main technology evolution directions, with patent proportions of approximately 27.37% and 25.33%.

The H01J11/00 main group (stereoscopic video systems; multi-viewpoint video systems; parts thereof) revision transferred subgroups H01J11/02 and

H01J11/04 to H01J11/10-H01J11/54 in the 2012 IPC classification. The pre- and post-revision technology structures and technology evolution are shown in Figures 12 [Figure 12: see original paper] and 13 [Figure 13: see original paper]. Among these, H01J11/22 (electrodes for gas-filled discharge tubes, e.g., special shape, material, or construction) and H01J11/46 (discharge tubes with liquid-pool cathodes structurally associated with one or more circuit elements) are the main technology evolution directions, with patent proportions of approximately 53.34% and 17.72% respectively.

The H04N13/00 main group (stereoscopic video systems; multi-viewpoint video systems; parts thereof) revision transferred H04N13/02 and H04N13/04 to H04N13/10-H04N13/198 in the 2018 IPC classification. The pre- and post-revision technology structures and technology evolution are shown in Figures 14 [Figure 14: see original paper] and 15 [Figure 15: see original paper]. Technology evolution results show that H04N13/30 (image reproducers), H04N13/20 (image signal generators), and H04N13/10 (processing, recording, or transmission of stereoscopic or multi-viewpoint image signals) are the main technology evolution directions, with patent proportions of approximately 25.92%, 13.03%, and 10.13% respectively.

3.5 Technology Evolution Analysis of Inter-class Transfer Classifications

Section H inter-class transfer classifications from 2009 to 2018 mainly involved classifications in H02 (generation, conversion, or distribution of electric power) and H04 (electric communication technique), including H04W (wireless communication networks) revised in 2009, H04N21/00 (selective content distribution, e.g., interactive television or video on demand) revised in 2011, and H02S (generation of electric power by conversion of infrared radiation, visible light, or ultraviolet light, e.g., using photovoltaic [PV] modules) revised in 2014.

The H04W subclass was created by transferring the H04Q7/00 main group to the H04W subclass. The pre- and post-revision technology structures and technology evolution are shown in Figures 16 [Figure 16: see original paper] and 17 [Figure 17: see original paper]. Among these, H04W4/00 (services specially adapted for wireless communication networks; facilities therefor), H04W4/02 (services using location information for wireless communication networks), H04W4/12 (messaging; mailboxes; announcements), H04W4/16 (supplementary services, e.g., call transfer or call hold), and H04W4/06 (broadcast or multicast services, e.g., multimedia broadcast multicast service [MBMS]; to groups of users; one-way selective calling) are the main technology evolution directions, with patent proportions of approximately 48.16%, 10.44%, 10.54%, 7.47%, and 6.65% respectively.

The H04N21/00 main group revision transferred H04N7/58, H04N7/60, and H04N7/62 to H04N21/00. The pre- and post-revision technology structures and technology evolution are shown in Figures 18 [Figure 18: see original paper] and 19 [Figure 19: see original paper]. Among these, H04N21/43 (reception or interaction with content at the user's home terminal), H04N21/242 (synchronization processes, e.g., processing of program clock references) are the main technology evolution directions, accounting for 13.12% and 10.0% respectively.

The H02S subclass revision transferred the H02N6/00 and H01L31/045 and H01L31/058 subclasses to the H02S subclass. The pre- and post-revision technology structures and technology evolution are shown in Figures 20 [Figure 20: see original paper] and 21 [Figure 21: see original paper]. Among these, H02S20/32 (specially adapted support structures for solar tracking PV modules), H02S40/22 (light-reflecting or light-concentrating devices for PV modules), and H02S30/20 (foldable PV modules) are the main technology evolution directions, accounting for approximately 12.47%, 9.74%, and 6.73% respectively.

Conclusion

Technology evolution has a complete research system, but studies from the perspective of patent classification revision are limited. This article conducts preliminary exploration from this angle. First, starting from patent classification revision, we summarized four revision types based on Section H IPC revisions from 2009 to 2018: new classification, deleted classification, intra-class transfer classification, and inter-class transfer classification. We also introduced the concept of "archived documents" generated after classification revision. Second, to address the discontinuity between old and new patent classification tables caused by archived documents, we proposed a Word2vec+TextCNN model-based patent reclassification method. By reclassifying archived documents, we connected old and new patent classification tables. Finally, by analyzing changes in technology structure before and after revision and conducting statistical analysis of reclassified patents, we described major technology evolution directions in classification revisions. This study not only provides new perspectives for technology evolution research and expands the technology evolution research system, but also offers reference for archived document reclassification work, representing a beneficial exploration.

This study has several limitations: First, it fails to provide a panoramic description of technology evolution in Section H or realize a "technology evolution panorama" through visualization, instead interpreting revised classifications from 2009-2018 separately, lacking comprehensiveness and integrity. Second, the Word2vec+TextCNN model accuracy currently only reaches about 80%, and reclassification accuracy needs improvement, which significantly impacts the accuracy of technology evolution intelligence provided in this study. Finally, the study fails to propose measurement indicators for describing technology evolution, using only proportions to describe major technology evolution directions, making it difficult to deeply reveal technology evolution information. Future work should further improve this.

References

[1] Xie Shoufeng. Research on Technology Evolution and Forecasting Based on Patent Analysis[D]. Nanjing: Nanjing University of Science and Technology, 2014.

[2] Liu Yun, Liu Lu, Yan Zhe. Analysis of Global Carbon Nanotube Technology Innovation Characteristics Based on Patent Bibliometrics[J]. Science Research Management, 2016(S1): 337-345.

[3] Suzuki J, Kodama F, et al. Technological Diversity of Persistent Innovators in Japan: Two Case Studies of Large Japanese Firms[J]. Research Policy, 2004, 33(3): 531-549.

[4] Liao Liefa, Le Fugang. Patent Technology Evolution Research Based on LDA Model and Classification Numbers[J]. Modern Information, 2017(5): 15-20.

[5] Chen Wei, Lin Chaoran, Li Jinqiu, et al. Technology Theme Evolution Trend Analysis Based on LDA-HMM Model: A Case Study of Marine Diesel Engine Technology[J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(7): 732-741.

[6] Yoon J, Kim K. Identifying Rapidly Evolving Technological Trends for R&D Planning Using SAO-based Semantic Patent Networks[J]. Scientometrics, 2011, 88(1): 213-228.

[7] Sasaki H, Kajikawa Y, Sakata I, et al. Predicting Potential Industrial Fields of Technological Spin-offs by Using IPC in Patent Analysis[C]//Proceedings of Technology Management for Emerging Technologies. Vancouver: IEEE, 2012: 975-991.

[8] Zheng J, Chen DZ, Huang MH, et al. Industry and Technology Development in China from 2003 to 2008: A Perspective from Patent Classification Analysis[C]//Technology Management for Global Economic Growth. Phuket: IEEE, 2010: 2256-2263.

[9] Zhou X, Zhang Y, Porter AL, et al. A Patent Analysis Method to Trace Technology Evolutionary Pathways[J]. Scientometrics, 2014, 100(3): 705-721.

[10] Jun S. IPC Code Analysis of Patent Documents Using Association Rules and Maps-Patent Analysis of Database Technology[C]//International Conference on Database Theory and Application/International Conference on Bio-Science and Bio-Technology. Jeju Island: Springer-Verlag Berlin, 2011: 21-30.

[11] Huang Bin, Huang Lucheng, Wu Feifei, et al. Technology Correlation Characteristics Identification Based on Patent Co-classification[J]. Intelligence Magazine, 2015(1): 44-48.

[12] Chen JKC, Pham VK, Lin FC, et al. Studying the Patent of Technology Development on Dye Sensitized Solar Cell[C]//Technology Management in the IT-Driven Services. San Jose: IEEE, 2012: 985-996.

[13] Li Ruixi, Chen Xiangdong. Research on Key Technology Identification and Technology Development Patterns Based on Patent Co-classification[J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(5): 49-56.

[14] Kraft J, Quatraro F, Saviotti PP. The Knowledge-base Evolution in Biotechnology: A Social Network Analysis[J]. Economics of Innovation and New Technology, 2011, 20(5): 445-475.

[15] Fang Shu, Hu Zhengyin, Pang Hongqian, et al. Research on Technology Evolution Analysis Methods Based on Patent Literature[J]. Library and Information Service, 2011, 55(22): 42-46.

[16] Wu Hong, Yi Huifang, Ma Yongxin, et al. Research on WI-LDA Model for Patent Technology Theme Analysis[J]. Library and Information Service, 2018, 62(17): 68-74.

[17] Lei Tao, Chen Xiangdong. Blockmodel Analysis of Cross-disciplinary Correlation in ICT Field: Based on University-Enterprise Cooperative Patents in China[J]. China Soft Science, 2011(1): 67-76.

[18] Miao Hong, Wang Xiaoyu, Huang Lucheng, et al. Measurement and Analysis of Cross-domain Technology Characteristics Based on Patents[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(5): 484-492.

[19] Lafond F, Kim D. Long-run Dynamics of the U.S. Patent Classification System[J]. Journal of Evolutionary Economics, 2019, 29(2): 631-664.

[20] Eusebi A, Silberglitt R. Identification and Analysis of Technology Emergence Using Patent Classification[R]. California: RAND Corporation, 2014.

[21] Wang CC, Sung HY, Huang MH. Technological Evolution Seen from the USPC Reclassifications[J]. Scientometrics, 2016, 107(2): 537-553.

[22] Wang Wenjing. Preliminary Study on the Correlation Between Patent Classification Revision and Industrial Development in New Energy Vehicle Industry[J]. Science and Technology Innovation and Application, 2017(14): 25-26.

[23] Niu Li, Xu Kai, Wang Yangping. IPC Revision in Air Conditioning Field and Its Application in Search[J]. China Invention & Patent, 2012(S1): 136-139.

[24] World Intellectual Property Organization (WIPO)[EB/OL]. [2020-11-14]. http://www.wipo.int/classifications/ipc/zh/.

[25] World Intellectual Property Organization. Guidelines for Revision of the IPC[EB/OL]. [2020-11-14]. https://www.wipo.int/export/sites/www/classifications/ipc/en/general/guidelines

[26] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. arXiv preprint. arXiv:1301.3781, 2013.

[27] Kim Y. Convolutional Neural Networks for Sentence Classification[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014: 1746-1751.

Author Contributions: Xu Jinglong: conceptual structure, research implementation, paper writing and revision; Zhao Yajuan: topic selection, guidance on research direction, paper revision and final approval.