Identification and Tracking of Technological Innovation Portfolios Based on Deep Learning and Semantic Mining (Postprint)

Authors: Zhou Xiao, Xu Yinbiao, Shi Yi

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

Abstract

[目的/意义]With the rapid development of strategic emerging technology industries, identifying technological innovation combinations with potential synergistic effects and clarifying the core innovation relationships within these combinations constitute important prerequisites for effectively planning industrial development routes and enhancing industrial competitive advantages. [方法/过程]Guided by technological combination evolution theory and integrating deep learning, SAO semantic mining, and the CFDP algorithm, this study proposes an identification scheme for technological innovation combinations and evolutionary relationships based on patent data. This research scheme consists of three steps: first, constructing a domain retrieval strategy based on keywords and patent classification codes, and implementing data cleaning and segmentation for the acquired data; subsequently, constructing a word vector semantic network for domain technology topics through Word2Vec, and utilizing the CFDP algorithm to identify potential innovation elements and combination patterns; finally, deeply mining the core SAO structures within each combination, classifying their evolutionary relationships through the LSTM deep learning algorithm, excavating the core innovation approaches of technologies, and thereby effectively identifying potential technological opportunities in the domain. [结果/结论]Taking the speech recognition domain as an example, through in-depth mining of DII patent text data in this domain, five potential technological innovation combinations and core innovation approaches are identified and tracked. The study finds that China's current speech recognition domain possesses substantial innovation potential in areas such as intelligent chip design, speech recognition algorithms, and new scenarios and applications.

ChinaRxiv 中X

Full Text

Identifying and Tracking Technological Innovation Combinations Based on Deep Learning and Semantic Mining

Zhou Xiao[1], **Xu Yinbiao**[1], **Shi Yi**[2] [1]School of Economics and Management, Xidian University, Xi'an 710126 [2]College of Computer Science, Xi'an Shiyou University, Xi'an 710065

Abstract

[**Purpose/Significance**] With the rapid development of strategic emerging technology industries, identifying technological innovation combinations with potential synergistic effects and clarifying the core innovation relationships within these combinations constitute an important prerequisite for effectively planning industrial development routes and enhancing industrial competitive advantages. [**Method/Process**] Guided by technology portfolio evolution theory, this study proposes a patent data-based approach for identifying technological innovation combinations and their evolutionary relationships by integrating deep learning, SAO semantic mining, and the CFDP algorithm. The research framework consists of three steps: First, constructing a domain search strategy based on keywords and patent classification numbers, followed by data cleaning and segmentation. Second, building a word vector semantic network of domain technology topics through Word2Vec, and utilizing the CFDP algorithm to identify potential innovation elements and combination patterns. Finally, deeply mining the core SAO structures within each combination, classifying their evolutionary relationships through an LSTM deep learning algorithm, and uncovering core innovation approaches to effectively identify potential technological opportunities in the field. [**Result/Conclusion**] Taking the speech recognition domain as an example, through in-depth mining of DII patent text data in this field, the study identifies and tracks five potential technological innovation combinations and core innovation methods. The findings reveal that China's speech recognition field currently possesses significant innovation potential in areas such as intelligent chip design, speech recognition algorithms, and new scenarios and applications.

Keywords: technological innovation combination identification; deep learning; SAO method; semantic mining; patent analysis **Classification Number:** G305
DOI: 10.13266/j.issn.0252-3116.2022.10.003

As global technological innovation enters an unprecedented period of intense activity, a new round of scientific and technological revolution and industrial transformation is reshaping the global innovation landscape and economic structure. To achieve high-level technological self-reliance, China must accelerate the improvement of its innovation capabilities and standards. In September 2020, President Xi Jinping profoundly elucidated the major strategic significance of accelerating technological innovation development in today's challenging world situation, particularly emphasizing that in key core emerging technology fields, China must build high-end, leading first-mover advantages by grasping the sources and directions of innovation, concentrating resources and strength to achieve breakthroughs. However, due to the rapid iteration of emerging technologies and the continuous emergence of new technologies and applications, accurately positioning technological innovation directions and planning technological innovation approaches have become two major scientific challenges facing current governments and industries. Solving these challenges is crucial for shortening innovation cycles and enhancing industrial competitive advantages.

Academia has conducted preliminary explorations into solutions for these problems, attempting to use quantitative analyses such as bibliometrics and text mining to identify potential technological innovation opportunities, yielding some valuable research results. However, current research mostly focuses only on outcomes—identifying future valuable technological innovation directions—while lacking in-depth exploration of the principles and mechanisms underlying technological innovation. Consequently, these studies cannot answer the focal question of "how technology can be effectively innovated," nor can they track the process of technological innovation to help industries effectively adjust their innovation directions.

In recent years, some scholars have conducted identification research on innovation themes from a dynamic development perspective centered on "evolutionary innovation." Among them, the technology portfolio evolution theory proposed by technology thinker Brian Arthur represents a typical example. Arthur proposed that "all technologies are derived from previous technology sets" and that "technology creates itself from itself," articulating a "self-generating" viewpoint and systematically elaborating on the modular thinking within technology and its combinatorial evolutionary relationships. This demonstrates that combinatorial evolution between technologies represents an important pathway for technological innovation. Building upon this foundation, the development of deep learning technologies has made it possible to mine implicit technological evolutionary relationships from massive scientific and technological data.

Current artificial intelligence stands at the forefront of new-generation information technology development and represents a key focus in the new round of international competition. China attaches great importance to the innovative development of the artificial intelligence industry, defining it as the top priority industry for the next decade in its 14th Five-Year Plan. Among AI domains, speech recognition constitutes a core R&D direction and has become a focal point for industry attention in recent years. However, due to the interdisciplinary nature of this field and its numerous sub-technologies and sub-applications, identifying promising development directions from continuously emerging new technologies and applications has become a key concern at both national and industrial levels. Against this backdrop, this study systematically analyzes the combinatorial patterns and evolutionary characteristics of technology modules in the speech recognition field, constructs a technological innovation combination identification model based on deep learning and SAO semantic mining methods, and achieves effective exploration of technological innovation themes and approaches in this domain, providing support for promoting the R&D and cultivation of China's speech recognition technology.

2 Literature Review

2.1 Core Theories and Methods for Technological Innovation Theme Identification

With continuous in-depth research on the conceptual characteristics, development patterns, identification, and evaluation methods of emerging technologies, studies on identifying emerging technological innovation themes have gradually formed two research perspectives: the bibliometrics-based identification perspective and the evolution theory-based identification perspective. The bibliometrics-based perspective represents the current mainstream research approach. Such studies primarily mine macro-level patterns of technological development from the quantity of scientific achievements and relationships among R&D entities shown in scientific literature to explore potential technological themes and development opportunities. For instance, Du Jian et al. used paper co-citation and patent coupling to identify cutting-edge technologies in the clinical medicine field, while J. Liu et al. employed bibliometrics and thematic analysis to identify emerging technology themes in gene editing. However, because these studies do not consider the endogenous effects of technological development, they struggle to explain technological innovation methods resulting from the combined influence of internal innovation, external integration, and other social factors.

The evolutionary perspective, by contrast, approaches from the angle of technological evolution, considering the fusion, evolution, and regeneration relationships between technology themes over time. Particularly after introducing relationship-based analysis methods such as technological morphology analysis and SAO semantic mining, this perspective can intuitively discover the bottom-up dynamic formation process of emerging technologies, compensating for the bibliometrics perspective's insufficient explanation of innovation elements and patterns. W. Schoenmakers et al. analyzed different types of disruptive technologies, attempting to explore cross-innovation generated by mature technologies and multidisciplinary knowledge during technological evolution from a fusion perspective. M. Karvonen et al. constructed citation networks to discriminate the convergence characteristics of technology fusion, achieving effective prediction of initial-stage technology fusion methods.

This study focuses on the combinatorial innovation mechanism of technology. Therefore, this section will specifically review relevant achievements in technological innovation theme identification research based on combinatorial evolution theory.

2.2 Technological Innovation Theme Identification Research Based on Combinatorial Evolution Theory

The earliest research on combinatorial evolution theory can be traced back to Joseph Schumpeter's "combination-driven innovation" theory proposed in 1912. This theory suggested that introducing new combinations of production factors and conditions into production systems, resulting in any changes, could be considered as originating from new combinations of production methods. Subsequently, in the 1940s, American astrophysicist Fritz Zwicky further proposed "morphological analysis theory," emphasizing that technology domains (systems) could be decomposed into several functional components, and that complex scientific problems could be solved by combining technological means (morphologies) of these components. In 2011, American scholar Brian Arthur proposed in *The Nature of Technology* that "only a small portion of innovations are primary and fundamental, while a considerable portion are combinatorial innovations generated through transplantation and fusion of technologies within and outside the field." Domestic scholars have also discussed the essence of combinatorial innovation. Sun Bing proposed from the perspective of "technological innovation dynamic mechanism" that self-organization among innovation elements can effectively describe complex innovation relationships among components within systems. Chi Xueqin et al. believed that the combinatorial innovation method involves researchers cleverly combining or reorganizing two or more independent technical principles to obtain a new invention with complete and unified functions. Wu Hong et al. proposed that during technological development, connections (combinations) may occur between existing old technologies, between new technologies, or between old and new technologies, thereby forming new scientific and technological innovations. Technology thus achieves gradual evolution through the cycle of "combination-cumulation-recombination-recumulation." In summary, technological innovation cannot be separated from the fusion and recombination of technological elements, and effective exploration of technological innovation mechanisms and combinatorial patterns can better identify valuable technological opportunities.

Current research on technology combination identification primarily employs network analysis and clustering analysis methods. For example, Zhou Liying et al. constructed patent cooperative R&D networks to identify patent technology combinations based on technological domain relevance. Zhang Zhengang et al. decomposed technology into combinations of knowledge elements, mining potential technology combinations through two approaches: reusing existing knowledge combinations and exploring new ones. Wang Xianwen et al. used the Girvan-Newman algorithm for clustering analysis of technology co-classification networks, combined with social network analysis and information visualization techniques to explore key technology fields and network structures, and further identify relevant technology combination patterns. Our research team has also attempted to divide technology domains into subsystems and integrate network analysis and text mining methods to achieve technology recombination and potential evaluation.

Although these methods consider domain technological attributes and demonstrate significant effectiveness in identifying potential technology combinations, they still fail to systematically explore the innovation relationships between technologies, and thus cannot plan potential technological innovation approaches. In recent years, some scholars have attempted to use the SAO method (Subject-Action-Object) from a semantic perspective to deeply explore the evolutionary and constitutive relationships between technology themes, thereby identifying potentially related technological innovation combinations. Li Xin et al. analyzed core SAO semantic structures in patents to reveal association relationships between technology themes and propose an effective methodology for identifying technological innovation domains. Miao Hong et al. innovatively proposed a function-oriented technology fusion analysis framework, analyzing and predicting overall technology fusion trends within the same function cluster based on the comprehensive utilization of SAO and "Technology-Relationship-Technology" (TRT) structural relationships, providing new ideas for technological innovation combination identification research.

With the rise of deep learning algorithms, the SAO method combined with deep learning can deeply explore complex evolutionary relationships contained in massive unstructured texts, demonstrating great potential in large-scale data processing and deep semantic mining. Based on this, we systematically analyze the composition principles of technology, integrate deep learning and SAO semantic analysis methods, explore identification schemes for technological innovation combinations, and conduct in-depth mining of their technological innovation relationships.

In summary, current academia has conducted preliminary explorations on how to identify technological innovation combinations and clarify their evolutionary relationships, achieving some valuable research results. However, existing research still has two major problems: (1) Most studies utilize explicit relationships (such as co-occurrence relationships between technology points) to define associated elements of technology combinations, lacking in-depth exploration of potential associations triggered by "functional similarity" or "application similarity" between technological elements; (2) They lack intelligent mining of complex relationships between technological elements in massive texts, and thus cannot effectively identify core innovation approaches between technologies.

3 Research Methods

This study aims to explore solutions to two major scientific questions: (1) "How to accurately position technological innovation directions," i.e., identifying what technological innovation combinations are; and (2) "How to effectively plan technological innovation approaches," i.e., how innovation elements should be associated to achieve technological innovation. Therefore, the research is divided into two major phases. The research framework is shown in Figure 1 [Figure 1:

see original paper].

3.1 Technological Innovation Combination Identification

The first research question this study addresses is "where technology will innovate" based on massive patent data. Preliminary investigations found that individual technology themes have limited support for business decisions because they cannot effectively reflect domain innovation elements and approaches. Therefore, how to identify technological innovation combinations with potential synergistic effects based on the basic laws of technological innovation constitutes an important issue. To address this research need, this section attempts to construct a contextual semantic network reflecting potential associations between technologies based on word vector methods. Unlike traditional networks built on word co-occurrence relationships, node positions in this network reflect the "potential" research associations between these technology points—i.e., technology points have similar functions or application methods and have the potential to be combined to achieve a certain innovation process (application). To realize this idea, the research will proceed through the following four steps:

First, formulate a domain search strategy to obtain domain patent data. Subsequently, utilize the text segmentation, cleaning, and merging functions of ITGinsight to complete preliminary screening of domain theme words.

Second, use initial keywords as input to build a domain semantic network based on the Word2Vec model. Word2Vec is a neural network probabilistic language model used to calculate word vectors, capable of fully mining contextual semantic information from texts.

Third, based on the CFDP algorithm, identify core nodes in the semantic network from two indicators: node density and distance. It should be emphasized that core nodes should satisfy two characteristics: (1) the density of surrounding nodes is lower than that of the node itself; (2) surrounding nodes are closer to this node than to other nodes. For local density calculation, we adopt the Cut-off kernel method, as shown in formula (1):

$$\rho_i = \sum_{j \in IS\ \{i\}} \chi(d_{ij} - d_c)$$

where $\chi(x)$ indicates whether the distance d_{ij} between nodes is greater than or equal to the preset cutoff distance d_c, taking 0 if greater than or equal and 1 if less, as shown in formula (2):

$$\chi(d_{ij} - d_c) = \chi(x) = \begin{cases} 0, & x \geq 0 \\ 1, & x < 0 \end{cases}$$

The calculation of inter-node distance δ_i is divided into two cases: (1) For non-local density maximum points i, the distance calculation involves two steps:

first, find all points with higher local density than point i, then find the point j closest to i among these points, and the distance between i and j is the value of δ_i. (2) For local density maximum points, δ_i is the maximum distance between this point and all other points, as shown in formula (3):

$$\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij})$$

Based on formulas (1)-(3), nodes with significantly greater density and distance than other points are selected as core nodes, and several nodes closely associated with these core nodes are identified based on distance as potential combinatorial innovation elements.

Fourth, under the guidance of literature review and expert opinions, construct a domain conceptual model, and classify the innovation elements in each combination according to the classification and composition relationships of technology modules in the conceptual model.

3.2 Combinatorial Innovation Approach Identification

The second research question this study addresses is "how technology can be effectively innovated." To solve this problem, based on the identification of combinatorial innovation elements, this study further explores the combinatorial evolution patterns among innovation elements to effectively plan technological innovation approaches. The SAO (Subject-Action-Object) method can reveal semantic associations between technology points and is a representative method for identifying technological evolution relationships. In SAO structures, S refers to the subject and O to the object, both representing technology themes; A refers to the verb, reflecting the technological evolution approach. Preliminary research has shown that the SAO method combined with deep learning can deeply mine semantic information in massive texts and is an important method for exploring innovation approaches between technology themes. Therefore, this stage constructs a combinatorial innovation approach identification scheme based on the combination of the SAO method and deep learning.

Since technological evolution relationships in SAO structures are primarily represented by predicate verbs A, identifying these evolution relationships is transformed into classifying the predicate verbs A. We employ the Long Short-Term Memory (LSTM) deep learning algorithm to identify technological evolution relationships in SAO structures. LSTM is a deep learning algorithm based on Recurrent Neural Networks (RNN) that adds "gates" with memory functions to control information flow in neurons, demonstrating good performance in accumulating semantic relationships in short texts that change over time and subsequently classifying them. Therefore, we propose to use the LSTM algorithm to mine innovation relationships in massive SAO structures. The specific steps are as follows:

First, use the TextBlob package to segment patent texts into sentences, use the core innovation elements obtained in the previous stage as core corpora constituting S and O, and then extract short sentences containing both S and O from patent texts using basic regular expression syntax.

Second, identify verbs a_i between S and O in sentences through part-of-speech tagging to form a verb list $A = [a_1, a_2, \ldots, a_n]$.

Third, simplify sentences. Concatenate the extracted S, A, and O in the order they appear in the sentence to form a concise SAO structure. Repeat these three steps to complete simplification of all sentences.

Fourth, group SAO structures with the same S and O, and then learn and classify the technological evolution relationships represented by A in the same SAO group. Through domain investigation and expert interviews, we categorize core innovation relationships in the field into four types: promotion, reduction, inclusion, and application, with meanings and explanations for each relationship shown in Table 1 .

Fifth, randomly select a certain proportion of SAO structures as training sets and manually classify their innovation relationships (A) with expert assistance (assigning labels). It should be noted that since A in SAO structures extracted from patent texts is typically an objective verb describing evolution approaches (such as "promote," "work together," "include"), without strong emotional tendencies, unsupervised classification algorithms prove ineffective. Therefore, we adopt supervised learning by manually classifying some structures in advance to achieve effective learning of core innovation relationships.

Sixth, use the manually classified data from the previous step as input samples for the LSTM model. After vector transformation and cyclic iteration, logistic regression is used to obtain category distribution vectors, thereby completing effective training of the classification model.

Seventh, input the remaining unclassified SAO structures into the trained model to ultimately obtain relationship classifications between all S and O pairs.

To systematically demonstrate the research logic of this section, we provide a detailed illustration of this process in Figure 2 [Figure 2: see original paper].

4 Case Study

Speech recognition is an important development direction in the artificial intelligence field. After experiencing explosive growth, its growth rate has gradually slowed, and the market urgently needs to find new entry points for industrial innovation. Against this background, this study takes the speech recognition field as a research object to explore potential technological innovation approaches in this domain and thereby test the feasibility of the proposed research framework.

4.1 Domain Data Acquisition and Cleaning

To comprehensively track R&D progress in the field and mine potential technology combinations, we developed a domain search strategy from both "keywords" and "patent classification" perspectives under the guidance of information retrieval theory. Keyword-based search schemes can comprehensively obtain overall domain data, while secondary retrieval based on patent classification can effectively "refine" the data collection to ensure retrieval accuracy. Guided by this construction 思路 and combined with literature research results, we formulated a domain search formula (as shown in Table 2) and obtained 8,818 DII patent family data entries in this field from 2011-2021.

On the basis of the initial retrieval results, we used ITGinsight software and Python for data cleaning and merging: (1) segment patent data through ITGinsight software; (2) filter stop words and common words by establishing word lists; (3) use the C-value term extraction algorithm and automatic grouping functions integrated in ITGinsight software to perform semantic fuzzy matching for merging words with the same stem; (4) with the participation of domain experts, use term degree greater than 15 and word frequency greater than 10 as screening criteria for preliminary selection of domain core keywords; (5) semantic merging. For synonyms that cannot be automatically merged by machines, such as "intelligent furniture," "intelligent household appliance," and "smart home device," we establish merging rules based on expert knowledge and perform automatic merging through Python. After a series of cleaning steps, we ultimately obtained 1,053 domain core themes. The specific segmentation and cleaning process is shown in Table 3 .

4.2 Identification of Technological Innovation Combinations in Speech Recognition

To establish a semantic association network between technology themes, we used the cleaned domain text data as corpus input, trained a Word2Vec model using the gensim package, and transformed the screened 1,053 keyword groups into word vectors containing contextual information. It should be noted that since Word2Vec by default generates word vectors for single words and cannot handle phrases consisting of two or more words, this study treats term groups (phrases) as a whole for vector transformation, which preserves the semantic information of phrases and is more conducive to establishing semantically associated technology networks.

After constructing the basic semantic network based on phrase vectors, we used the CFDP method to explore core technology themes and associated technological innovation elements. Figure 3 [Figure 3: see original paper] (left panel) shows the density and distance distribution of 1,053 core technology points, where the horizontal axis represents density and the vertical axis represents distance. Nodes in the upper-right portion of the left panel indicate high centrality and large distances to other central nodes. In-depth analysis revealed that six nodes—"wearable device," "neural network algorithm," "voice control module," "speech recognition section," "multi-pass," and "speech recognition accuracy"—hold core positions in the speech recognition field. Since "speech recognition section" and "speech recognition accuracy" are closely related, we merged them into one core node.

Based on this, we further screened technological elements with combination potential around core nodes using the distance indicator. Points closer to core nodes are more likely to form potentially valuable technology combinations with them. With expert assistance, we identified 115 technology points and formed five potential technological innovation combinations based on their distance and density relationships: "voice control technology," "speech recognition model algorithms," "intelligent voice products and applications," "new scenarios for speech recognition technology," and "speech recognition technology performance," as shown in the right panel of Figure 3.

To deeply mine the innovative evolutionary relationships between elements within each technology combination, this study needed to fully understand the domain's technological composition and divide its basic modules. Through literature review and expert interviews, we divided the field into five sub-modules and constructed a domain conceptual model based on the composition and dependency relationships between modules, as shown in Figure 4 [Figure 4: see original paper]. The five sub-modules in this field are: support platforms and technologies, algorithms and models, technical performance, products and applications, and market demand. Among them, the "algorithms and models" sub-module is the core of the speech recognition field. New products and applications depend on algorithmic breakthroughs, while new products in turn meet specific market demands. Across all sub-modules, "support platforms and technologies" form the foundation, safeguarding industry technological innovation and product updates.

Based on this, we further grouped the technological elements in each combination according to their sub-module affiliation. This grouping approach effectively helps stakeholders understand domain technological evolution patterns and innovation mechanisms from a technological composition perspective. Table 4 systematically shows the grouping of technological elements in each combination.

4.3 Exploration of Combinatorial Innovation Approaches and Technological Opportunities

Using the technology combinations obtained in the previous stage as research objects, this section employs SAO theory to mine the technological evolutionary relationships between innovation elements within each combination. This study focuses on two innovation mechanisms: evolutionary innovation within group elements and fusion innovation brought about by inter-group technology fusion and application transfer.

Using the 115 innovation elements (phrases) obtained previously as input, we extracted short sentences containing two innovation elements from original texts using regular expression rules. For example, in the sentence "Microphone arrays capture the spatial characteristics of the sound field to improve voice signal quality," S is "microphone arrays" and O is "voice signal." We then used part-of-speech tagging methods from the nltk package to identify verbs in short texts, thereby simplifying them into SAO structures containing only two innovation elements (nouns) and their relationship (verb). For instance, the above example was ultimately simplified to "Microphone arrays capture voice signal." Through this processing, we obtained 14,548 SAO structures.

Based on domain investigation and expert discussion, we believe that in the speech recognition field, the innovation relationships of greatest concern to industry are four types: technological evolution and upgrading (promotion relationship), technology fusion (inclusion relationship), negative effect elimination (reduction relationship), and product application (application relationship). Due to the large volume of SAO structural data involved, we employed deep mining algorithms to explore the innovation relationships among the 115 core technological innovation elements. To ensure the accuracy of SAO relationship identification, we randomly selected 3,000 SAO structures for manual annotation as test and training sets for the classification algorithm. We then trained an LSTM model based on the manually annotated data, adjusting model parameters to improve prediction accuracy, with the final classification accuracy reaching 85%.

To verify the superiority of LSTM in short text classification performance, we measured the classification accuracy of four other commonly used classification algorithms (Naive Bayes, Logistic Regression, Support Vector Machine, and Random Forest) based on the same test data. The results are shown in Table 5 . The data in Table 5 demonstrate that compared with other machine learning algorithms, the LSTM algorithm improves classification accuracy for SAO evolutionary relationships by 10%-20%. Finally, we used the trained model to classify the remaining SAO structures. Table 6 shows partial classification results.

We visualized the innovation relationships among the 115 technological elements, with results shown in Figure 5 [Figure 5: see original paper]. In Figure 5, technology nodes from different clusters represent their affiliation with different technology combinations, and the thickness of connections between nodes represents the strength of evolutionary relationships between associated nodes. Figure 5 shows that thematic associations in the speech recognition field are high. This association is evident not only within groups but also significantly between groups. To deeply mine intra-group innovation and inter-group evolution patterns, combined with the technology grouping in Table 4, we conducted in-depth exploration of core technological innovation relationships.

Cluster 1 represents "speech recognition model algorithms," with related technologies divided into two sub-domains: (A-1) algorithm models and (A-2) industry demands and applications. Deep mining of SAO structures in this combination revealed that "speech recognition model" is the most fundamental and core technology in this combination, with SAO relationships related to it accounting for 65% of all relationships in this combination. "Language model," "acoustic model," and "pronunciation dictionary" are its main components, directly determining speech recognition effectiveness. According to investigations, with breakthroughs in speech recognition model performance, intelligent voice technology has entered a commercialization period, with domestic speech recognition leaders such as iFLYTEK, Sogou, and YITU Technology conducting in-depth explorations of speech recognition algorithm models and achieving outstanding results in international competitions. In the industry demand and application sub-domain, the SAO structure "neural network algorithm 'promotes' voice search" appears 45 times, indicating that neural network algorithms can enhance voice search effectiveness. Notably, artificial intelligence algorithms, voice interaction, and voice control technology have become focal points in this field in recent years.

We focused on the core technology transfer in this technology combination. Among the 1,522 patents involved in this combination, Google (352 patents) and Baidu (160 patents) are the two most representative patent holders. Analysis of related SAO structures revealed that the two companies have very similar patent layouts in this field. Google holds an advantageous position in overall patent quantity and comprehensive technological strength. However, in the deep learning algorithm domain, Baidu's patent quantity is basically on par with Google, leading in Chinese speech recognition and dialect recognition. Further investigation found that Baidu has narrowed the gap with Google by cooperating with the Institute of Acoustics of the Chinese Academy of Sciences and appointing AI expert Andrew Ng as the company's chief scientist. However, Google still leads Baidu in multilingual recognition, voice search, and voice input methods. Notably, another domestic enterprise with years of R&D in this field, iFLYTEK, also has a similar patent technology structure. Exploring a joint development model between Baidu and iFLYTEK would effectively enhance China's international competitive position in this field, as shown in Figure 6 [Figure 6: see original paper].

Cluster 2 represents "intelligent voice chip technology," with related innovation elements belonging to two sub-modules: software and hardware support (B-1) and products and applications (B-2). SAO semantic mining revealed that the two most important SAO relationships in this combination are: (1) "microphone array 'promotes' microphone module," which appears 37 times, indicating that microphone array technology can effectively improve voice signal acquisition capability of microphone modules and enhance speech recognition accuracy; (2) "MCU 'replaces' FPGA, applied in voice control module" (appearing 18 times), reflecting the upgrading of speech recognition chips and providing possibilities for new scenarios and applications. FPGA hardware offers advantages such as programmability, reconfigurability, and parallel computing, enhancing circuit performance, design flexibility, and efficiency of voice control devices, and has become an important platform for speech recognition system hardware acceleration. Although MCU hardware is less performant than FPGA, its lower power consumption and smaller size make it highly suitable for small embedded speech recognition devices, with mature applications in intelligent lights, smart door locks, and smart wearable devices in (B-2).

Notably, intelligent voice chips and integrated voice control technology have become focal points in the field in recent years. Among the 1,522 patents involved in this combination, Nvidia, Intel, Altera, Cambricon, Huawei, and Inspur hold the most domain patents. In-depth mining revealed that Nvidia firmly holds the discourse power in traditional GPU hardware acceleration technology, while Intel hopes to achieve transformation by exploring a "CPU+FPGA" technical solution through its acquisition of FPGA giant Altera. Facing the patent barriers built by Nvidia and Intel in the voice chip field, domestic intelligent chip manufacturer Cambricon has chosen to avoid more mature technical routes such as CPU and GPU, leveraging its own advantages to cooperate with Huawei, HiSilicon, Inspur, and the Chinese Academy of Sciences, seeking technological breakthroughs in MCU and Application-Specific Integrated Circuit (ASIC) domains. It is foreseeable that competition between traditional hardware manufacturers and new entrants in this technological field will be inevitable, as shown in Figure 7 [Figure 7: see original paper].

Cluster 3 represents "intelligent voice products and applications," with related technologies belonging to two sub-modules. (C-1) constitutes the software and hardware foundation for speech recognition technology, where the development of artificial intelligence and IoT technologies has greatly promoted the application of speech recognition technology in smart home devices, smart wearable devices, navigation equipment, self-service terminals (kiosks), and other products (C-2). In-depth mining revealed that the cross-fusion relationships between technological elements in this group and those in Cluster 2 (voice control technology) and Cluster 4 (new scenarios for speech recognition technology) are very significant.

For example, the development of voice control systems in Cluster 2 (B-2) has changed human-computer interaction methods, making speech recognition products an important entry point for controlling smart home devices such as smart air conditioners and televisions. Digital assistants, intelligent voice interaction, and other functions have also seen many commercial transformations on mobile terminals such as mobile phones, cameras, and wearable devices (C-2). The SAO structure "digital assistant 'applied to' terminal device" appears most frequently in Cluster 4 "speech recognition demands and scenarios," indicating that user scenarios of interacting with terminal devices through voice are becoming increasingly common. The development of speech recognition technology brings infinite possibilities for meeting new scenario demands and interaction experiences. Additionally, industry demands such as "multimodal," "multi-device," "multilingual," "multi-target," "real-time," and "feedback" appear frequently in this technology combination and are closely associated with Cluster 2 "intelligent voice products and applications" (C-2). It can thus be predicted that real-time speech recognition, multimodal speech recognition, and other technologies will spawn more intelligent products applying speech recognition technology to meet diverse demands in complex scenarios. Regarding currently popular speech recognition and interaction application technologies, in May 2021, the national standard plan "General Safety Technical Requirements for Intelligent Voice Controllers" jointly compiled by China Electrical Equipment Research Institute and Midea and other home appliance leading enterprises was approved, providing guarantees for the healthy development of voice controller-related industries.

Cluster 5 represents "speech recognition performance," containing two subdomains: (E-1) performance indicators and (E-2) industry demands. Through text semantic mining, the most frequently appearing SAO structure in this combination is "noise reduction technology 'enhances' speech recognition accuracy," which appears 24 times. This indicates that in complex scenarios, applying noise reduction technology during voice signal acquisition and processing stages can reduce the impact of background noise on voice quality, thereby improving speech recognition accuracy.

Additionally, the inter-group association between Cluster 5 and Cluster 2 in Figure 5 is also significant. Intelligent chips enhance voice signal processing speed, while algorithm models guarantee speech recognition accuracy. As speech recognition technology penetrates more industries, complex scenarios such as healthcare and transportation also impose higher requirements on various aspects of speech recognition technology performance. Technologies such as robust speech recognition, far-field speech recognition, and speech recognition in reverberant environments have great development potential. In recent years, national major R&D plans and provincial key R&D projects have provided funding support for relevant enterprise research projects, such as AISpeech Information Technology Co., Ltd. conducting research on "speech recognition in medical scenarios," iFLYTEK conducting research on "courtroom virtual assistant technology based on multi-person and multi-dialect speech recognition and judicial trial information resource databases," and the State Key Laboratory of Visual Synthesis Graphics and Image Technology of Sichuan University jointly conducting research with Sichuan Wisesoft Co., Ltd. on "key technologies and applications of air traffic control speech recognition and semantic understanding engines in complex environments."

Through in-depth mining of intra-group and inter-group innovation relationships in the speech recognition field, several meaningful phenomena warrant attention:

(1) Competition in the speech recognition field is relatively intense in the algorithm and model domain, with limited patent cooperation among manufacturers. Although domestic enterprises have taken the lead in some areas such as Chinese speech recognition, their overall technological strength still lags behind international leaders.

(2) In speech recognition software/hardware and product application do-

mains, speech recognition service providers cooperate closely with chip companies and traditional manufacturing industries, with the "cloud-chip" integration model becoming a development trend. The application fields of speech recognition technology are also becoming increasingly broad.

(3) Domestic speech recognition technology manufacturers should conduct cooperative innovation during the R&D stage to break through foreign technological blockades and market monopolies; during the technology application stage, they should strengthen cooperation with traditional manufacturing industries to jointly develop intelligent speech recognition products and applications for new scenarios and demands.

5 Conclusion

This study proposes a patent data-based method for identifying and tracking technological innovation combinations and verifies its effectiveness using the speech recognition field as an example. This method constructs a domain semantic network through the Word2Vec algorithm, uses the CFDP algorithm to obtain potentially associated technology combinations, identifies relationships between technology themes based on the SAO method, and achieves accurate classification of massive SAO structures by training an LSTM deep learning model. This enables exploration of combinatorial evolution patterns between innovation elements and prediction of potential technological opportunities. Based on this research framework, we conducted a case study in the speech recognition field, verifying the effectiveness of this method in identifying technological innovation combinations and approaches, while also providing strong support for future R&D and policy formulation in the speech recognition field.

This study has certain limitations: it only identifies and tracks innovation combinations from a technology-driven perspective without perfecting the identification model from a multi-factor-driven perspective. Additionally, this study only selects patent data for analysis, resulting in a relatively single dimension. Future research will expand and mine data from multiple channels including journal data, policy documents, and social media.

References

[1] Li Xiaoman, Zhang Xuefu, Song Hongyan, et al. Research on patent literature technology element identification methods: Taking the nanofertilizer field as an example[J]. Library and Information Service, 2020, 64(6): 59-68.

[2] Yuan Ye, Wu Chaonan, Li Qiuying. Analysis of international competition landscape for core technologies in the artificial intelligence industry[J]. Journal of China Academy of Electronics and Information Technology, 2020, 15(11): 1128-1138.

[3] Lu Xiaobin, Yang Guancan, Xu Shuo, et al. Review of research progress on

emerging technology identification from metrological and evolutionary perspectives[J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(6): 651-661.

[4] Du Jian, Sun Yinan, Li Yongjie, et al. Identifying innovation frontiers from the intersection of science and technology: Methods and empirical evidence[J]. Information Studies: Theory & Application, 2019, 42(1): 94-99.

[5] Liu J, Wei J, Liu Y. Technology forecasting based on topical analysis and social network analysis: A case study focusing on gene editing patents[J]. Journal of Scientific and Industrial Research, 2021, 80(5): 428-437.

[6] Schoenmakers W, Duysters G. The technological origins of radical inventions[J]. Research Policy, 2010, 39(8): 1051-1059.

[7] Karvonen M, Kassi T. Patent citations as a tool for analysing the early stages of convergence[J]. Technological Forecasting and Social Change, 2013, 80(6): 1094-1107.

[8] Schumpeter J, Backhaus U. The theory of economic development[M]. Boston: Springer, 2003.

[9] Arthur W B. The nature of technology: What it is and how it evolves[M]. New York: Simon and Schuster, 2009.

[10] Sun Bing. Review of research on technological innovation motivation mechanisms[J]. East China Economic Management, 2010, 24(4): 143-147.

[11] Chi Xueqin, Sun Shuxian, Wei Xueyan, et al. Technological combinatorial innovation and laboratory medicine[J]. Industrial Enterprise Medical Journal, 2013(3): 276.

[12] Wu Hong. Analysis of combinatorial patterns in technological invention: Also discussing that the C919 large aircraft belongs to Chinese innovation[J]. Journal of Wuhan University of Science and Technology (Social Science Edition), 2018, 20(4): 456-460.

[13] Zhou Liying, Leng Fuhai, Zuo Wenge. Research on patent portfolio identification methods based on innovation teams[J]. Information Studies: Theory & Application, 2016, 39(11): 116-120.

[14] Zhang Zhengang, Luo Taiye. Technology opportunity discovery based on knowledge combination theory[J]. Science Research Management, 2020, 41(8): 220-228.

[15] Wang Xianwen, Xu Shenmeng, Peng Lian, et al. Research on technology network structure based on patent co-classification analysis: 1971-2010[J]. Journal of the China Society for Scientific and Technical Information, 2013, 32(2): 198-205.

[16] Zhou X, Huang L, Zhang Y, et al. A hybrid approach to detecting technological recombination based on text mining and patent network analysis[J].

Scientometrics, 2019, 121(2): 699-737.

[17] Li Xin, Wang Jingjing, Yang Zi, et al. Research on emerging technology identification based on SAO structural semantic analysis[J]. Journal of Intelligence, 2016, 35(3): 80-84.

[18] Miao Hong, Wang Yan, Huang Lucheng, et al. Research on technology fusion trends oriented toward functions[J]. Journal of Intelligence, 2020, 39(5): 51-58.

[19] Ma T, Zhou X, Liu J, et al. Combining topic modeling and SAO semantic analysis to identify technological opportunities of emerging technologies[J]. Technological Forecasting and Social Change, 2021, 173: 121159.

[20] Huang Lu, Zhu Yihe, Zhang Yi. Research on emerging technology theme identification based on weighted network link prediction[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(4): 335-341.

[21] Liu Yuqin, Wang Xuefeng, Lei Xiaoping. Design and implementation of a scientific research relationship construction and visualization system[J]. Library and Information Service, 2015, 59(8): 103-125.

[22] Wang X F, Zhang S, Liu Y Q. ITGInsight-discovering and visualizing science, technology and innovation information for generating competitive technological intelligence[C]//Proceedings of the 1st workshop on AI+ informetrics (AII2 2021) co-located with the iConference. Aachen: RWTH Aachen University, 2021: 202-219.

[23] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[J]. ArXiv, 2013, abs/1301.3781.

[24] Rodriguez A, Laio A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.

[25] Li Wenhui, Zhang Yingjun, Pan Lihu. Improved biLSTM network for short text classification[J]. Computer Engineering and Design, 2020, 41(3): 880-886.

[26] Shanghai iResearch Consulting Co., Ltd. Gazing at the bright starry river: China intelligent voice industry research report 2020[C]//iResearch consulting series research reports (2020 No. 2). Shanghai: Shanghai iResearch Consulting Co., Ltd., 2020: 81-127.

[27] Ying Zhao, Mai Yongfeng, Wang Yanmin. Development prospects of smart home voice control systems[J]. Intelligent Building Electrical Technology, 2019, 13(1): 30-33.

[28] Li Yang. Research on sound-based foreign object detection technology in electric energy meters[D]. Mianyang: Southwest University of Science and Technology, 2019.

[29] Wu Yanxia, Liang Kai, Liu Ying, et al. Progress and trends of deep learning FPGA accelerators[J]. Chinese Journal of Computers, 2019, 42(11): 2461-2480.

[30] Laurie A H. Artificial intelligence: Background, selected issues, and policy considerations[R]. Washington: Congressional Research Service, 2021.

[31] Gao Dongmei. Cambricon Technologies: China's "chip" hope?[EB/OL]. [2022-03-28]. https://tech.sina.com.cn/csj/2019-04-06/doc-ihvhiiewr3525252.shtml.

[32] Standardization Administration of China. National public service platform for standard information[EB/OL]. [2022-03-28]. http://std.samr.gov.cn/gb/search/gbDetailed?id=A365251D80

[33] Liu Qingfeng, Gao Jianqing, Wan Genshun. Research progress and challenges in speech recognition technology[J]. Frontiers of Data & Computing, 2019, 1(6): 26-36.

[34] Institute of Scientific and Technical Information of China. National science and technology report service system[EB/OL]. [2022-03-28]. https://www.nstrs.cn/index.

Author Contributions: Zhou Xiao: Conceptualized the research, constructed models, and wrote the paper; Xu Yinbiao: Analyzed data, constructed models, and wrote the paper; Shi Yi: Constructed models and analyzed data.

Identifying and Tracing Technological Innovation Combination Based on Deep Learning and Semantic Mining

Zhou Xiao[1], Xu Yinbiao[1], Shi Yi[2]

[1]School of Economics and Management, Xidian University, Xi'an 710126
[2]College of Computer Science, Xi'an Shiyou University, Xi'an 710065

Abstract: [Purpose/Significance] With the rapid development of strategic emerging technology industries, identifying technological innovation combinations with potential synergistic effects and clarifying core innovation relationships within combinations is an important prerequisite for effectively planning industrial development routes and enhancing industrial competitive advantages. [Method/Process] Guided by technology portfolio evolution theory and combining deep learning, SAO semantic mining, and the CFDP algorithm, this paper proposes a patent data-based scheme for identifying technological innovation combinations and evolutionary relationships. The research protocol is divided into three steps: First, design a domain search strategy based on keywords and patent classification numbers, and complete data cleaning and segmentation. Second, build a word vector semantic network of domain technology topics through Word2Vec, and use the CFDP algorithm to identify potential innovation elements and combination methods. Finally, deeply mine core SAO structures in each combination, classify their evolutionary relationships through the LSTM deep learning algorithm, explore core innovation approaches, and thereby effectively identify potential technological opportunities in the field. [Result/Conclusion] Taking the speech recognition field as an example, through in-depth mining of DII patent text data in this field, five potential technological innovation combinations and core innovation methods are identified and tracked. The study finds that China's current speech recognition field has great innovation potential in smart chip design, speech recognition algorithms, new scenarios, and applications.

Keywords: technological innovation combination identification; deep learning; SAO; semantic mining; patent analysis

Note: Figure translations are in progress. See original paper for figures.