energies

MDPI

Article

Discovery and Analysis of Key Core Technology Topics in Proton Exchange Membrane Fuel Cells Through the BERTopic Model

Yurong Gou [1,2] and Qimei Chen [1,2,*]

[1] National Science Library, Chinese Academy of Sciences, Beijing 100190, China; gouyurong@mail.las.ac.cn
[2] University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: chenqm@mail.las.ac.cn; Tel.: +86-1861-031-3267

Abstract: As a core component of clean energy technology, proton exchange membrane fuel cells (PEMFC) play a crucial role in promoting the evolution of energy structures and realizing sustainable development, representing an environmentally friendly energy conversion strategy. This paper identifies the key core technology themes in the field of the proton exchange membrane fuel cells by constructing patent and paper datasets in the field, applying the BERTopic model for theme identification, and calculating the key core technology scores of each theme using the importance, innovativeness, and high competitiveness barriers to identify the key core technology themes in the field, so as to provide guidance and references for the relevant research and practice. The results of the study show that patent documents and academic papers show obvious differentiation in technical themes: the key core technologies identified in patent texts include 'battery separator materials', 'rubber sealing materials', and 'porous carbon fibre materials'. The key core technologies identified in the academic paper of the thesis include 'palladium-based electrocatalys', 'graphene oxide composite film', and 'platinum-graphene oxide catalyst'.

Keywords: key core technology; the proton exchange membrane fuel cell; index system

check for updates

Citation: Gou, Y.; Chen, Q. Discovery and Analysis of Key Core Technology Topics in Proton Exchange Membrane Fuel Cells Through the BERTopic Model. *Energies* **2024**, *17*, 5418. https://doi.org/10.3390/en17215418

Academic Editor: Antonino S. Aricò

Received: 3 September 2024
Revised: 14 October 2024
Accepted: 18 October 2024
Published: 30 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The proton exchange membrane fuel cell (PEMFC) is currently one of the most mainstream fuel cell types in China, featuring high efficiency, fast startup, low temperature operation, and higher power density [1], and it has a wide range of potential in many applications.

From the perspective of identification methods, the key core technology identification methods can be roughly divided into five kinds of methods, method namely based on complex networks, methods based on indicators, methods based on machine learning, methods based on expert wisdom, and methods based on composite classes [2]. Complex network-based methods mainly refer to the use of text to establish a complex knowledge network by analyzing various attributes and contents in the network to identify technologies in the field [3,4], and the current commonly used methods include citation network analysis [5], network structure feature analysis [6], network main path analysis [7], and K-kernel analysis. Indicator-based methods are mainly used to identify technologies in the field by selecting external features of documents that can represent technological development as the external features of the literature, and determine the technology to be identified through the calculation and comparison of the indicators [8–10]. The method based on machine learning refers to the use of various types of machine learning methods to extract and train the features of data and then complete the technology identification [11,12]. The method based on the wisdom of experts refers to the method of filtering out specific technologies based on the empirical knowledge of domain experts through the sorting to screen out specific technologies, including the Delphi method and hierarchical analysis method [13], in order to better solve complex decision-making problems. Combining the Delphi method with the hierarchical analysis method [13] and adopting the fuzzy hierarchical analysis

method [14] have also been increasingly applied in the field of key core technology identification [15,16]. A composite class of methods refers to the combination of different identification methods to mitigate the defects of single identification methods, such as combining LDA models with other methods [17] or combining complex network-based methods with indicator-based methods [18].

After reading, it was found that most existing studies adopt a single perspective, and few studies synthesize multiple perspectives to identify complex scientific research topics while taking into account the differences in the theme performance of different data sources. In addition, most of the existing studies use various literature measurement indicators and network centrality indicators to identify key core technologies, but there is no effective measurement of the high competition barriers of technologies, especially the lack of measurement from the time dimension, and it is rare to comprehensively determine the weight of indicators after considering the emphasis of different methods on indicator measurement. Therefore, through the definition and feature discrimination of key core technologies, this paper extracts the technical characteristics of their importance, innovation, and high competition barriers, combines patents and papers to design discrimination indicators and measurement methods, takes into account the variability and correlation of indicators, and uses a variety of methods to give weight to build a relatively complete key core technology identification system.

This paper selects patent and paper results in the field of the proton exchange membrane fuel cell since 2001 as the research object and constructs a suitable index system to identify the key core technologies in this field.

2. Study Design

2.1. Theme Identification

BERTopic is a topic model based on Bert embedding, which makes full use of Bert's semantic expression ability and contextual understanding ability to obtain semantically relevant topics in large-scale text data. BERTopic does not need to pre-determine the number of topics and is able to discover some niche topics in complex text. The specific process includes the following: (1) embedding word summaries; (2) reducing dimensionality and clustering; and (3) creating topic representation based on the C-TF-IDF algorithm, as shown in Figure 1.

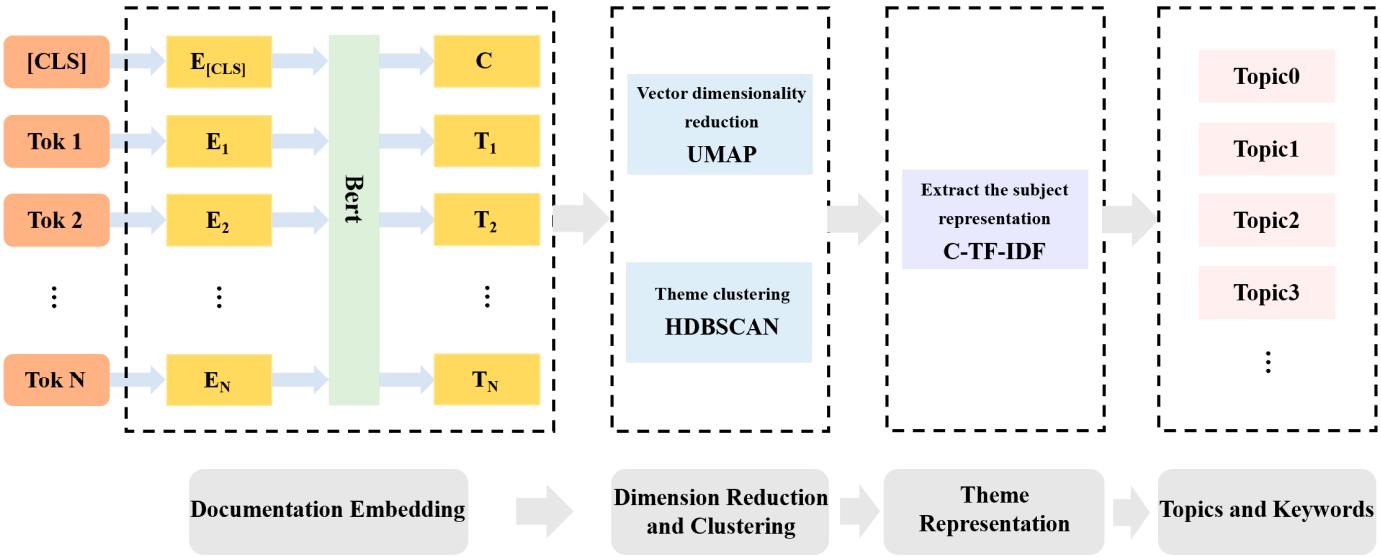


Figure 1. BERTopic theme modelling process.

2.2. Construction of the Indicator System

At present, there are various terminologies and definitions regarding 'what constitutes key core technologies', and the scope of coverage is relatively broad. There is no complete consensus on the 'knowledge-intensity' [19,20], 'complexity of technology system' [21] and '(quasi) public goods' [22,23] of the key core technology, but many studies agree that the key core technology should be 'in the core position, difficult to replace, play an important role and indispensable technology'. Therefore, this paper defines key core technology from two dimensions of the industrial chain and technical attributes as the technology that occupies an important position in a certain technical field, which is generally innovative has high competitive barriers, and builds a key core technology indicator identification system based on this.

2.2.1. Importance

This paper selects the node crossover degree and time-weighted citation frequency to calculate the importance.

(1) Node Crossover Degree

This paper introduces the node cross degree index (*NC*) [24] using the node's in-degree and out-degree to measure the role of the patent/paper node network status, reflecting the node's importance status in the knowledge network, which is calculated by the following formula:

$$NC_i = \frac{\sum_{j=1}^{n} (k_{in}(j)+1)^{\alpha} (k_{out}(j)+1)^{1-\alpha} - 1}{n}$$

where n is the number of patents/papers in subject i, j is the jth patent/paper in subject i, $k_{in}(j)$ is the incidence of the j_{th} patent/paper, $k_{out}(j)$ is the out-degree of the jth patent/paper, and α is a quantity constant.

(2) Time-Weighted Citation Frequency

The measurement of topic impact is a key issue in measuring the importance of topics [25]. Due to the existence of time accumulation in the number of citations, this paper incorporates the time weight parameter in the measurement of theme influence and sets the time weight as t_i to calculate the time-weighted citation frequency (*TWCF*), which is given in the following formula:

$$TWCF_i = \frac{\sum_{j=1}^{n} (t_j \times C_j)}{n}$$

$$t_j = \frac{2s}{y \times (y+1)}$$

where n is the number of patents/papers in subject i, j is the jth patent/paper in subject i, C_j is the citation frequency of the jth patent/paper, t_j is the time weight of the jth patent/paper, s is the quantity constant, and y is the year span of the dataset.

2.2.2. Innovativeness

In technology evaluation, the assessment of technological innovativeness is an important basis for identifying key core technologies [26]. In this paper, the average public/publication year and similarity to existing topics are used to analyze the innovativeness of technology in terms of time and citation, respectively.

(1) Average Public/Publication Year

Generally speaking, recently disclosed/published patents/papers usually have a high degree of novelty in terms of time, so the technical innovativeness in the time dimension

can be expressed by the average year of occurrence (*AYO*) of the relevant paper, which is calculated by the following formula:

$$AYO_i = \frac{\sum_{j=1}^{n} YO_j}{n}$$

where n is the number of patents/papers in subject i, j is the jth patent/paper in subject i, and YO_j is the year of disclosure/publication of the jth patent/paper.

(2) Similarity to Existing Topics

Text similarity reflects the degree of similarity between the semantic connotations of the two texts and the size of the difference; the higher the similarity, the stronger the relevance of the content contained in the two texts. In this paper, the cosine similarity algorithm is used to measure the similarity between retrieved technical topics and existing technical topics (*TTS*) by calculating the angle between two vectors. The specific calculation formula is as follows:

$$TTS_{ab} = \frac{\sum_{1}^{n+m} c_i}{\sqrt{\sum_{1}^{n+m} c_i^2} \times \sqrt{\sum_{1}^{m+n} c_i^2}} (c = a + b)$$

where a is a vector of the retrieved patent/paper's textual content, b is a vector of the reference textual content, n is the number of patents/papers in subject i, and m is the number of reference texts corresponding to the patent/paper.

2.2.3. High Competitiveness Barriers

This paper introduces the concept of a "monopoly industry" into the field of technology identification and defines a "monopoly technology subject" as follows: Due to the existence of technical barriers, only a small number of patentees/institutions have mastered the technology and applied for/published relevant patents/papers. Based on this, this paper tries to measure the high competitive barriers of technical subject matter from the perspective of technical monopoly, in terms of monopoly intensity and monopoly duration.

(1) Monopoly Year

In economics, the Concentration Rate 4 (*CR*4), an indicator of the concentration of the top four shares of an industry, can be used to classify the degree of competition and monopoly of an industry.

Drawing on the Bain's classification of the market structure, this paper considers technical topics with a $CR4 \geq 30\%$ as non-competitive technical topics, and obtains the time when the technical topic belongs to non-competitive technical topics, i.e., monopoly year (*MY*), by calculating the number of years in which the *CR*4 is greater than or equal to 30%, as shown in the specific calculation formula below:

$$CR4_i = \frac{n_{i1} + n_{i2} + n_{i3} + n_{i4}}{\sum_{p=1}^{m} n_{ip}}$$

$$MY_i = year(CR4_i \geq 30\%)$$

where n_{i1}, n_{i2}, n_{i3}, and n_{i4} denote the institution-specific disclosure/publication numbers of the top four institutions with the highest number of patent disclosures/publications in subject i, n_{ip} is the number of patent disclosures/paper publications of the p_{th} organization under the i_{th} technical subject, and m is the number of organizations included under the i_{th} technical subject.

(2) Topic Concentration

The Herfindahl–Hirschman Index is a comprehensive index that expresses the level of industrial concentration by the sum of the squares of the market shares of all enterprises in

a particular market. Some scholars draw on this index [27] to determine whether most of the patents/papers are held in a few institutions by calculating the sum of the squares of the patent disclosures/papers published by each institution under a certain topic over the total number of patent disclosures/papers published by all the institutions under the topic, i.e., topic concentration (*TC*), which is calculated by the following formula:

$$TC_i = \sum_{p=1}^{m} (\frac{n_{ip}}{n_i})^2$$

where TC_i denotes the theme concentration index for the ith technology theme, n_{ip} is the number of patent disclosures/paper publications of the pth organization under the ith technical subject, n_i is the number of all patents/papers under the ith technical subject, and m is the number of organizations included under the ith technical subject.

The above indicators can be summarized as shown in Table 1:

Table 1. Key core technology identification indicator system.

Character	Perspective	Indicator	Formula
Importance	Status	Node Crossover Degree	$NC_i = \frac{\sum_{j=1}^{n} (k_{in}(j)+1)^{\alpha} (k_{out}(j)+1)^{1-\alpha}-1}{n}$
	Influence	Time-Weighted Citation Frequency	$TWCF_i = \frac{\sum_{j=1}^{n} (t_j \times C_j)}{n}$
Innovativeness	Time	Average Public/Publication Year	$AYO_i = \frac{\sum_{j=1}^{n} YO_j}{n}$
	Element	Similarity to Existing Topics	$TTS_{ab} = \frac{\sum_{1}^{n+m} c_i}{\sqrt{\sum_{1}^{n+m} c_i^2} \times \sqrt{\sum_{1}^{m+n} c_i^2}}$
High competitiveness barrier	Monopoly period	Monopoly Year	$MY_i = year(\frac{n_{i1}+n_{i2}+n_{i3}+n_{i4}}{\sum_{p=1}^{m} n_{ip}} \geq 30\%)$
	Monopoly power	Topic Concentration	$TC_i = \sum_{p=1}^{m} (\frac{n_{ip}}{n_i})^2$

3. Empirical Research

3.1. Data Acquisition and Preprocessing

3.1.1. Data Acquisition

The patent data were obtained from Derwent Patent Database (DII). By locating the relevant information in the field of proton exchange membrane (PEM) fuel cells and combining the opinions of experts, we searched the patent data with the search terms 'Proton exchange membrane* fuel cell*', 'polymer electrolyte membrane* fuel cell*', 'Polymer electrolyte* fuel cell*', etc., and combined them with the Derwent manual codes L03-E04A2 and L03-E04F to search the patent data. Considering that the number of citations of patents is commonly used to measure the value of patents in practice [3], in this paper we attempt to initially screen the patent data from the number of citations to obtain a high-quality patent dataset. Through reading the relevant literature, existing research suggests that the value of patents conforms to the 'law of two or eight' in the field of sociology, and the frequency of citations also conforms to this law [4]. Therefore, based on the above research results, in this paper we take the top 20% of patent records with the citation frequency in each year after 2001, and after removing the blank and seriously incomplete data, a total of 5683 patent data points are obtained, which lays the foundation for the subsequent research.

The paper data were obtained from the Web of Science platform (WOS) by mining the relevant information in the field of proton exchange membrane fuel cells and synthesizing experts' opinions, and the terms 'Proton exchange membrane* fuel cell*', 'polymer electrolyte membrane* fuel cell*', 'Polymer electrolyte* fuel cell*', etc., were used as search terms. According to the conclusion of related research, the low citation criterion defined by the 'law of two or eight' can effectively distinguish the academic value of papers, and the derived results are more stable [5]. Considering that the citation frequency as an objective index for evaluating academic quality has become a consensus in the academic community [6], in this paper we also adopt the 'law of two-eight' to select the top 20%

of the papers with the citation frequency in each year after 2001 and remove the blanks and seriously incomplete data; a total of 7203 papers were obtained for the subsequent experimental study.

3.1.2. Data Preprocessing

The title and abstract of a patent record the main technical programme and the technical effect achieved by the patent, so in this paper we extract the method, use, and novelty fields in the title and abstract of each patent and then splice them together to form the text content field of the patent. The title, abstract, and keywords of a paper summarize the core theme, research purpose, and research results of the paper, so in this paper we also extract and splice the title, abstract, and keyword fields of each paper to form the content field of the paper.

After constructing the text dataset, in this paper we carry out preprocessing operations including word splitting, synonym replacement, the deletion of deactivated words using common deactivation lists and self-constructed deactivation lists, the removal of punctuation marks, lexical reduction, etc., and finally obtain the text corpus to be trained.

3.2. Topic Identification

3.2.1. Patent Topic Identification

Based on the patent corpus obtained by preprocessing, using the BERTopic model for topic identification, our results suggest that the research topics of patents in the field of proton exchange membrane fuel cells mainly include the following categories, as shown in Table 2:

Table 2. Patent research topics in the field of proton exchange membrane fuel cells.

Level 1 Technology	Level 2 Technology	Level 3 Technology	Percentage (%)
Stack	Membrane electrode	Polymer membrane	27.06
		Platinum carbon catalyst	22.02
		Electrode catalyst	5.39
		Porous carbon fibre material	4.79
		Graphene-Pt catalyst	0.76
		Sulphonated graphene oxide preparation of the proton exchange membranes	0.62
	Bipolar plate	Bipolar plate seal	7.38
		Bipolar plate corrosion resistance	1.09
		Graphite bipolar plate	0.56
	Seal	Rubber sealing material	0.60
	Other material	Battery separator material	3.45
		Battery electrode	0.46
	Algorithmic model	Parameter prediction modelling	1.02
System	Hydrothermal management system	Air supply and cooling system	3.28
		Hollow fibre membranes for humidifier	0.65
	Fuel system	Methanol fuel supply system	0.65
Practical application	Hydrogen fuel cell vehicle		18.69

3.2.2. Paper Topic Identification

Based on the paper corpus obtained by preprocessing, using the BERTopic model for topic identification, we suggest that the research topics of patents in the field of proton exchange membrane fuel cells mainly include the following categories, as shown in Table 3:

Table 3. Paper research topics in the field of proton exchange membrane fuel cells.

Level 1 Technology	Level 2 Technology	Level 3 Technology	Percentage (%)
PEMFC system hydrogen generation/storage	Hydrogen generation	Steam reforming to hydrogen	0.93
	Hydrogen storage	Metal hydride hydrogen storage	0.84
Stack	Membrane electrode	Platinum carbon catalyst	37.54
		Nafion composite proton conduction membrane	19.14
		Gas diffusion and microporous layer	6.62
		Platinum-reduced graphene oxide catalyst	4.48
		Membrane electrode assembly degradation	1.94
		Graphene oxide composite membrane	1.63
		Methanol cross effect	1.59
		Non-precious metal catalyst	0.99
		Palladium-based electrocatalyst	0.72
	Bipolar plate	Corrosion resistance of the bipolar plate	2.18
		Graphite bipolar plate	0.91
	Algorithmic model	Temperature influence model	12.26
		Fault detection diagnostic model	0.62
	Testing technology	Neutron radiography	0.67
System	Energy management	Energy management of hybrid system	4.15
	Algorithmic model	Battery performance improvement model with metal foam as dispenser	0.36
Practical application	Hydrogen fuel cell vehicle		2.44

3.2.3. Comparative Analysis

It is well known that a patent is the innovative result of technology at the application level, while a paper is the research result of technology at the scientific level, and there is a certain difference between the two in terms of their focus on technical topics. This study uses a butterfly diagram to visualize and compare the recognition results, as shown in Figure 2.

As can be seen from Figure 2, in terms of the focus of the research topics, there are some differences in the topics and the percentage of the patent texts and paper texts identified under each classification, with the most researched topic in the patents being polymer membranes and most of the papers' research focusing on platinum–carbon catalysts; in terms of the content of the research, the patents are generally more inclined to solving the problems of practical applications or focusing on the future commercial potentials of the technologies, such as material improvement and cost reduction, while papers focus more on basic scientific research, such as the development of new materials or methods and performance optimization.

3.3. Index Calculation

After calculating the importance, innovativeness, and high competitive barriers of 17 patent research topics and 19 paper research topics in the field of the proton exchange membrane fuel cell, respectively, the Entropy Method and the Criteria Importance Through Intercriterion Correlation (CRITIC) assign different weights to the three indexes, and the overall assessment is made by applying the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). The comprehensive scores of key core technologies for each theme are shown in Tables 4 and 5.

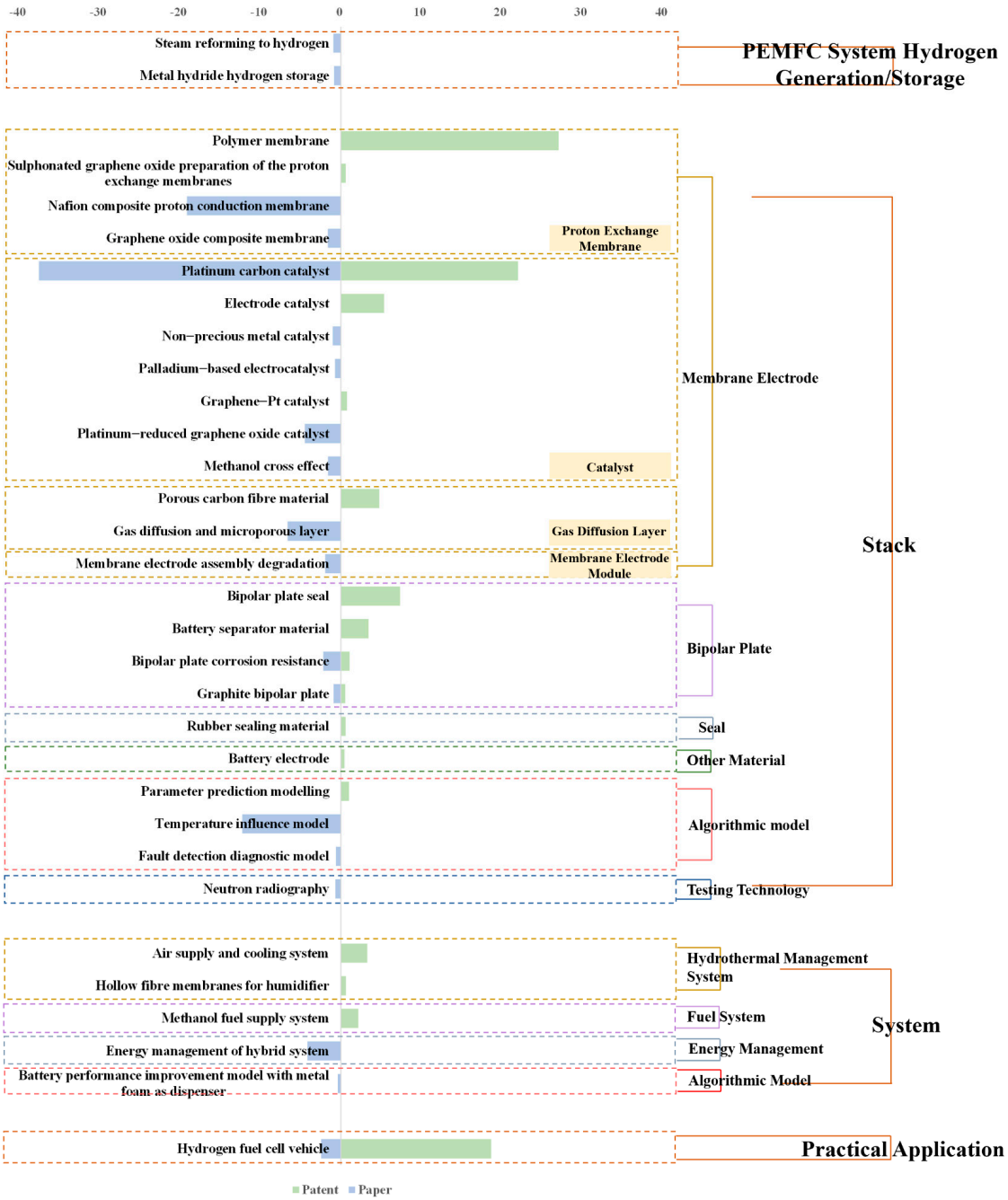


Figure 2. Visualization comparison of patent and paper identification results.

Table 4. Patent text key core technology topic composite score.

No.	Overall Score	No.	Overall Score
Topic 0	0.405307	Topic 9	0.380223
Topic 1	0.380850	Topic10	0.479992
Topic 2	0.373719	Topic11	0.422042
Topic 3	0.423837	Topic12	0.402619
Topic 4	0.438897	Topic13	0.426572
Topic 5	0.490420	Topic14	0.497500
Topic 6	0.529310	Topic15	0.312286
Topic 7	0.325339	Topic16	0.349631
Topic 8	0.356908		

Table 5. Paper text key core technology topic composite score.

No.	Overall Score	No.	Overall Score
Topic 0	0.501489	Topic10	0.519005
Topic 1	0.426902	Topic11	0.399432
Topic 2	0.381690	Topic12	0.463090
Topic 3	0.410762	Topic13	0.395977
Topic 4	0.519146	Topic14	0.400389
Topic 5	0.391863	Topic15	0.588108
Topic 6	0.505548	Topic16	0.433947
Topic 7	0.398494	Topic17	0.516251
Topic 8	0.473416	Topic18	0.513738
Topic 9	0.553236		

The TOP3 key core technology topics in the patent text in the field of proton exchange membrane fuel cells, namely Topic6 (cell spacer materials), Topic14 (rubber sealing materials), and Topic5 (porous carbon fibre materials), were selected for specific analysis.

Battery spacer material is a key component used inside a battery to provide isolation between the cathode and anode of the battery to prevent short circuits and combustion explosions, as well as to allow ionic movement to enable the charging and discharging processes of the battery, and it plays an important role in ionic conduction, electronic insulation, and maintenance of the chemical and thermal stability of the battery. The battery spacer material covered in this article mainly refers to the bipolar plate, and a review of the title and summary information of each patent under Topic6 reveals that the topic focuses on the formation of separator materials with excellent properties such as high strength, high corrosion resistance, and high electrical conductivity.

Fuel cell sealing is an important link to ensure the performance and safety of the fuel cell, and sealing material plays a crucial role in the fuel cell, not only to provide a gas barrier, but also due to the need to have environmental insulation, a rubber elastomer, coolant resistance, etc., and the optimization of their performance and application process is essential to ensure the long-term stable operation of the battery system. The rubber sealing material involved in this study mainly includes sealing rubber compositions for proton exchange membrane fuel cell diaphragms, in addition to reaction-curing silicone rubber adhesive compositions and sealing materials consisting of silicone rubber compositions and layered double hydroxide compounds.

In proton exchange membrane fuel cells, porous carbon fibre material is usually used as part of the gas diffusion layer, which mainly consists of sheets containing randomly dispersed carbon staple fibres and carbide resins with better electrical conductivity and higher production efficiency, and can be regarded as a key core technology in proton exchange membrane fuel cells. The porous carbon fibre material involved in this study is mostly used to fabricate gas diffusion layers with high tensile strength, good compressibility and durability, and excellent mechanical properties.

The TOP3 key core technology topics in the paper text in the field of proton exchange membrane fuel cells, namely Topic15 (palladium-based electrocatalysts), Topic9 (graphene oxide composite membranes), and Topic4 (platinum-graphene oxide catalysts), were selected for specific analyses.

Electrocatalysts are an important factor in determining the efficiency and cost of fuel cell devices, and the preparation of catalysts with high catalytic activity and low prices is key to the wide application of fuel cells. Generally speaking, electrocatalysts can be divided into metal catalysts and non-metal catalysts, and metal catalysts can be subdivided into precious metal catalysts and non-precious metal catalysts. Palladium-based electrocatalysts, as one kind of noble metal catalysts, can effectively solve the problems of high cost and limited reserves of platinum-based catalysts, and belong to the key technology in the field of proton exchange membrane fuels. The related papers involved in this study mainly investigated the excellent performance of palladium-based electrocatalysts

in portable power applications and how to improve the performance of palladium-based electrocatalysts.

Graphene oxide can be compounded with ionic polymers due to the presence of reactive oxygen functional groups to prepare various types of composite proton exchange membranes including graphene oxide nanocomposite membranes, graphene oxide/Nafion composite membranes, and polybenzimidazole/graphene oxide composite membranes [28]. Graphene oxide-based composite proton exchange membranes can improve the proton conductivity of the membrane under high-temperature and low-humidity conditions, reduce the methanol permeability, and improve the power density of the cell when applied to fuel cells [29], which is a key technology in proton exchange membrane fuel cells. The graphene oxide composite membrane paper involved in this study mainly focuses on the innovation of the composite membrane preparation method and its advantages in enhancing the battery performance and other aspects.

Graphene-loaded noble metal nanoparticles are one of the most widely used classes of electrocatalysts for fuel cells, and high dispersion of metal nano-catalysts can be achieved by depositing platinum and low-platinum nanomaterials on graphite nanosheets. Most of the papers under this theme in this study focus on topics such as methods for forming stable metal–metal oxide–graphene triple junctions and the preparation of graphene-loaded Pt and Pt alloy electrocatalysts.

Through the above analysis, it can be found that in the field of PEMFCs, there are some differences in the focus of key core technologies between patents and papers. Generally speaking, patents are more inclined to solve problems in practical applications or focus on the future commercial potential of the technology, such as material improvement and cost reduction, while papers are more focused on basic scientific research, the development of new materials or new methods, and performance optimization.

Specifically, the battery spacer materials identified in the patents mainly refer to the bipolar plates, the core component of the proton exchange membrane fuel cell, and it can be seen from the specific content that the patents usually focus on improving the conductivity and corrosion resistance of the spacer materials, as well as reducing the overall cost and improving the stability of the battery; the papers pay more attention to another core component—the proton exchange membrane—and focus more on the development of new materials, the exploration of new methods, and the optimization of the battery performance. Meanwhile, catalysts are an important component of PEMFCs; how to prepare better catalysts and how to improve the chemical performance of catalysts are also the key technologies in the current papers.

In the patents, the rubber sealing material is an important part of the practical application of the battery due to its key role in ensuring the sealing of the battery and preventing gas leakage, while the papers discuss relatively little about the rubber sealing material; the patents focus more on the application of different materials in the battery, such as the application of porous carbon fibre materials in the gas diffusion layer, which are chosen due to their high electrical conductivity and good mechanical properties, while the papers focus more on the microstructure design and optimization of the materials.

In addition, the algorithmic model category is more likely to become a key core technology identified in the paper text, such as fault detection diagnostic models, for which research focuses more on improving the diagnostic prediction accuracy to better troubleshoot or predict the causes of battery performance degradation or failure.

4. Conclusions

In this paper, the BERTopic theme model is applied to achieve theme discovery and a set of key core technology theme identification index systems is constructed. In the key core technology identification of proton exchange membrane fuel cells, research themes with importance, innovativeness, and high competitiveness barriers are successfully identified by designing the comprehensive score index of key core technology, and the identification results of patents and papers are compared and analyzed.

In the future, with the growing global demand for clean energy, research on proton exchange membrane fuel cells will focus more on performance improvement, cost reduction, and the development of new materials. At the same time, with the further development of artificial intelligence and machine learning technologies, future research methods will be more efficient and precise. In addition, as global environmental policies continue to tighten, the role of PEMFCs in alternative energy will become more and more important, and their applications in electric vehicles, static power stations, and portable power supplies will be further expanded.

Author Contributions: Research framework design, experiment implementation, paper writing, and revision, Y.G.; putting forward research ideas and suggestions for revision, revising and approving the paper, Q.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Loreti, G.; Facci, A.; Ubertini, S. High-Efficiency Combined Heat and Power through a High-Temperature Polymer Electrolyte Membrane Fuel Cell and Gas Turbine Hybrid System. *Sustainability* **2021**, *13*, 12515. [CrossRef]
2. Zhou, B.; Leng, F. Advances in technology identification research. *Adv. Inf. Sci.* **2022**, *14*, 315–348.
3. Xu, X.; Wu, F.; Wang, B. Research on key core technology identification based on international patent classification. *J. Intell.* **2022**, *41*, 74–81.
4. Wang, Y.; Han, T.; Chen, F. Research on key technology identification method integrating literature knowledge clustering and complex network. *Libr. Inf. Serv.* **2020**, *64*, 105–113.
5. Tian, J.; Qi, Y.; Lv, W.; Zhang, J. Research on core technology identification of intelligent networked automobile industry based on patent big data mining. *China Invent. Pat.* **2023**, *20*, 4–9+17.
6. Luke, D.; Harris, J. Network analysis in public health: History, methods, and applications. *Annu. Rev. Public Health* **2007**, *28*, 69–93. [CrossRef]
7. Hummon, N.; Doreian, P. Connectivity in a Citation Network—The Development of Dna Theory. *Soc. Netw.* **1989**, *11*, 39–63. [CrossRef]
8. Luan, C.; Wang, X.; Liu, Z. A core technology identification method based on the database of Derwent. *Sci. Sci. Manag. S. T.* **2008**, *6*, 32–34.
9. Sun, T.; Jin, B. Key Technology Mining and Enterprise Technology Competitive Intelligence--The Case of DVD Laser Head Technology. *Libr. Inf. Serv.* **2008**, *5*, 129–132.
10. Yang, W.; Yang, D. Research on the identification of industrial core technology based on patent data--taking 5G mobile communication industry as an example. *J. Intell.* **2019**, *38*, 39–45+52.
11. Zhao, X.; Wu, D.; Wu, W.; Sun, Z.; Hu, J.; Lian, Y.; Shan, J. A High-Quality Patent Identification Model for 'Necked' Technologies Based on Deep Learning and Multi-Classification Polling Mechanism--Taking Patent Application Documents as the Main Research Body. *Data Anal. Knowl. Discov.* **2023**, *7*, 30–45.
12. Wang, X.; Gao, M. BERT-LDA based key technology identification method and its empirical research—Taking agricultural robots as an example. *Libr. Inf. Serv.* **2021**, *65*, 114–125.
13. Zheng, G.; Li, W.; Liu, G.; Lian, G. A study on the selection of 'necked' key technologies based on Delphi method and hierarchical analysis—A case study of biomedical field. *World Sci.-Tech. R D* **2021**, *43*, 331–343.
14. Buckley, J. Fuzzy hierarchical analysis. *Fuzzy Sets Syst.* **1985**, *17*, 233–247. [CrossRef]
15. Sun, L.; Zhu, L.; Zhang, X.; Liu, C.; Wang, Y. Identification of key technology elements based on fuzzy hierarchical analysis. *Sci. Technol. Eng.* **2020**, *20*, 9816–9821.
16. Ma, D.; Hung, S. An Integrated Framework for the Selection and Acquisition of Core Technologies: The Case of Taiwan's LED Industry. *Long Range Plan.* **2015**, *48*, 381–397. [CrossRef]
17. Lv, K.; Chen, X.; Jing, J. Research on key technology identification of blockchain financial industry based on combinatorial disambiguation method and LDA model. *Libr. Inf. Serv.* **2022**, *69*, 110–121.
18. Lai, K.; Chen, H.; Chang, Y.; Kumar, V.; Bhatt, P. A structured MPA approach to explore technological core competence, knowledge flow, and technology development through social network patentometrics. *J. Knowl. Manag.* **2021**, *25*, 402–432. [CrossRef]
19. Yuan, Y.; Wang, S.; Tao, Y. Construction of the Measurement System of Key Core Technology Innovation Capability of Artificial Intelligence: Based on the Perspective of Innovation Ecosystem. *Sci. Technol. Prog. Policy* **2021**, *38*, 84–93.
20. Yu, J.; Chen, F.; Zhang, Y.; Liu, R. Casting a strong national weapon:the law exploration and system construction of key core technology breakthrough. *Bull. Chin. Acad. Sci.* **2019**, *34*, 339–343.

21. Zhang, Y.; Tan, L. Definition of the concept of key core technology, analysis of its characteristics and path of breakthrough. *Forum Sci. Technol. China* **2023**, *2*, 20–29.
22. Hu, X.; Yuan, C. Key core technologies: Concepts, characteristics and breakthrough factors. *Stud. Sci. Sci.* **2022**, *40*, 4–11.
23. Yang, S. Government Promotion of Key Core Technology Innovation: Theoretical Basis and Practical Solutions. *Economist* **2020**, *9*, 85–94.
24. Zhou, L.; Liu, C.; Huo, L.; Wang, Y. Research on center node identification algorithm of directed network based on crossover degree. *Comput. Appl. Res.* **2016**, *33*, 3299–3302+3306.
25. Wang, X.; Liu, Z.; Chen, Z.; Shan, X. Research on hot spots of public concern in the development of COVID-19 vaccine in China based on Weibo. *Inf. Explor.* **2022**, *2022*, 125–134.
26. Wang, K.; Chen, Y. Research on early identification of disruptive technologies based on heterogeneous patents. *Stud. Sci. Sci.* **2023**, *41*, 1364–1375.
27. Cao, K.; Wu, X.; Bai, G.; Zheng, Y.; Jin, B.; Li, L. Research on "stuck neck" technology recognition based on patent literature—Taking CNC machine tool field as an example. *Libr. Inf. Serv.* **2023**, *67*, 80–91.
28. Shi, Y.; Yang, L.; Jiang, Z.; Jia, Z.; Jiang, Z. Research progress on the application of graphene oxide in proton exchange membranes. *Membr. Sci. Technol.* **2015**, *35*, 114–121.
29. Fu, F.; Zang, J.; Chen, J.; Zhang, S.; Zhang, Y.; Fan, J. Application of graphene oxide in proton exchange membranes for fuel cells. *Chem. Ind. Eng. Prog.* **2019**, *38*, 2233–2241.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.