Article

Core Technology Topic Identification and Evolution Analysis Based on Patent Text Mining—A Case Study of Unmanned Ship

Yan Lin [1], Xuelei Wang [1], Jing Yang [1] and Shutian Wang [2],*

[1] School of Shipping Economics and Management, Dalian Maritime University, Dalian 116026, China; linyan@dlmu.edu.cn (Y.L.); wangxuelei_2021@dlmu.edu.cn (X.W.); 18035260962@163.com (J.Y.)
[2] School of Economics and Management, Dalian Ocean University, Dalian 116026, China
* Correspondence: 13258845726@163.com

Abstract: Accurate identification and evolutionary analysis of core technology topics within patent texts play a crucial role in enabling enterprises to discern the development trajectory of core technologies, optimize research and development (R&D) strategies, and foster technological innovation. Based on the perspective of time series dynamic analysis, this study uses the Latent Dirichlet Allocation (LDA) topic modeling and TF-IDF text vectorization methods to comprehensively mine and identify patent technology topics in the field of unmanned ships. This study deeply analyzes the dynamic evolution of unmanned ship technology topics from two aspects: the evolution of technology theme intensity and the evolution of technology theme content. We refine the development characteristics and future development directions of unmanned ship technology. The findings reveal two hot technologies, six growth technologies, and six declining technologies in unmanned ship technology. Furthermore, the analysis of technical topic evolution illustrates a pattern of fragmentation, inheritance, and integration. This study advances the methodologies used for identifying and analyzing patent technology topics and helps to grasp the development rules and evolutionary trends of core technologies. In addition, this paper has reference value for the research and practice of core technology topic identification and evolution analysis methods based on patent text mining.

Keywords: patent text; technology topic identification; technology topic evolution; unmanned ships

1. Introduction

The fast-paced evolution of the science and technology industry has made technological innovation a key driver of socio-economic development and enterprise competitiveness. However, navigating the complex R&D landscape, characterized by rapid technological growth and interdependencies, poses challenges for enterprises. Therefore, it is vital for them to swiftly and accurately identify core technology topics and track development trends to seize market opportunities and maintain a competitive edge. Patents play a pivotal role in this process [1,2]. Efficient analysis and interpretation of patent literature are crucial for identifying and analyzing core technology topics.

Researchers have proposed various methods to mine and identify technical themes in patents and unveil technological evolution processes, focusing on fine-grained analysis data. These methods include patent classification numbers, citation networks, and text content analysis [3,4]. However, relying solely on patent classification numbers may hinder accurate identification due to the vast array of technical topics. Similarly, analysis based on patent citation networks suffers from time lags and lacks continuity in tracking technical evolution. In contrast, analyzing patent text content offers a more objective approach, free from subjective biases. By delving into patent text, this method can effectively uncover potential technical themes and reveal their evolution patterns [5].

This study employs LDA topic modeling and TF-IDF text vectorization methods to comprehensively identify core technical topics in unmanned ships through a time-series dynamic analysis of patent text content. The study offers insights into technology

development characteristics and future directions by delving into the evolution intensity and content changes of these core technology topics. Through meticulous analysis of patent text content, we aim to uncover core technical themes and their evolution patterns, providing valuable insights into the development trends in this technical field and offering new directions for technological innovation. Furthermore, this study presents a novel perspective on theoretical research regarding core technology identification and evolution, addressing limitations observed in previous studies.

The main contributions of this paper are as follows. (i) Theoretical implications: Potential technological development opportunities are explored in technological innovation. Identifying the dynamic evolution of relevant technology development can provide more real-time insight and help capture changes in technology development in a timely manner, thus accelerating the innovation process. It reveals the evolutionary path of its technical themes and promotes innovation and development between technologies. (ii) Practical implications: It can excavate the core technology of unmanned ships, master the development path of core technology, and provide suggestions for enterprises to innovate unmanned ship technology. This paper applies theoretical research to real-life situations and identifies unmanned ship hotspot technologies that help shipping companies make decisions on unmanned ship technology development.

This paper is structured as follows. Section 2 reviews the theoretical background of the main methods and techniques, including the current state of development of unmanned ship technology, patent technology identification, and evolution analysis. Section 3 explains the research thinking and the process of technology theme identification and evolution. Section 4 describes the application process using the example of unmanned ships. Section 5 discusses unmanned ship technology development trends. Finally, Section 6 concludes the paper, reviewing its limitations and recommending subsequent research.

2. Literature Review

Identifying patented technologies and analyzing their evolution is crucial for understanding technological trends and fostering innovation. Table 1 provides an overview of the state of the art in topic modeling and recognition methods. The academic community has extensively researched these topics, focusing on three main analysis methods: patent classification numbers, patent citation networks, and patent text content. The research objectives of this paper are to review the development of unmanned ship technology and describe the current path of growth in the field. By constructing a framework for mining technologies, the readers can better understand the detailed technologies under each category. The proposed methodological framework can be used further to explore other technological aspects of unmanned ship technology.

Table 1. Overview of related works.

	Paper	Model Used	Real Dataset	Special Features	Limitations
Patent classification numbers	Zhang et al. [6]	LDA, text similarity calculation	blockchain field patents	Technological evolution	Inaccurate, Without objectivity
	Zhou et al. [7]	IPC co-category analysis	hydrogen fuel cell patents	map the knowledge map of hotspots	Inaccurate, Insufficient information
Patent citation networks	Liu et al. [8]	Main path analysis, Evolutionary trajectory	electric vehicles patents	co-opetition situation analysis	Inaccurate, Time lag
	Yang et al. [9]	structural holes, SAO network	Graphene patents	identifying technology development trends	maybe errors in explaining the relationship

Table 1. *Cont.*

	Paper	Model Used	Real Dataset	Special Features	Limitations
Patent Text content	Wang et al. [10]	LDA	Communication patents	Increased institutional topic probability hierarchy	Without dynamic characteristics
	Dotsika et al. [11]	keyword network analysis, visualization approach	3D Printing, Big Data, Bitcoin, Cloud Technologies,	Technical forecasts	Inaccurate
	Xue et al. [12]	technology evolution path	hydrogen energy patent	visualization	Uncovered potential information
	Lin et al. [13]	SAO, text similarity	mechanical structures patent	fast, accurate	Without dynamic characteristics
	Altuntas et al. [14]	k-means, text mine	Renewable energy patent	data-driven analysis	Unclear evolutionary path
	Di et al. [15]	topic modeling, clustering	scientific papers	Automated determination of parameters	take a long time, long text

The method based on patent classification numbers aims to grasp technical connections and evolutionary relationships across various fields by analyzing patent classification numbers. This helps uncover related technologies' development history and trends in different domains [16,17]. For example, Zhang et al. [6] used LDA to conduct technical evaluation and roadmap analysis in the blockchain field based on patent data and predicted future development trends. Zhou et al. [7] used IPC joint analysis to map the hotspot knowledge map in the hydrogen fuel cell technology field and then identified hotspot technologies. Liu et al. [8] further analyzed the development trend of technology in electric vehicle charging by analyzing high-frequency IPC classification numbers and their time distribution. The analysis method based on patent classification numbers can analyze the technological evolution process to a certain extent [18]. However, classification based on patent classification numbers may be inaccurate or subjective. In addition, the analysis method does not go deep into the patent text content. It cannot cover the specific details and characteristics of the technology, which may affect the accurate understanding of technological evolution.

The patent citation network analysis method constructs networks between technical fields using citation relationships among patent documents. This approach calculates network attributes to unveil potential technical themes [19,20]. For example, Small et al. [21] used the difference function to identify emerging technology topics based on direct citations and co-citations. Yang et al. [9] constructed an SAO network and identified graphene technology development trends by analyzing structural holes, node degree distribution, and network centrality changes. Zheng et al. [22] used point degree centrality and structural holes to identify and compare emerging technology fields across the Taiwan Strait based on patent citation network analysis. Ning et al. [23] constructed a mobile phone chip technology patent citation network based on topological structure parameters to identify core technologies. They used the SPNP algorithm to analyze the main path of technology diffusion. The patent citation network analysis method is an effective technology identification and analysis tool [24,25]. However, this method relies on citation relationships between patent documents, making it difficult to identify specific technical subjects and keywords and ignoring the diversity of technological evolution paths. Moreover, this analysis method pays close attention to highly cited patent documents and easily ignores low-cited or emerging technology patent documents.

The analysis method based on patent text content is a method that uses the content of patent documents to identify and classify technical topics. The distribution and evolution of different technical topics are revealed by analyzing keywords, technical terms, invention abstracts, and other information in patent documents [26]. For example, Wang et al. [10] used the LDA topic model to identify communication technology topics from patent titles and abstracts. They incorporated an institution–topic probability hierarchy to ascertain the distribution probabilities of competing enterprises within each topic and their technological standing. Dotsika et al. [11] used the LDA topic model and Word2vec vector model to identify core patent technology topics in artificial intelligence and analyzed the technical topic intensity and topic content evolution. Analyzing patent text content bypasses subjective judgment, enabling a comprehensive grasp of topic distribution and evolution in the technical realm. It facilitates timely detection and tracking of the latest technological advancements and trends [27]. However, semantic ambiguity in the patent text information analysis may affect the accuracy of the analysis results.

Evelina et al. [15] designed and developed ESCAPE (enhanced self-tuning characterization of document collections after parameter evaluation). It integrated two different solutions for document clustering and topic modeling: the joint approach and the probabilistic approach. Both approaches are able to correctly identify meaningful partitions of a given document corpus by grouping them according to topics. Escape minimizes user intervention and the method works well for large texts. For small text, the parameters need to be chosen in the context of the target domain. Patents belong to short text data and TF-IDF applies to short text of patents. Therefore, in this paper, we use this method to set weights to remove some high-frequency but not technically meaningful words.

Xue et al. [12] conducted research on identifying the technology evolution path in the field of hydrogen energy, adopted text mining methods to mine patent data, and introduced document vectorization and phrase mining algorithms to improve the mining depth while increasing the interpretability of the results. Wei et al. [28] constructed, through the unsupervised LDA topic clustering method and the main fields of activity, blank fields, and fields of agricultural machinery were visually analyzed. Lin et al. [13] used the SAO structure extraction technique for the patent text to obtain the effective content of the text, and acquisition of core technologies in the field of mechanical structures. Xin et al. [29] used SAO semantic mining methods to mine technical problems and solutions contained in scientific and technical papers and patents. Altuntas et al. [14] mined the abstracts of wind energy technology patent documents and used the k-means clustering algorithm to determine the distribution of the keywords among the clusters.

Unmanned ship technology presents unprecedented prospects, but existing research on the identification and evolution analysis of unmanned ship technology topics is limited. Ghaffari et al. [30] utilized the LDA topic model to extract technical themes and keywords from unmanned ship patent text. They constructed a co-occurrence network based on unmanned ship technology keywords to analyze the evolution process of technical topics and keywords, identifying technology opportunities. Wang et al. [31] used BERT to classify keywords in unmanned ship patent texts and used TEMPEST to cluster technical topics. Next, the function-oriented search (FOS) method is applied to extract fresh technical elements from patents across various fields. This process involves reorganizing both original and novel technology elements to pinpoint technical opportunities. Previous studies have identified and analyzed unmanned ship technology from various perspectives, guiding this study in understanding the development direction of unmanned ship technology. However, the above research is lacking in terms of specific technical theme systems, identification of evolution rules, and future development trends in the field of unmanned ship technology.

Deep learning models typically require large amounts of labeled data for training. It is more practical for large corpora or real-time processing tasks, which require large computational resources or long training times. However, LDA and TF-IDF are typically more computationally efficient and less resource-intensive than deep learning models. It achieves better performance on smaller datasets for lower data volumes and provides

relatively intuitive and interpretable results. The patent data of unmanned ships are small and the abstract text is short; this paper chooses LDA and TF-IDF for text topic mining.

In conclusion, the current patent technology theme identification and evolution analysis research has achieved specific results. However, the intricacy and variability of patent data pose a persistent challenge in conducting more precise and efficient identification and analysis of patent technology subjects and their evolution. In particular, the pattern of identification and evolution of patented technologies in unmanned ships has not yet been fully revealed. Therefore, this paper uses LDA topic modeling and TF-IDF text vectorization methods to mine and identify technical topics and keywords from patent texts; at the same time, it deeply analyzes the changes in the evolutionary intensity and content of the technology theme and further refines the characteristics of unmanned ship technology development and the future development direction.

3. Methods

3.1. Research Framework

This paper explores the application of the LDA topic model in identifying technical topics within unmanned ships. It elucidates their dynamic evolution process and further delineates the development characteristics and directions of pivotal technologies in unmanned ships. The research framework is shown in Figure 1. The main steps include:

- Patent data collection and preprocessing. First, we collect and download patent data in unmanned ship technology, then clean the collected data, delete irrelevant and duplicate data, and standardize the data.
- Technical topic identification. Mining and identifying technical topics and keywords through the LDA topic model and TF-IDF text vectorization method, and drawing word cloud diagrams.
- Evolution results of technical theme intensity. Draw a heat map of technical topics based on the intensity of technical topics, and analyze the attention and rise and fall of technical topics at each stage.
- Evolution results of technical theme content. This paper divides technology development stages according to the technology life cycle theory, uses cosine distance to measure the similarity between technical topics, draws the content evolution diagram of technical topics, and analyzes the evolutionary relationship between technical topics at each stage.
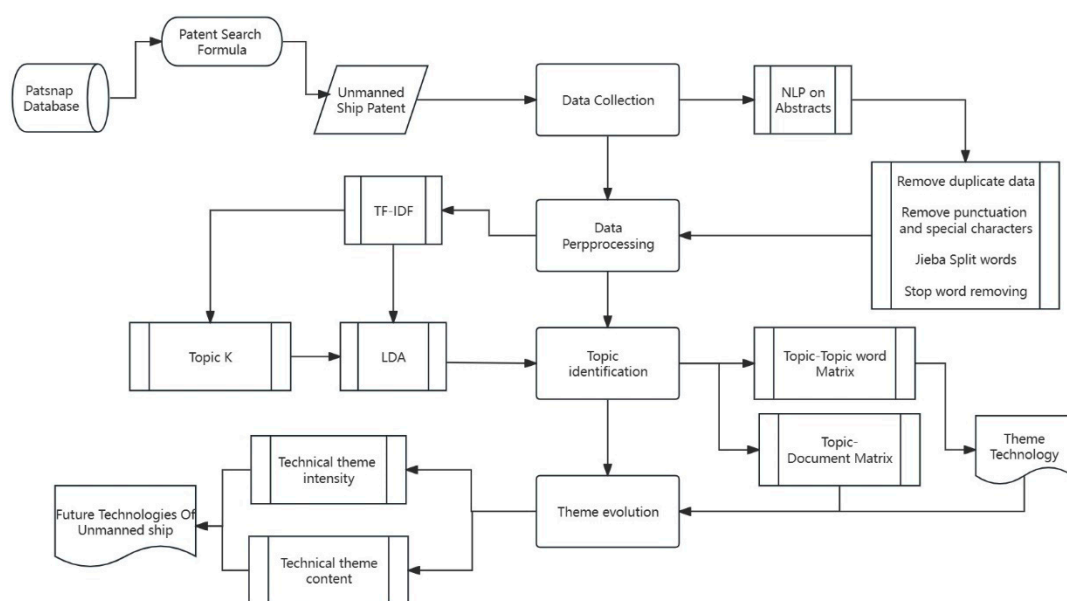

Figure 1. Research framework.

3.2. Data Acquisition

This article focuses on patent data related to unmanned ships, conducting empirical research utilizing the Patsnap global patent information database as the data source. The Patsnap global patent information database covers patent information around the world. It is updated regularly and promptly, including the latest patent information. It also provides a multi-dimensional search function to facilitate users in conducting flexible searches based on keywords, patent numbers, applicants, and other conditions and locate the required information more accurately. Therefore, the Patsnap global patent information database can be used as the data source for this study.

This study uses "unmanned ship" as the keyword to search for all patents in unmanned ships before 2023. It retrieved a total of 3902 invention applications, utility models, authorized inventions, and design patents in the field of unmanned ship technology. The annual number of patent applications related to unmanned ships is shown in Figure 2. The results show that the development of unmanned ship technology was relatively slow from 2013 and the number of patent applications was relatively small. Unmanned ship technology began to develop, and the number of patent applications gradually increased in 2013. The number of applications skyrocketed, and unmanned boat technology developed rapidly after 2016. This study selected the patent data of unmanned ships from 2013 to 2022 for analysis to ensure the robustness of the experimental results. After cleaning the patent data and removing duplicate and irrelevant data, 3586 valid patent data were finally obtained. The search expressions for unmanned ship patents are shown in Table 2.
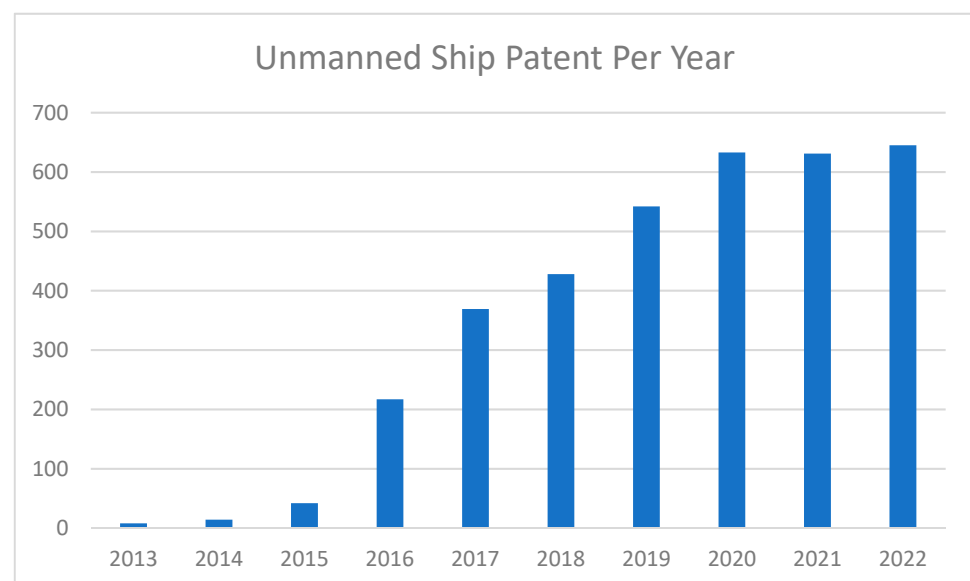

Figure 2. Number of patent applications for unmanned ships from 2013 to 2022.

Table 2. Unmanned ship patent search expression.

Type	Content
Database	Patsnap Database
Search scope	2013–2022
Patent search scope	Invention application patent, Utility model patent, Authorized invention patent, and Design patent
Search expression	Key words = (unmanned ship)

Python 3.9.13 programs are used to preprocess patent texts, which mainly include the following three steps to better mine patent texts. We used the sklearn library in Python for topic modeling, NumPy 1.22.4 for data preprocessing, and pyLDAvis 3.4.0 for visualization. pyLDAvis uses scatter plots or bubble plots to display topics. Each point or

bubble represents a topic. Their position and size then reflect the similarity and importance between topics. The pyLDAvis tool allows the user to view these keywords directly and the user can more easily identify the core concepts or subject areas of each topic. The pyLDAvis tool can also display the relationship between topics and documents. This is usually performed by presenting the documents (usually as columns or rows) and the strength of their association with each topic in a separate chart. We look at the topic distribution and document–topic relationship graphs through pyLDAvis, which allows the user to identify possible outliers or noise. For example, some documents may be only weakly related to all topics, which may be due to the fact that the content of these documents does not match the goal of topic modeling or contains too much noise.

First, before word splitting, Jieba will perform some preprocessing operations on the input text, such as removing punctuation marks, special characters, etc., in the text. Second, the Jieba splitter loads a built-in dictionary containing a large number of Chinese words. The loading of the dictionary determines which words can be recognized. Adding a deactivation word list removes irrelevant words and removes words with low contribution value, which greatly improves the accuracy and efficiency of text processing. Then, some words commonly used in patents have a high frequency of occurrence in each text. However, these words do not have any value for text analysis, so we need to be clear about these words in text pre-processing, for example, 'invention', 'innovation', 'patent', and other words.

3.3. Methods

3.3.1. LDA Topic Model

The LDA topic model is widely used in text topic mining. The LDA model offers advantages for semi-structured documents like patent documents, including fast mining speed, comprehensive coverage, and high accuracy [32,33]. Therefore, this paper applies the LDA topic model to mine and analyze the patent text to extract the technical topics and topic words in the patent text. This study uses the LDA topic model to analyze patent texts, which helps to reveal the implied subject terms within the text. It analyzes the evolution of technology and identifies trends in technology. The principle of the LDA topic model is shown in Figure 3. The solid circular nodes represent the observable variables and the directed edges between these nodes depict the conditional probabilistic dependencies between the variables. The rectangular boxes represent the process of repetition; the outer rectangular box reflects the process of repeatedly generating topic distributions for each document, and the inner rectangle reflects the process of repeatedly generating words from these topic distributions. LDA deconstructs the relationship between the input document and the original word vectors in two steps: selecting topics from the document and selecting words from the selected topics.
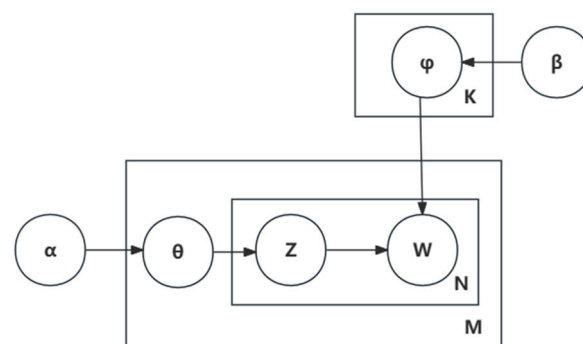

Figure 3. LDA theme model.

'M' represents the total number of documents, 'N' represents the number of words in a single document, 'K' represents the number of topics, 'W' represents all the words in the corpus, and 'Z' represents all possible topics. The parameter θ represents the distribution between documents and topics; φ represents the distribution of topics over words.

LDA extracts potential themes from unstructured text, which is useful for understanding and analyzing patent content. Patent texts are often highly diverse and complex, containing a great deal of specialized terminology and technical details. LDA models are able to adapt to different types of text data and deal with noise and redundant information in the text.

At the same time, the text vectorization method of term frequency–inverse document frequency (TF-IDF) is used to allocate weights to the words in the patent text. TF-IDF tends to filter out common words and retain important words, which can reduce the impact of invalid high-frequency words on technical topics. It can be calculated as follows.

$$TF - IDF(w_i, D_{text}) = \frac{F_{w_i}}{F} \times ln \frac{D_{text}}{(d_i : w_i \in d_i) + 1} \tag{1}$$

where D_{text} represents the entire text set. d_i is the text in set D_{text}. w_i represents the words. F represents the total number of times a word appears. F_{w_i} represents the number of times the word appears in D_{text} text. The higher the TF-IDF value of a word, the greater the importance of the word in the text, and the more likely it is a keyword of the text.

3.3.2. Technical Topic Extraction

This paper uses perplexity to determine the optimal number of topics. Perplexity is the evaluation function Perplexity, which is used to measure the model's generalization ability. The lower the perplexity value, the stronger the model's generalization ability. It can be calculated as follows.

$$Perplexity(D_{text}) = \exp \left\{ -\frac{\sum\limits_{d=1}^{D} \log p(w_d)}{\sum\limits_{d=1}^{D} N_d} \right\} \tag{2}$$

$$p(w_d) = \prod_{i=1}^{N_d} \sum_{k=1}^{K} p(z_k|d) \times p(w_i|z_k) \tag{3}$$

where D_{text} represents the entire text set. D represents the number of documents in the text set. w_d represents the sequence of words in the d document. The non-repetitive vocabulary of the d document is denoted by N_d. $p(w_d)$ is the probability of the word sequence w_d, calculated by multiplying the probabilities of all words in the d document. The i-th word in document d is noted as i. K represents the assumed number of topics. $p(z_k|d)$ represents the probability of matching topic K in a given document. And $p(w_i|z_k)$ represents the probability that the topic z_k contains word i. It can obtain the optimal number of topics in the text collection and the document collection under each topic through Formulas (2) and (3).

3.3.3. Technical Theme Intensity

Technical topic intensity refers to the degree of attention a certain technical topic receives in a certain period. The greater the intensity of a technical topic, the higher the level of attention that the technical topic receives [34]. It is possible to better observe the development trends of related technologies by analyzing the changes in the intensity of technical topics over time. It can be calculated as follows.

$$Q(Z_{t,k}) = (\sum_{d=1}^{D_t} \theta_{d,k})/D_t \tag{4}$$

where $Q(Z_{t,k})$ denotes the intensity of topic k in the current time slice t, $\theta_{d,k}$ presents the probability of the topic k in the d document, and D_t indicates the number of documents in time slice t.

3.3.4. Evolution of Technical Theme Content

The evolution of technical topic content reflects the change process of a specific technical topic in the time dimension. This method calculates the similarity between two technical topics at similar time stages and then analyzes the content evolution of technical topics [35]. Cosine distance, Kullback–Leibler (KL) difference distance, and Jenson–Shannon (JS) distance are the most commonly used methods to measure topic similarity [36]. The origins of technical topics vary across different periods, leading to differences in the content and quantity of subject words represented by these topics. KL and JS distance cannot accurately measure the similarity of two topic distributions. In this paper, we use text similarity to calculate topic similarity. Cosine similarity has the problem of dealing with sparse matrices that will have large pinch points but low similarity. This paper solves the problem of text sparsity by vectorizing the topics of each stage using word2vec. Therefore, this paper measures the similarity between technical topics by calculating the cosine value of the angle between the two technical topic vectors. A smaller angle implies a value closer to 1, indicating a higher degree of similarity between the two technical topics [37]. This article uses cosine distance to calculate the similarity between technical topics at similar time stages. It can be calculated as follows.

$$sim(D_i, D_j) = \frac{D_i * D_j}{\|D_i\| * \|D_j\|} \tag{5}$$

where $sim(D_i, D_j)$ represents the similarity between the two different topics of D_i and D_j. D_i represents technical topic vectors in adjacent time periods, respectively.

4. Results and Analysis

4.1. Technical Theme Identification

This study further identifies technical themes through data screening and preprocessing. The training of the LDA topic model requires first determining the number of technical topics. It determines the optimal number of technical topics by calculating the perplexity values of different technical topics. Figure 4 shows the perplexity curve under a different number of topics. It can be seen from the figure that when the number of topics is 1–15, the confusion curve shows a rapid downward trend. The number of topics is between 15 and 35, and the decline rate of the confusion curve slows down. The number of topics is between 35 and 50, and the confusion curve shows a smooth fluctuation trend. The lower the perplexity value, the stronger the model's generalization ability and the optimal number of topics. Therefore, based on the needs of this research and the perplexity value, the number of topics is optimal when it is 35.

According to the weight of the subject words contained in each technical topic, combined with the specific content of the patent abstract corresponding to each technical topic, the 35 identified technical topics were named. And we conducted repeated checks to ensure that the patents under each technical subject can represent that subject. The unmanned boat technology subject terms are shown in Table 3. The specific naming is as follows: Speed control (Topic #0), Wireless communications (Topic #1), Navigation decision-making (Topic #2), Path planning (Topic #3), Hull structure (Topic #4), Automatic parking (Topic #5), Intelligent ship network (Topic #6), Anti-collision technology (Topic #7), Fault detection (Topic #8), Automatic recovery (Topic #9), Positioning and navigation (Topic #10), Remote control system (Topic #11), Surface rescue (Topic #12), Maritime monitoring (Topic #13), Ship body (Topic #14), Module control (Topic #15), Collaborative control (Topic #16), Hydrological observation (Topic #17), Propulsion technology (Topic #18), Anti-overturning technology (Topic #19), Attitude control (Topi c#20), Task processing (Topic #21), Acquisition technology (Topic #22), Movable platform (Topic #23), Intelligent aquaculture (Topic #24), Image processing (Topic #25), Remote control (Topic #26), Surveying and mapping technology (Topic#27), Motion optimization (Topic #28), Antenna technology (Topic #29), Renewable

energy technology (Topic #30), Object detection (Topic #31), Garbage cleanup (Topic #32), Water quality/Water area detection (Topic #33), Regional orientation (Topic #34).
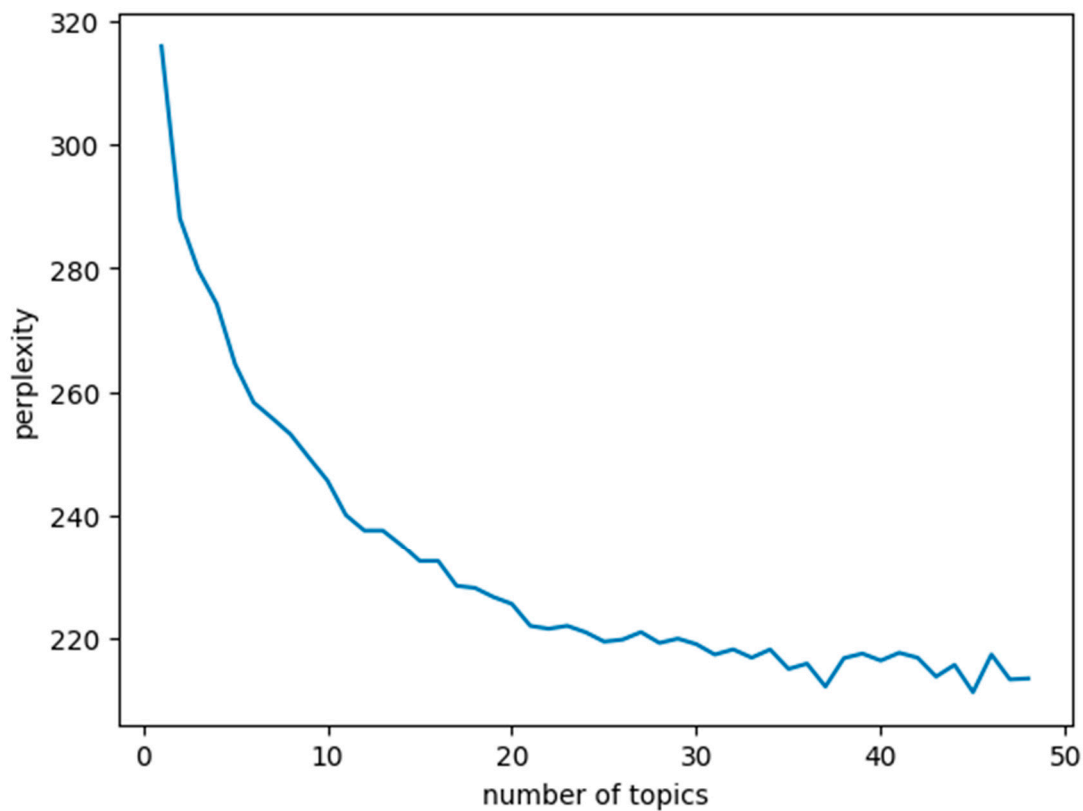

Figure 4. Perplexity curve chart under different number of topics.

Table 3. Identification results of technical themes.

Topic	Topic Word
Topic #0.	Controller, Design, Velocity, Time, Error, Disturbance, Observer, Performance, Guidance, Surface
Topic #1.	Wireless, Terminal, Robot, Base station, Ship, Cable, Technology, Communication, Transmission, User
Topic #2.	Navigation, Ship, Command, Action, Experimentation, Status, Risk, Assessment, Status, Capability
Topic #3.	Path, Planning, Obstacle, Steps, Route, Algorithm, Goal, Dynamic, Environment, Information
Topic #4.	Hull, Structure, Power, Field, Technology, Tail, Water Surface, Anti-collision, Camera, Effect
Topic #5.	Pump, Airbag, Pushrod, Bottom, Plate, Hydraulic, Fascia, Inflatable, Rod, Inlet, Pressure
Topic #6.	Data, Intelligence, Servers, Environment, Networks, Analytics, Data Processing, Remote, Integrated, Transmission
Topic #7.	Body, Spring, Active, Slider, Camera, Motor, Slot, Support, Rod, Slide, Slide Rail
Topic #8.	Information, Shore-based, Video, Failure, Status, Transmission, Node, Oil, Spill Control Command, Messages
Topic #9.	Boxes, Floats, Openings, Pipes, Grooves, Rings, Components, Rails Water, plants, Containers
Topic #10.	Sensors, Positioning, Heading, Distance, Angle, Adjustment, Attitude, Speed, Information, Trajectory
Topic #11.	Bracket, Circuit, Mode, Remote, Control, Power, Supply, Flexible, Work, Function, Voltage
Topic #12.	Signal, Rescue, Status, Dock, Inertial, Rotor, Personnel, Momentary, Sensing, Gain
Topic #13.	Module, Space, Ground, Power, Camera, Size, Object, Electric, Control, Body Management System
Topic #14.	Component, Hull, Adjustment, Component, Driver, Tubing, Structure, Fuselage, Impact, Damping
Topic #15.	Module, Controller, Positioning, Data, Communication, Attitude, Up, Drive, Transmission, Remote
Topic #16.	Formation, Ship, Collaboration, Mothership, Buoyage, Orbit, Network, Distributed, Shelf, Formation
Topic #17.	Measurement, Subsystem, River, Hydrology, Observation, Carrier, Data, Acoustic, Work Steps
Topic #18.	Thruster, Automatic, Structure, Cabin, Direction, Deck, Helm, Horizontal, Motion, Field
Topic #19.	Body, Limit, Battery, Support, Frame, Structure, Sheet, Elasticity, Function, Collection, Box, Technology
Topic #20.	Unit, Mechanical, Attitude, Storage, Centre, Winch, Base, Crossbar, Automatic, Monitoring
Topic #21.	Partial, task, processor, calibration, test, catheter, indicator, monitoring, point, computational, structure
Topic #22.	Powerplant, Mast, Battery, Bow, Sludge, Central, Navigation, System, Electrode, Somewhat, Corresponding
Topic #23.	Platform, Movement, Sonar, Emitter, Marker, Direction, Range, Adjustment, Motion, Terminal

Table 3. *Cont.*

Topic	Topic Word
Topic #24.	Drives, Motors, Propellers, Power, Farming, Feeding, Shafts, Transmission, Fields, Feeding
Topic #25.	Image, Ultrasound, Camera, Receiver, Image-processing, Pixel, Classification, Area, Object, Body
Topic #26.	Equipment, Operation, Monitoring, Remote, Water, Area, Technology, Field, Communication, Method, Work
Topic #27.	Surface, Case, Screw, thread, Slide, Motor, Map, Gear, Structure, Frame, Shaft
Topic #28.	Model Move Trajectory Parameter Problem Environment Predict Optimization Algorithm State
Topic #29.	Antenna Chassis Test Shell Cable Catamaran Connector Circuit Board Rotating Shaft Structure
Topic #30.	Solar, Energy, Battery, Power, Generation, Electricity, Electric, Battery, Panel, Generator, Wind, Wave, Utilisation
Topic #31.	Target, Water, surface, Detection, Radar, Features, Vision, Area, Cohesion, Data, Utilization
Topic #32.	Mechanisms, Garbage, Surface, Float, Regulating, Drive, Field, Technology, Floaters, Dynamics
Topic #33.	Detection, Monitor, Water, quality, Water, body, Waters, Analyze, Watery, Sensor, Automation, Pollution
Topic #34.	Area, Guidance, Water, Preset, Bathymetry, Terrain, Technology, Surface, Water, Sample, Operations

4.2. Technical Theme Intensity Evolution

The evolution of technical theme intensity can fully demonstrate changes and development trends over time. We calculate the intensity value of each technical topic in different years based on Formula (4) using the 35 identified technical topics, and draw a technical topic heat map, as depicted in Figure 5. The ordinate on the right is the intensity value; yellow indicates lower intensity, and red indicates higher intensity. The vertical axis on the left represents each technical theme. The horizontal axis represents the time window period.

It can be seen from the technical theme heat map that the intensity of most technical themes fluctuates, which shows that most technical topics in the field of unmanned ship technology are not fixed. Still, over time, the research interest in related technologies has been adjusted to adapt to the development trends of the times. As can be seen from Figure 5, Hull structure (Topic #4) and Remote control (Topic #26) are hot-spot technologies. It has received widespread attention in various periods, and the importance of technical topics has always remained high. Further, a line chart is drawn to show the evolution trends of growth and decline in technology topics over time. Figure 6 reports the evolution trend of growth technology topics. It can be seen that Path planning (Topic #3), Surveying and mapping technology (Topic #27), Motion optimization (Topic #28), Target detection (Topic #31), Garbage cleaning (Topic #32), Water quality/water area detection (Topic #33) are growing. The intensity of these six technical themes shows an upward trend. The technical theme intensity is at a relatively high level. This indicates that these six technical topics are hot technologies in the current field of unmanned ship research. Related technologies are developing rapidly, and many related patents have been generated. Figure 7 reports the evolution trend of declining technology topics. Wireless communication (Topic #1), Automatic recovery system (Topic #9), Remote control system (Topic #11), Attitude control system (Topic #20), Acquisition technology (Topic #22), Image processing (Topic #25) showed fluctuating trends in different periods. This shows that these six technical topics are relatively mature technologies currently being researched in unmanned ships, and the research interest is gradually decreasing.
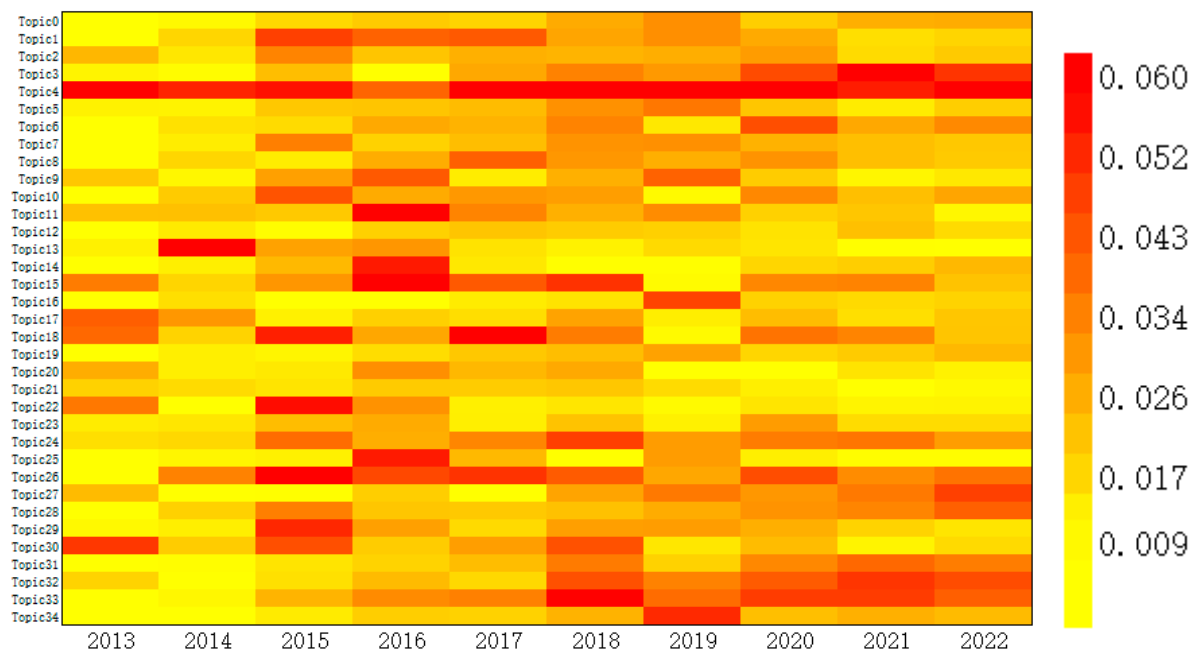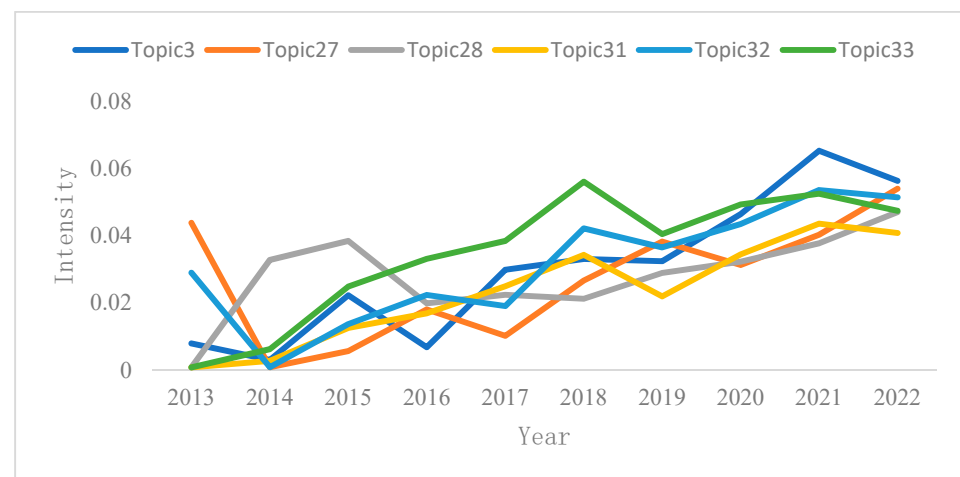
Figure 5. Technology topic heat map.

Figure 6. Evolutionary trends of growth technology topics.
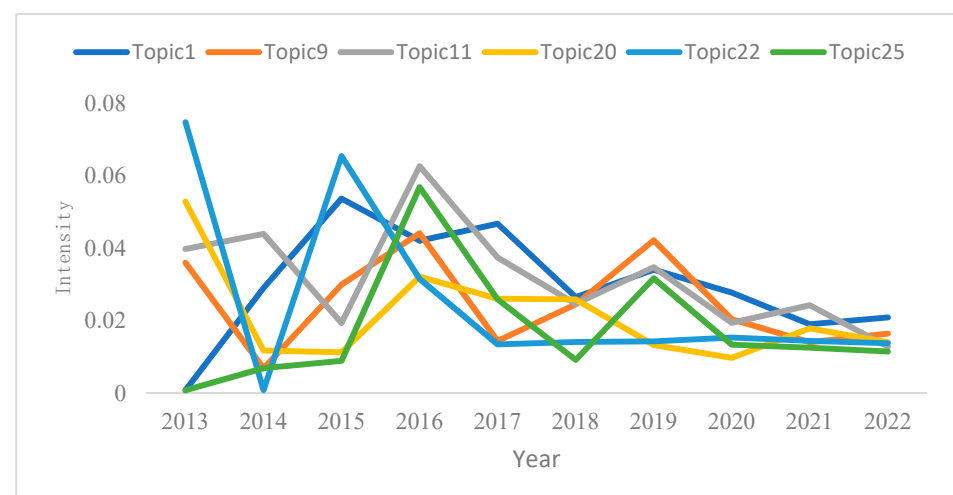

Figure 7. Evolution trend of declining technology topics.

4.3. Technical Theme Content Evolution

According to the theory of technology life cycle and combining the change in patent applications per year of unmanned ship technology from 2013 to 2022, Figure 8 shows the fitted curve of year and annual patent applications. The development of unmanned ship technology is categorized into the following three phases.
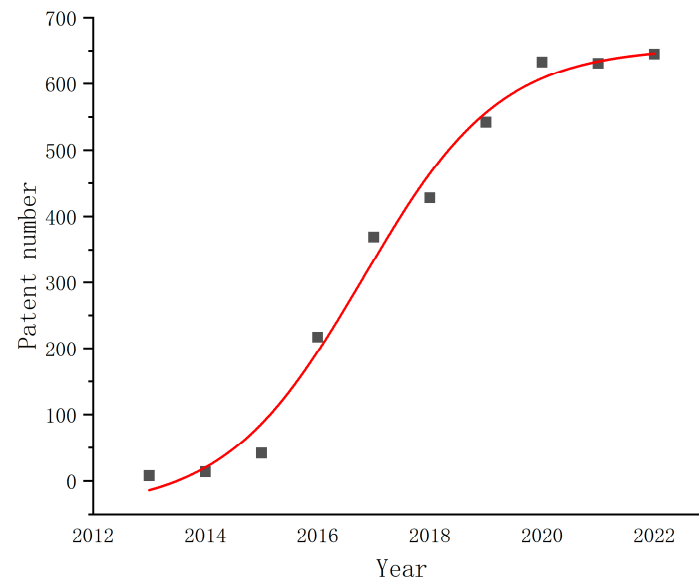

Figure 8. Fitting curve of the number of patents for unmanned ship technology.

- In the embryonic stage (2013–2015), the number of annual patent applications in unmanned ships began to increase, and the annual growth rate increased slowly, indicating that the technology in unmanned ships was gradually developing.
- In the rapid development stage (2016–2019), the number of annual patent applications in unmanned ship technology increased rapidly, and the annual growth rate was relatively high, indicating that the technology in unmanned ships was developing rapidly.
- In the stable development stage (2020–2022), the annual number of patent applications in the unmanned ship field technology was at a high level, and the annual growth rate was low, indicating that after the rapid development stage of the unmanned ship field technology, unmanned ship technology technology was still evolving.

Based on the technology life cycle theory and annual fluctuations in patent applications for unmanned ship technology, we categorize the development of unmanned ship technology into three stages: embryonic, rapid development, and stable development. Table 4 shows the subject content obtained by LDA training on patent data at each stage. In the embryonic stage of technological research in unmanned ships, there are fewer patent applications and fewer technical topics of concern in the first stage. In the rapid development stage of technological research in unmanned ships, the number of patent applications is gradually increasing, and the number of technical topics of concern rises in the second stage. In the steady development stage of technological research in unmanned ships, the number of patent applications and technical topics of concern has gradually stabilized in the third stage.

The cosine similarity between technical topics at each stage is calculated using Formula (3). To improve the accuracy of topic association, the similarity threshold T needs to be designed. If the threshold between two topics is less than T, the correlation between the two topics is considered insufficient. This study invalidates the topic association when the similarity is less than 0.3. Figure 8 shows that the correlation between the two technical topics is stronger. The more comprehensive the connection line between each technical topic, the stronger the correlation between the two technical topics.

Table 4. Thematic content of unmanned ship technology at different stages.

Stage 1 (2013–2015)	Stage 2 (2016–2019)	Stage 3 (2020–2022)
Theme Content	Theme Content	Theme Content
(A-Topic 0) Remote control	(B-Topic 0) Power Battery	(C-Topic 0) Sampling technology
(A-Topic 1) Sensor Technology	(B-Topic 1) Image Processing	(C-Topic 1) Solar Battery
(A-Topic 2) Water monitoring	(B-Topic 2) Garbage removal	(C-Topic 2) Garbage removal
(A-Topic 3) Autonomous navigation	(B-Topic 3) Water quality testing	(C-Topic 3) Path planning
(A-Topic 4) Measurement technology	(B-Topic 4) Hull structure	(C-Topic 4) Intelligent Aquaculture
(A-Topic 5) Hull structure	(B-Topic 5) Attitude control	(C-Topic 5) Speed control
(A-Topic 6) Drive System	(B-Topic 6) Target Detection	(C-Topic 6) Hull structure
	(B-Topic 7) Remote control	(C-Topic 7) Remote control
	(B-Topic 8) Solar battery	(C-Topic 8) Movable platform
	(B-Topic 9) Advancing technology	(C-Topic 9) Water quality testing
	(B-Topic 10) Path planning	(C-Topic 10) Propulsion equipment
	(B-Topic 11) Surveillance system	(C-Topic 11) Surveying and mapping technology
	(B-Topic 12) Autonomous navigation	(C-Topic 12) Positioning technology
	(B-Topic 13) Motor drive	(C-Topic 13) Power Battery
	(B-Topic 14) Surveying and mapping technology	(C-Topic 14) Simulation Technology

All technical themes of unmanned ships are related in three stages, showing an evolutionary relationship of division, inheritance, and integration in Figure 9. Each research direction gradually becomes stable and concrete, and the research depth gradually increases. The technical topics of remote control (A-Topic 0, B-Topic 7, C-Topic 7) and hull structure (A-Topic 5, B-Topic 4, C-Topic 6) have always been hot technologies in the field of unmanned ships. Autonomous navigation and attitude control technology topics only exist in the first stage (A-Topic 3, A-Topic 1) and the second stage (B-Topic 12, B-Topic 5). These two technologies are gradually maturing, and their research interest is declining. However, autonomous navigation technology mainly splits positioning technology (C-Topic 12), making the research more detailed. Some technologies are inherited and developed based on previous technologies, such as water quality/water area detection (A-Topic 2, B-Topic 3, C-Topic 9), and surveying and mapping technology (A-Topic 4, B-Topic 14, C-Topic 11). Some technologies have been inherited and developed in the second stage after integrating the technologies of the previous stage, such as battery management system (B-Topic 0, C-Topic 13), garbage cleaning (B-Topic 2, C-Topic 2), solar power generation (B-Topic 8, C-Topic 1), path planning (B-Topic 10, C-Topic 3). Some technologies appear in the third phase after integrating technologies from the previous phase, such as speed control (C-Topic 5) and simulation technology (C-Topic 14).

In order to verify the accuracy of the results of this paper, unmanned boat patent data were collected for the year 2023 to validate the results of this paper. The unmanned ship results derived above were validated by collecting unmanned ship patent data for the study. The feasibility of the technical opportunity to verify the results is obtained in the patent (CN117193321A) for an autonomous decision-making system for unmanned ship navigation and autonomous navigation method. The combination of path planning and autonomous navigation technologies may accelerate the development of intelligent navigation systems in the future. The feasibility of water quality monitoring in the future is verified in the patent (CN116991097A) for a water information intelligent monitoring system based on the unmanned ship technology, and in the patent (CN219565417U) for an intelligent river cleaning boat device, the driving force of the boat is used to automatically collect rubbish on the river surface so as to improve salvage effect. The patent (CN117614310A) for a composite nano power generation device for realizing self-driven power positioning of a ship was verified. The use of renewable energy sources such as solar energy may be a new development direction for unmanned ship energy technology in the future.
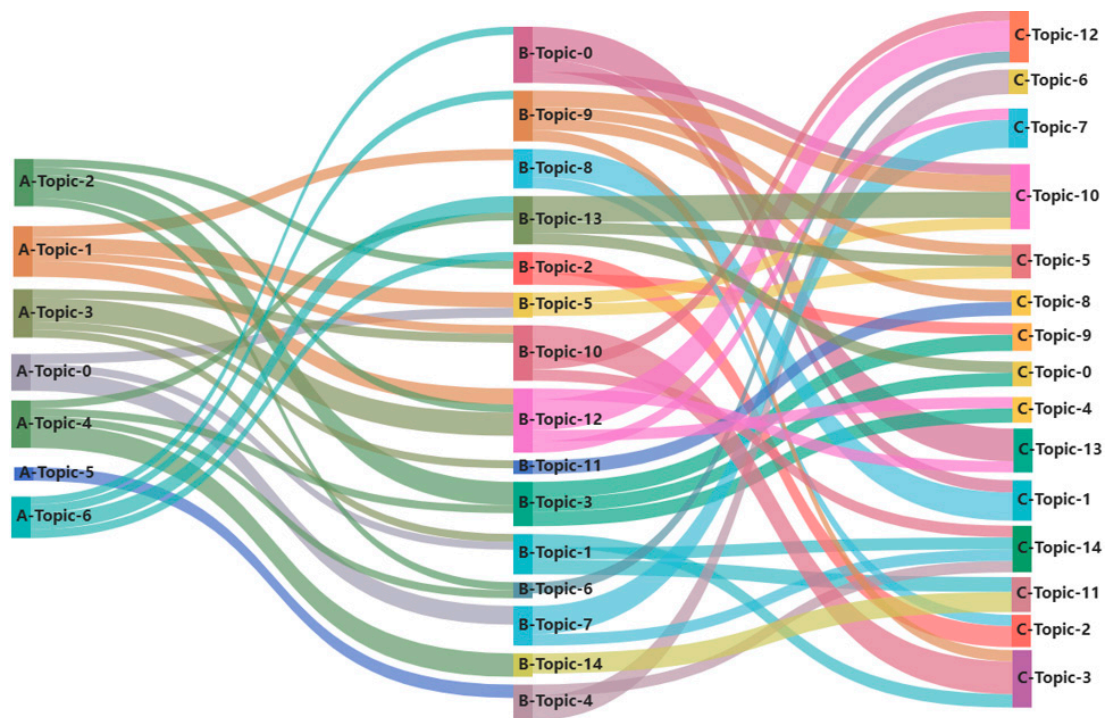
Figure 9. Technology roadmap.

5. Discussion

This study uses the LDA topic model and TF-IDF text vectorization method to mine and identify unmanned ship patent technology topics, and analyze the intensity evolution analysis and content evolution of technical topics in unmanned ship technology. Unmanned ship technology has a large development space in the fields of perception technology, navigation technology, control technology, and power technology. TF-IDF and LDA are two commonly used methods in text processing, and they each have different characteristics and applicable scenarios. In some cases, combining TF-IDF with LDA may bring better results, while in other cases, it may not be necessary or appropriate to use them together.

(A)   Against the use of TF-IDF with LDA:

When computational resources are limited, it may not be necessary to use TF-IDF and LDA at the same time. Both methods require a certain amount of computational resources, especially when dealing with large amounts of text data. If resources are tight, one of the methods can be chosen according to specific needs. For some simple text processing tasks, such as simple keyword search or basic text categorization, it may be sufficient to use only one of the TF-IDF or LDA methods. In such cases, using a combination of both methods may add unnecessary complexity.

(B)   In favor of using TF-IDF with LDA:

In some cases, combining TF-IDF with LDA may give better results. When we need to extract topics and identify keywords from the text at the same time, we can use TF-IDF in combination with LDA. TF-IDF can help identify keywords in the text, while LDA can extract potential topics from the text. In this way, we can obtain the topic distribution and keyword distribution of the text at the same time, so as to understand the text content more comprehensively. TF-IDF and LDA can also be used in combination with the task of text classification or clustering. TF-IDF can be used for feature extraction to convert text into numeric vectors, while LDA can be used to further extract the underlying thematic structure of the text on this basis. This helps to improve the accuracy of classification or clustering. In text summarization or recommender systems, the combined use of TF-IDF and LDA can help extract key information and recommend content on relevant topics. TF-IDF identifies

keywords and phrases in the text, while LDA makes content recommendations based on potential topics.

In this paper, we obtain 35 unmanned boat technology topics. Two hot technologies, six growing technologies, and six declining technologies are obtained through theme evolution. By analyzing the changes in technologies between time periods, this paper suggests future technology development for unmanned ships.

In the field of perception technology, there is a need to strengthen the development and optimization of perception technology for unmanned ships and focus on exploring potential technologies related to water quality/water area detection, surveying and mapping technology, target detection, and garbage cleanup. In the research and development and application of unmanned ship sensing technology, it is necessary to actively research and develop various sensors and sensing algorithms to accurately obtain environmental information around unmanned ship technology, including water quality, waters, terrain, water surface objects, etc. This sensing information provides a decision-making basis for the autonomous navigation of unmanned vessels and data support for monitoring and managing the water environment. At the same time, it is necessary to optimize the target detection algorithm further in computer vision. These algorithms can detect objects, obstacles, fairways, etc., that surround the unmanned ship in real time, helping the unmanned ship to navigate autonomously and avoid obstacles. The current focus of research is to use sensing technology to use sensors, cameras, and other equipment to identify, locate, and clean up garbage and other obstacles on the water surface.

In the field of navigation technology, it is necessary to design an autonomous navigation system with high accuracy, stability, and reliability, focusing on exploring potential technologies related to path planning and positioning. In the process of R&D and application of unmanned ship navigation technology, autonomous navigation systems are carefully designed and optimized according to different mission requirements and environmental conditions. In particular, it is necessary to conduct in-depth research on positioning technology to improve the precise positioning and navigation capabilities of unmanned ships. In addition, current research focuses on optimizing the path planning algorithm for unmanned vessels, strengthening the planning and decision-making capabilities of unmanned ship navigation paths, and avoiding obstacles and dangerous areas to achieve efficient and reliable navigation in various complex and dynamic environments.

In the field of control technology, it is necessary to strengthen the research and development of unmanned ship control technology, focusing on exploring potential technologies related to remote control, speed control, motion optimization, and movable platforms. In the process of R&D and unmanned ship control technology application, it is necessary to strengthen real-time monitoring of the ship's dynamics and position and promptly obtain parameters and information. At the same time, we actively develop specific control algorithms and actuators to adjust and control the speed of unmanned ships accurately. We need to strengthen the optimization and adjustment of the motion parameters of unmanned ships to improve the navigation performance and mission execution capabilities of unmanned ships in different waters and environments. In addition, it is also necessary to improve the stability and maneuverability of the movable platform and actively carry out the design and development of equipment such as hulls, propellers, and navigation sensors to achieve stable navigation and precise control of unmanned ships in the water.

In the field of power technology, it is necessary to focus on exploring potential technologies related to propulsion devices, power batteries, and solar batteries. In the field of power technology, it is essential to focus on exploring potential technologies related to propulsion devices, power batteries, and solar batteries. The propulsion device is the equipment that provides power for the unmanned ship. Different propulsion device types must be accurately selected, such as motor propulsion, water jet propulsion, propeller propulsion, etc. Power batteries are the energy source for unmanned ships. It is also necessary to develop efficient energy systems such as wave energy, lithium batteries, solar charging systems, solar batteries, or fuel cells to extend the endurance of unmanned ships.

6. Conclusions

This study starts from the perspective of time series dynamic analysis and uses the Patsnap global patent information database as the data source. We use the LDA topic model and TF-IDF text vectorization method to comprehensively mine and identify technical topics and keywords in unmanned ships, which provide reference and guidance for technology research and development in the field of unmanned ships. Based on the above research, the findings of this article are as follows. First, the evolution of related technologies in unmanned ships progresses through a budding stage, a phase of rapid development, and is presently in a stable development stage. Second, technological development in unmanned ships encompasses three themes: growth, decline, and hot spots, analyzed based on the intensity of technological themes. Then, the technical themes of unmanned ships are linked within the three stages of development, presenting an evolutionary relationship of splitting, inheritance, and fusion through the content analysis of technical themes. Finally, it was found that unmanned ship technology has more room for development in perception, navigation, control, and power technology by analyzing the evolution of technology topics.

This study also has some shortcomings. First, data sources are limited. This study only considered patent data in unmanned ships; the data source is relatively singular. The data source can be further expanded in the future, such as science and technology news, academic papers, and science and technology reports. Second, the research scenario is limited. This study only takes the field of unmanned ships as the research object. In future research, verifying the universality of the method in other fields is necessary.

Author Contributions: Conceptualization, Y.L. and X.W.; methodology, J.Y.; software, S.W.; validation, Y.L., X.W. and J.Y.; formal analysis, Y.L.; investigation, J.Y.; resources, J.Y.; data curation, X.W.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L.; visualization, S.W.; supervision, X.W.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

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

References

1. Ampah, J.D.; Jin, C.; Fattah, I.M.R.; Appiah-Otoo, I.; Afrane, S.; Geng, Z.; Liu, H. Investigating the evolutionary trends and key enablers of hydrogen production technologies: A patent-life cycle and econometric analysis. *Int. J. Hydrogen Energy* **2023**, *48*, 37674–37707. [CrossRef]
2. Miao, Z.; Du, J.; Dong, F.; Liu, Y.; Wang, X. Identifying technology evolution pathways using topic variation detection based on patent data: A case study of 3D printing. *Futures* **2020**, *118*, 102530. [CrossRef]
3. Liu, Z.; Feng, J.; Uden, L. From technology opportunities to ideas generation via cross-cutting patent analysis: Application of generative topographic mapping and link prediction. *Technol. Forecast. Soc. Chang.* **2023**, *192*, 122565. [CrossRef]
4. Baumann, M.; Domnik, T.; Haase, M.; Wulf, C.; Emmerich, P.; Rösch, C.; Weil, M. Comparative patent analysis for the identification of global research trends for the case of battery storage, hydrogen and bioenergy. *Technol. Forecast. Soc. Chang.* **2021**, *165*, 120505. [CrossRef]
5. Lee, P.C.; Su, H.N.; Wu, F.S. Quantitative mapping of patented technology—The case of electrical conducting polymer nanocomposite. *Technol. Forecast. Soc. Chang.* **2010**, *77*, 466–478. [CrossRef] [PubMed]
6. Zhang, H.; Daim, T.; Zhang, Y.P. Integrating patent analysis into technology roadmapping: A latent dirichlet allocation based technology assessment and roadmapping in the field of Blockchain. *Technol. Forecast. Soc. Chang.* **2021**, *167*, 120729. [CrossRef]
7. Zhou, H.; Dai, J.; Chen, X.; Hu, B.; Wei, H.; Cai, H.H. Understanding innovation of new energy industry: Observing development trend and evolution of hydrogen fuel cell based on patent mining. *Int. J. Hydrogen Energy* **2024**, *52*, 548–560. [CrossRef]
8. Liu, Z.; Xiang, X.; Feng, J. Tracing evolutionary trajectory of charging technologies in electric vehicles: Patent citation network analysis. *Environ. Dev. Sustain.* **2023**, *26*, 12789–12813. [CrossRef]
9. Yang, C.; Huang, C.; Su, J. An improved SAO network-based method for technology trend analysis: A case study of graphene. *J. Inf.* **2018**, *12*, 271–286. [CrossRef]
10. Wang, J.; Cheng, L.; Feng, L.; Lin, K.Y.; Zhang, L.; Zhao, W. Tracking and predicting technological knowledge interactions between artificial intelligence and wind power: Multimethod patent analysis. *Adv. Eng. Inf.* **2023**, *58*, 102177. [CrossRef]

11. Dotsika, F.; Watkins, A. Identifying potentially disruptive trends by means of keyword network analysis. *Technol. Forecast. Soc. Chang.* **2017**, *119*, 114–127. [CrossRef]

12. Xue, D.; Shao, Z. Patent text mining based hydrogen energy technology evolution path identification. *Int. J. Hydrogen Energy* **2024**, *49*, 699–710. [CrossRef]

13. Lin, W.; Yu, W. Measuring Patent Similarity Based on Text Mining and Image Recognition. *Systems* **2023**, *11*, 294. [CrossRef]

14. Altuntas, F.; Gok, M.S. A data-driven analysis of renewable energy management: A case study of wind energy technology. *Clust. Comput.* **2023**, *26*, 4133–4152. [CrossRef]

15. Di Corso, E.; Proto, S.; Vacchetti, B.; Bethaz, P.; Cerquitelli, T. Simplifying text mining activities: Scalable and self-tuning methodology for topic detection and characterization. *Appl. Sci.* **2022**, *12*, 5125. [CrossRef]

16. Khachatryan, D.; Muehlmann, B. Measuring the drafting alignment of patent documents using text mining. *PLoS ONE* **2020**, *15*, e0234618. [CrossRef]

17. Garechana, G.; Rio-Belver, R.; Bildosola, I.; Cilleruelo-Carrasco, E. A method for the detection and characterization of technology fronts: Analysis of the dynamics of technological change in 3D printing technology. *PLoS ONE* **2019**, *14*, e0210441. [CrossRef]

18. Wu, H.; Yi, H.; Li, C. An integrated approach for detecting and quantifying the topic evolutions of patent technology: A case study on graphene field. *Scientometrics* **2021**, *126*, 6301–6321. [CrossRef]

19. Ashouri, S.; Mention, A.L.; Smyrnios, K.X. Anticipation and analysis of industry convergence using patent-level indicators. *Scientometrics* **2021**, *126*, 5727–5758. [CrossRef]

20. Hu, R.; Ma, W.; Lin, W.; Chen, X.; Zhong, Z.; Zeng, C. Technology topic identification and trend prediction of new energy vehicle using LDA modeling. *Complexity* **2022**, *2022*, 1–20. [CrossRef]

21. Small, H.; Boyack, K.W.; Klavans, R. Identifying emerging topics in science and technology. *Res. Policy* **2014**, *43*, 1450–1467. [CrossRef]

22. Zheng, X.; Aborisade, M.A.; Liu, S.; Lu, S.; Oba, B.T.; Xu, X.; Ding, H. The history and prediction of composting technology: A patent mining. *J. Clean. Prod.* **2020**, *276*, 124232. [CrossRef]

23. Ning, L.; Guo, R. Technological diversification to green domains: Technological relatedness, invention impact and knowledge integration capabilities. *Res. Policy* **2022**, *51*, 104406. [CrossRef]

24. Naumanen, M.; Uusitalo, T.; Huttunen-Saarivirta, E.; van der Have, R. Development strategies for heavy duty electric battery vehicles: Comparison between China, EU, Japan and USA. *Resour. Conserv. Recycl.* **2019**, *151*, 104413. [CrossRef]

25. Govindarajan, U.H.; Trappey, A.J.C.; Trappey, C.V. Intelligent collaborative patent mining using excessive topic generation. *Adv. Eng. Inf.* **2019**, *42*, 100955. [CrossRef]

26. Wang, B.; Liu, S.; Ding, K.; Liu, Z.; Xu, J. Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology. *Scientometrics* **2014**, *101*, 685–704. [CrossRef]

27. Han, X.; Zhu, D.; Lei, M.; Daim, T. R&D trend analysis based on patent mining: An integrated use of patent applications and invalidation data. *Technol. Forecast. Soc. Chang.* **2021**, *167*, 120691.

28. Wei, T.; Jiang, T.; Feng, D.; Xiong, J. Exploring the Evolution of Core Technologies in Agricultural Machinery: A Patent-Based Semantic Mining Analysis. *Electronics* **2023**, *12*, 4277. [CrossRef]

29. Li, X.; Wu, Y.; Cheng, H.; Xie, Q.; Daim, T. Identifying technology opportunity using SAO semantic mining and outlier detection method: A case of triboelectric nanogenerator technology. *Technol. Forecast. Soc. Chang.* **2023**, *189*, 122353. [CrossRef]

30. Ghaffari, M.; Aliahmadi, A.; Khalkhali, A.; Zakery, A.; Daim, T.U.; Yalcin, H. Topic-based technology mapping using patent data analysis: A case study of vehicle tires. *Technol. Forecast. Soc. Chang.* **2023**, *193*, 122576. [CrossRef]

31. Wang, J.; Zhang, Z.; Feng, L.; Lin, K.Y.; Liu, P. Development of technology opportunity analysis based on technology landscape by extending technology elements with BERT and TRIZ. *Technol. Forecast. Soc. Chang.* **2023**, *191*, 122481. [CrossRef]

32. Wei, C.; Chaoran, L.; Chuanyun, L.; Lingkai, K.; Zaoli, Y. Tracing the evolution of 3-D printing technology in China using LDA-based patent abstract mining. *IEEE Trans. Eng. Manag.* **2020**, *69*, 1135–1145. [CrossRef]

33. Ma, Y.; Kong, L.; Lin, C.; Yang, X. Research on the identification of generic technology of eco-friendly materials based on text mining. *Environ. Sci. Pollut. Res.* **2022**, *29*, 35269–35283. [CrossRef] [PubMed]

34. Pan, X.; Zhong, B.; Wang, X.; Xiang, R. Text mining-based patent analysis of BIM application in construction. *J. Civ. Eng. Manag.* **2021**, *27*, 303–315. [CrossRef]

35. Yun, J.; Geum, Y. Automated classification of patents: A topic modeling approach. *Comput. Ind. Eng.* **2020**, *147*, 106636. [CrossRef]

36. Wang, J.; Hsu, C.C. A topic-based patent analytics approach for exploring technological trends in smart manufacturing. *J. Manuf. Technol. Manag.* **2021**, *32*, 110–135. [CrossRef]

37. Hu, Z.; Fang, S.; Liang, T. Empirical study of constructing a knowledge organization system of patent documents using topic modeling. *Scientometrics* **2014**, *100*, 787–799. [CrossRef]