Research Article

Digitalization and Information Management Mechanism of Sports Events Based on Multisensor Node Cooperative Perception Model

Yi Liu,[1] **Yaodong Wang** ,[1] **Yuntong Tan,**[2] **Jie Ma,**[1] **Yan Zhuang,**[1] **and Xiangqian Zhao**[1]

[1]*Department of Physical Education Teaching, China University of Mining and Technology (Beijing), Beijing 100083, China*
[2]*Department of Physical Education Teaching, Guang Dong University of Technology, Guang Zhou 510006, China*

Correspondence should be addressed to Yaodong Wang; wyd@cumtb.edu.cn

In the process of developing major sports events, how to guide providers and users to provide and utilize the archives information resources of major sports events and realize the interaction between them is an important problem to be solved urgently in the development of major sports events and the archive service of major sports events. By analyzing the present situation of archive service of major sports events, especially the analysis of the opposite dependent subjects of service providers and users, we can see that the continuous development of archive services for major sports events will inevitably lead to constant changes in user groups and user needs, guided by the theory of information retrieval, knowledge management, and media effect. According to the service model of archive service of major sports events, the archive service model of specific sports events is constructed. In this paper, four kinds of event recommendation models are applied to the collected marathon event data for experiments. Through experimental comparison, the effectiveness of content-based recommendation algorithm technology in the event network data set is verified, and an algorithm model suitable for marathon event recommendation is obtained. Experiments show that the comprehensive event recommendation model based on term frequency–inverse document frequency (TF-IDF) text weight and Race2vec entry sequence has the best recommendation performance on marathon event data set. According to the recommendation target of the event and the characteristics of the event data type, we can choose a single or comprehensive recommendation algorithm to build a model to realize the event recommendation.

1. Introduction

Under the background of big data and artificial intelligence, sports and big data are in urgent need of integration and development, and the increment of sports-related information, especially sports information resources on the Internet platform, has risen sharply. In the new era, our people's demand for sports is characterized by pursuing more personalized, multilevel sports events, and sports services to meet their own needs with the improvement of living standards [1]. At the same time, facing a large number of demands, the number of sports events and services also shows a rapid growth trend. Taking marathon events as an example, in 2018, a total of 1,102 events were held nationwide, with nearly 50 million participants. Among them,

there are 350 certified events of the Chinese Association of Athletics, and various interesting and characteristic theme events are also booming [2]. Therefore, a large number of race supplies provide rich choices for many marathon entry-level runners or ordinary participants, but complex race classification standards and uneven race promotion information also increase the difficulty for runners to choose races and the promotion cost of race service organizations. Beginner runners need to spend a lot of time screening race information and evaluating content, and inspecting and selecting suitable race services [3]. Today, with the rapid development of big data, the above-mentioned problem of noncirculation of sports event information resources, taking marathon event information as an example, is no longer a case, and the problem of event information resources should

be regarded as an information overload problem in essence [4].

In the digital age, the change of media urges people to get the needed sports information more quickly through the network and computer technology. However, sports information, especially sports information resources on a large number of Internet platforms, such as event data display information and text evaluation information, have not been collected and applied reasonably [5]. Fragmented network sports information, especially the text state information, makes the irrationality of traditional sports information classification and storage begin to appear, which easily leads to the dislocation of information resources intercommunication when users search for sports information, and excessive dislocation information accumulation easily leads to information overload. Facing the problem of information overload, search engine, e-commerce, and other fields have carried out in-depth research and discussion, among which the most important solution is entity recommendation technology [6]. In 1990s, some American scholars put forward the concept of recommendation system, and realized the content recommendation to forum users through recommendation technology. At present, the application of recommendation technology in sports field is not very extensive, and it is still only recommended for specific entities such as sports goods or sports news [7]. At the same time, under the current industry development background of artificial intelligence and "Internet plus," the mining and processing of sports event information resources need the introduction of new technologies. Therefore, if the fragmented sports information resources can be fully mined and applied by effectively constructing the event recommendation model and introducing a variety of algorithms to process the event network information data, it will help to improve the use efficiency of users' Internet fragmented event information taking marathon event network data as an example [8]. This fundamentally meets the needs of runners for suitable event information supply and enhances the interactive experience of users in using sports information resources [9]. At the same time, the introduction of event recommendation algorithm can promote the utilization rate and relevance of sports information resources, which will not only provide more personalized and professional information service technical support for sports enthusiasts but also help to improve the overall information intelligence level of sports and promote the development of sports industry [10].

At present, as one of the most important information exchange channels, there is still a lot of fragmented sports-related information. Under the background of the rapid increase of sports events and service information, studying the recommendation technology of competition information and the comparative application of various algorithms will help to make statistical analysis of competition network information more efficiently, realize the demand of people for competition information recommendation, and provide basic theoretical and technical support for the research of network sports information resources mining and application in a more intuitive and effective way. Internet data of sports events is a part of sports information resources, which

is fragmented information data. This paper analyzes the characteristics of the current network event information to select the appropriate data recommendation algorithm, through the combination of algorithm and data to build a practical event recommendation model, in order to provide the basis and reference for the application research of fragmented sports information resources represented by competition Internet data and the use of related methods. Second, it gives the possibility of integration and development of sports event information and related information technology from the perspective of technology, which provides a broader idea for sports informationization research and enriches the technical means of sports research. For this reason, archive service providers must constantly develop service infrastructure, change service concepts, and innovate service methods: build an objective foundation of archive service that adapts to the development of the times and can continuously integrate new technologies, new equipment, and new concepts.

2. Related Work

At present, the definition of Internet data of sports events is rarely mentioned in academic circles, and the Internet data of sports events is essentially an expression form of information resources based on Internet platform in sports information resources [11]. Therefore, in the research status analysis, choose to belong to the upper level of network sports information resources for current research and analysis, searching in the full-text database of academic journals of China Knowledge Network (CNKI) with the retrieval format of "Subject = Network Sports Information Resources." Among them, there are 179 literatures in journals and master's and doctoral dissertations, among which 13 are cited more than 20, but only 3 are highly related to the research of online sports information resources, all of which are before 2005, and their reference value is not great under the current Internet development background [12]. Through all the relevant literature available for inquiry, the main research directions are divided into two categories: the construction and application of network sports information resources in colleges and universities, and the integration and development of network sports information resources [13]. Among them, the dominant research content is the integration of network sports information resources and how to use them efficiently [14].

Among them, the researchers made a clear exposition on the mining and acquisition of network sports information resources at that time, covering the use of search engines, sports authoritative websites, network databases, sports websites, or sports channels of comprehensive websites. In the research direction of improving the retrieval efficiency of network sports information resources, based on the previous information resource acquisition skills, researchers put forward a method of using professional database retrieval skills and file type retrieval on the network platform [15]. According to the problems existing in the development of network sports information resources, the researchers put forward some suggestions on building a sharing platform of network

sports information resources and analyzed the corresponding operation mechanism and feasibility [16]. However, from all the available literature, most of the research on network sports information resources is to summarize the current situation and analyze the possibility of its integration, there is a certain technical lag, and there is little in-depth research on its internal data structure and application.

To sum up, at present, there is little research on the application of Internet data of sports events in China. This study builds a reasonable event recommendation system by analyzing the characteristics of public event information of professional sports vertical websites in the Internet, which is an effective and innovative research on network sports information resources based on this [17]. At present, with the government departments at all levels vigorously promoting the development of sports and cultural undertakings, more and more major sports events are held in major cities in China. Academic circles and other people from all walks of life, government departments, and other workers on major sports events have gradually been put on the agenda and become a hot spot of current research [18]. Among them, many scholars, staff, and government departments have noticed the various influences of archives on the inheritance, holding, development, and dissemination of major sports events and have written books, expressed their views, and clarified their positions, showing a scene of a hundred flowers blooming.

In this paper, before studying the major sports event archive service, the relevant literature was consulted, and some existing journals, papers, and government documents related to the major sports event archive service were collected and combed to obtain the necessary literature support [19]. By searching the full-text database of electronic resources of China Periodical Network, Wanfang Database, domestic and foreign government, and research websites, we find the literature related to the archives of major sports events and study some related works. From the current point of view, domestic scholars are committed to study the archives of major sports events from different angles [20]. After sorting out and analyzing this paper, some conclusions are drawn. From the existing achievements, the research on the archives of major sports events by relevant workers and scholars mainly includes three aspects: the research on the concept and management of sports archives, the need for archive service support for the development of sports cultural undertakings, and the development and utilization of sports archives information resources [21]. This paper will summarize the current research status from these three aspects.

At present, the research on the application of sports information resources and sports information mainly focuses on sports information and sports literature, while the application of a large number of fragmented sports event information based on the Internet is rarely mentioned.

3. Multisensor Node Perception of Internet Data of Sports Events

3.1. Basic Characteristics of Internet Data of Sports Events.
In order to effectively meet the urgent needs of the masses to participate in sports events, it is necessary to display and recommend many event information to the masses reasonably and accurately. Under the current background of "Internet plus," the event data in the Internet is multiplying day by day, and the characteristics of event information hidden behind a large amount of data and the rules of user browsing are effective information when constructing the event recommendation model. Therefore, this chapter will take the Internet data of sports events as the research object, deeply discuss the network data characteristics and reasonable data collection framework of marathon events, and collect corresponding data sets according to the framework, so as to provide basic data reference for the construction of sports event recommendation model, as shown in Figure 1.

The Internet information data storage of sports events is large, and the Internet content corresponding to a single event includes official websites, portals, professional forums, new media, and other information dissemination platforms. For example, by the end of 2019, searching for "Marathon Events" on Baidu search engine can obtain more than 30 million related web pages, covering information such as publicity, communication, and popular science of marathon events. At the same time, the content between sites is relatively independent. To obtain the corresponding event information completely, visitors need to obtain information through multiple related keywords and multiple platforms. This fragmentation feature is becoming more and more obvious in the Internet data of events with a sharp increase in data volume. Because of relying on the Internet platform for information display, the Internet data of sports events shows diversified characteristics in data type distribution, including video data, numerical data, computer language data, and other forms besides traditional text data and picture data. Taking the retrieval of "NBA Games" in Baidu as an example, the data types displayed include the text data of the victory and defeat reports of the Games, the numerical data of the players' participation information, and the video data of the wonderful performance of the Games. At the same time, the angle of each site to spread events and the scope of data collection are different, the types of sports event data transmission are different, and there is a lack of label definition for content types.

3.2. Internet Data Acquisition Based on Multisensor.
The collection framework of Internet data set of sports events refers to the standardized data collection structure and rules that can be constructed according to the characteristics of network data of sports events. Among them, the common Internet data collection framework is generally realized by constructing database catalogue and metadata format. At the same time, the collection of event network data set is different from the current sports event state data collection and physiological state data collection, which is non-real-time and delayed. Common event data collection focuses on athletes' physiological state and real-time state data information during the event, while the Internet data set of sports events is generally public information set related to the event, which is published and non-real-time data information. In order to effectively collect the event network data and build
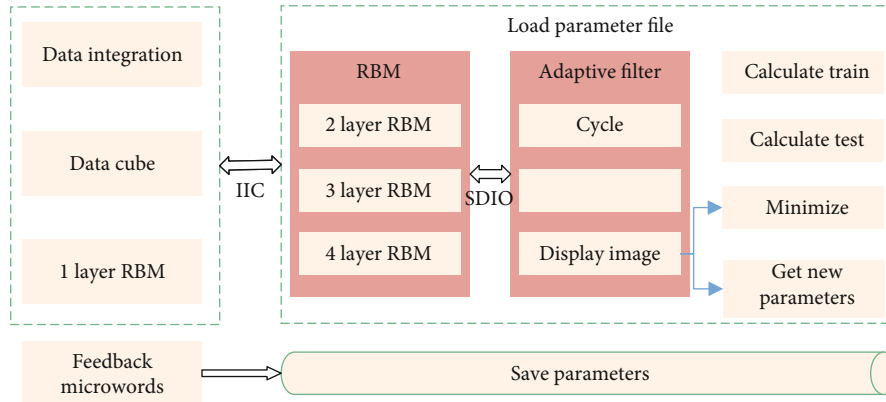
Figure 1: Recommended model for sports events.

an effective data storage warehouse, the attributes and categories of the event data in the current network should be clearly defined, so as to establish the corresponding data collection and storage table. According to the current scholars' research on the elements of sports events, events are generally composed of various elements such as event attributes, human resources, competition, and evaluation. From the perspective of sports information resources, the composition of competition network information includes many types of elements, such as competition news information, database resources, competition video resources, organizational social resources, and so on.

This paper studies the event recommendation model to meet such needs. The construction of event recommendation model first needs to clarify the current application scenarios of event recommendation and the data types required for recommendation, so as to select the appropriate recommendation model algorithm according to reasonable recommendation objectives and effective event data.

According to the research direction of this paper, the construction of event recommendation model needs to consider three elements: sports events, users, and algorithms. From the perspective of users participating in events or browsing event information, users browse event details or attribute information and upload event participation details when browsing event information. Therefore, from the perspective of the association between network users and events, the event network data can be divided into three categories: event attribute data set, user attribute data set, and user participation in events data set. As shown in Figure 2, this paper divides the event Internet data into three parts, thus constructing the corresponding data table. In each data set, there are field names to be collected under the data set. According to the standard of constructing the third normal form according to the data table, each field name is independent and does not repeat, which represents a data feature under the data set. At the same time, in the process of data storage, the network data of the three types of events contains the data corresponding to all the field information in the data table, which is regarded as the metadata of a data table, also known as tuples. In order to effectively collect the characteristic data in the network data of events and make the data set conform to the information details of

sports events, this study clarifies the inherent field names under the three data sets, so that the collection framework can be matched to the network data collection process of various events.

Among them, the event attribute network data set should include five field names: event number, event name, event venue, event date, and event introduction. The event number is the primary key of the data set, that is, the necessary field name. The user attribute network data set includes four field names: user number, user name, user gender, and user location, wherein the user number is the necessary field name of the data set. The network data set of user participation in the competition is associated with the other two tables and has the competition number, user number, and necessary user participation number. As shown in Figure 3, when the collection framework is applied to the network data collection of various events, the corresponding data sets and field names should be determined according to the event-related public contents to be collected. In the data table composed of three data sets, the event number, user number, and user entry number are the primary keys of each data table, that is, this field is the key field to determine the uniqueness of data in the data set. According to the different network data of various sports events, you can choose to add other fields to form a corresponding reasonable data table.

3.3. Digital Management of Event Information. In the era of big data, data analysis cannot be separated from reasonable search and collection of data. At present, the collection of large quantities of Internet data has entered the period of automatic collection, which is also called network data collection or network crawler. Web crawler technology has played a great role in scientific research, public opinion collection, and information security. Through Web crawler technology, regular data information can be obtained in large quantities according to the set program content. At present, the crawler technology based on python language is the most widely used, and the personalized website data collection framework can be written through python language. At the same time, a large number of data are collected with the help of plug-ins, among which the commonly used plug-in modules include web page request module, scrawny crawler framework, and selenium automated web page test
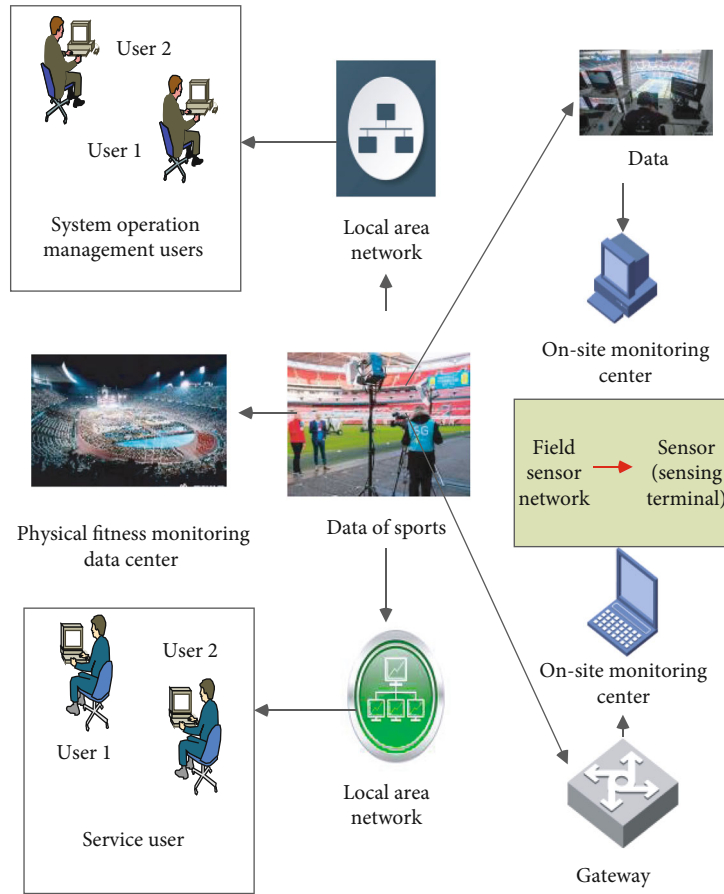
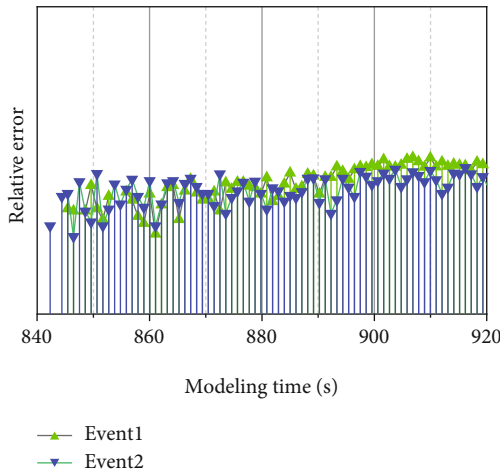Figure 2: Event network data collection framework.
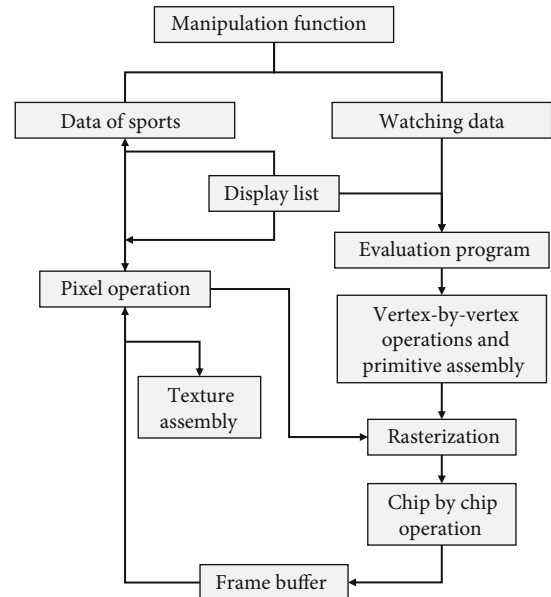

Figure 3: Collection of event attribute data sets.

Figure 4: Implementation of web crawler.

framework. The implementation flow of web crawler is shown in Figure 4.

The collection of Internet data of sports events can also be expanded according to web crawler technology. First of all, it is necessary to determine the target website and the target content that need to collect data, through Python language or other computer language to write the corresponding website content request module code, content analysis module code, and content collection module code. Then, batch event data collection is realized according to the order of requesting or taking event data website content, analyzing event data content, and collecting corresponding field data. At the same time,

it is necessary to make it clear that the collection of corresponding data should conform to the robots protocol of the Internet and only be used for academic research, so as to ensure the reasonable and legal collection behavior.

Based on the corresponding user needs when constructing the event recommendation model, the main function of the model after construction is to realize the recommendation between the same type of sports events and have a certain accuracy to meet the basic needs of users. From the use scenarios of the event recommendation model, there are uncertain differences in the event characteristics and data characteristics of different categories of events. According to different competition requirements, it is necessary to fine-tune the algorithm under a unified framework in the construction of the competition model, so as to meet the effective recommendation under different input data conditions, different competition project recommendation requirements, and different use scenarios. Therefore, the constructed event recommendation model needs to have a statistical basic framework, and at the same time, according to the project objectives, it needs to meet the recommendation needs of different sports characteristics. In terms of actual functional requirements, the functions that the event recommendation model should realize include feature extraction of events, similarity calculation of events, and recommendation list supply of events.

4. Experiences and Results Analysis

4.1. Data Requirements Analysis. The construction of the recommended model needs to meet its usage scenarios and performance requirements, which makes the construction of the model follow the target direction and data feature dimension of the entity project. The event recommendation model constructed in this paper is based on the analysis of network data characteristics of events, which makes it clear that the current network data of events has the characteristics of large amount of text data and uneven and diversified data distribution. Under such characteristics, the construction of recommendation model needs to be as close to the characteristics of most public network data as possible. In the selection of input data, we should select characteristic data with clear classification attributes and considerable quantity. Traditional recommendation models often need a large number of data sets with standardized structure to support them. In the past studies, the information research on sports events tends to be carried out on the theoretical framework and data structure, while the research on the characteristics of event data in the current Internet is rarely mentioned. The reason lies in the lack of effective analysis means and data support. The characteristics analysis and collection of network data of current marathon events as an example can provide more suitable basic data for the research of event recommendation model. With the corresponding basic data set, the construction of the event recommendation model will be more accurate. From the technical feasibility point of view, the rapid development of artificial intelligence and big data provides more feasible directions for the selection and comparison of model algorithms, as shown in Figure 5. The rapid evolution of natural language
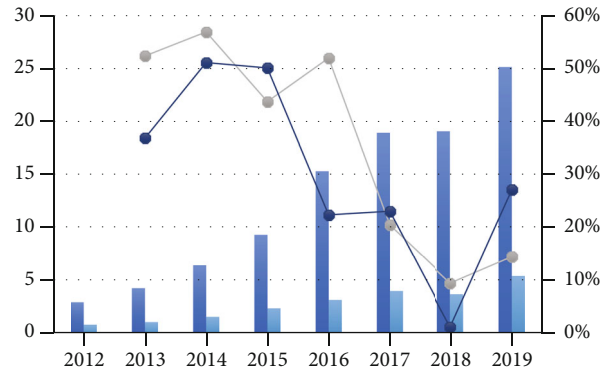

FIGURE 5: Model feasibility requirements.

processing technology and machine learning technology promotes the practice of various recommendation algorithms in more industrial production fields. According to the data characteristics and the needs of runners, the recommendation technology algorithm selected in this paper has been effectively proved in other fields such as commodity recommendation, text recommendation, and news recommendation, so as to ensure that the algorithm selected in the construction of event recommendation model in this study will be followed.

4.2. Recommendation System. In order to evaluate the performance of various recommendation algorithms or systems conveniently, academia and industry have a series of evaluation indexes which can be used to evaluate the reliability of recommendation algorithms or systems. Different evaluation indicators have different emphases in measuring recommendation performance and correspond to different evaluation approaches. This section summarizes some commonly used evaluation indicators in academic circles, including recommendation accuracy, recommendation coverage, and user satisfaction. Users' satisfaction with the recommended items is one of the important indicators to evaluate the recommendation model. However, user satisfaction cannot be obtained by offline calculation, which requires user survey and real-time collection. At the same time, in the online system, user satisfaction needs to be obtained by collecting some user behaviors and making statistical analysis.

Prediction accuracy is the most important index to measure the offline evaluation of recommendation system. Among them, it is mainly divided into scoring prediction accuracy and using prediction accuracy. According to different research directions, the commonly used prediction accuracy indicators are as follows: mean absolute error (MAE), precision, and recall. MAE uses absolute value to calculate the recommendation error and observes the gap between the predicted score of items given by the recommendation algorithm and the actual score of users to measure the performance of the recommendation system. Accuracy and recall rate are widely used in Top-N recommendation. Top-N recommendation gives users a recommendation list of corresponding items with the number of N, which is the

mainstream recommendation scheme of recommendation system at present. Coverage is to evaluate the mining ability of a recommendation system or recommendation algorithm for long-tail items. The most common definition is the proportion of items recommended by the recommendation system to the total collection of items. Coverage rate is often used to evaluate the recommendation performance of books, movies and other items with complex classification. Combined with the above-mentioned Internet data set characteristics of sports events taking marathon events as an example, in this study, the actual test of the event recommendation model will use the evaluation method of prediction accuracy to test the results of the event recommendation model based on the content recommendation algorithm.

According to the above research on the category of recommendation model, this paper chooses content-based recommendation model technology to build the corresponding event recommendation model according to the sports event data set on the Internet. The content-based recommendation model algorithm mainly uses the descriptive content features of the entity to be recommended and calculates the number of tags or the similarity of tags content through the tagging vector of content features. It can be seen that the content-based recommendation algorithm replaces entities with tags of feature content, and each tag has different corresponding values, thus transforming the feature distribution problem of entities into the vector value problem of entity tagging and realizing the calculation of vector value distance instead of similarity. In the Internet data of sports events, the descriptive information of sports events is mostly unstructured feature data, which is manifested as event name, event introduction, event location, and entry requirements. This kind of descriptive information is mostly distributed in text content, and the text length is different. By browsing this kind of text information, users or visitors can quickly form a preliminary understanding of the competition situation, as shown in Figure 6.

Therefore, for this kind of unstructured feature text, it is generally necessary to use the corresponding text processing algorithm to transform the feature content into space vectors. At the same time, by observing the event data set and data characteristics, we can find that there are a large number of user entry record sequence data in the Internet data of sports events. This kind of data belongs to unstructured data type, but the text in its sequence is mainly the name of the event, and the content features are hidden in the entry records. Therefore, for the application of this kind of data set, it is necessary to extract the hidden features from the sequence records by algorithms and attach them to the event entities or user entities. This transforms the feature similarity problem of events into the hidden feature distribution problem in the event sequence. In the current research, the text content feature extraction algorithms commonly used in content-based recommendation model include the LDA topic model algorithm, TF-IDF text weight model, and Word2vec model. These three algorithms have achieved the extraction of text features in sentence segments through different concepts and have achieved success in a large number of experiments and practical applications. At the same time, the algorithm idea of Word2vec model is also widely used in the data set of sequence data or behavior records, which can obtain the vector space values of each entity in the sequence. In the following, these three content text feature extraction algorithms are described and analyzed.

Vectorization of events introduces the concept of vector space model (VSM), which is defined as transforming traditional text content into dimension vectors in vector space, thus giving corresponding values for calculation, so that similar documents or paragraphs have similar vector spaces. Vector transformation makes entity content from text information to numerical information, which makes it easier to carry out statistics of entity content attributes. For example, in marathon events, "Shanghai International Marathon," "Beijing International Marathon," and "Beijing International Cross-country Running Challenge" are all top marathon events at present, but it is difficult to judge the similar intensity of the two events from the classification attributes of the events. For example, "Shanghai International Marathon" and "Beijing International Marathon" are both regular paved marathon events, but their venues belong to different cities. The "Beijing International Marathon" and "Beijing International Cross-country Running Challenge" are both held in Beijing, which are geographically similar events, but their classification is different, and the suitable participants are also different. However, as shown in the vectorization of event names in Figure 7, the vector space model transformation based on word frequency (the greater the word frequency, the weaker the feature performance and the smaller the value) can extract the values of similar texts from the text attributes and give effective space vector values to these three events. The corresponding event vector model can be constructed under a large amount of training data.

Event similarity calculation is a necessary step to achieve event recommendation, which transforms the similarity problem between events into the distance problem of spatial values, in which it is convenient to use mathematical and physical calculation forms to realize the similarity judgment of event entities. And the recommended list can be output conveniently according to the ranking of similarity between events. At present, there are many methods to calculate the similarity of VSM vector space models. Common vector space similarity algorithms include Euclidean distance, cosine similarity, and Pearson's correlation coefficient. At the same time, when using different algorithms to build vector models, the calculation of entity similarity also needs a specific adaptive similarity calculation method. For example, when using the LDA theme model to build "event-theme distribution matrix" in this study, it is necessary to use the corresponding theme proportion similarity calculation method to realize the similarity calculation of theme proportion. Therefore, in the following research, this paper will compare and analyze the commonly used vector similarity calculation methods and study their applicability, so as to choose the appropriate similarity calculation method and apply it to the modeling of each event recommendation model. As the mainstream topic generation and topic vectorization model in text content analysis, LDA model holds that a document corresponds to multiple topics, and each topic corresponds to a different vocabulary in the document.
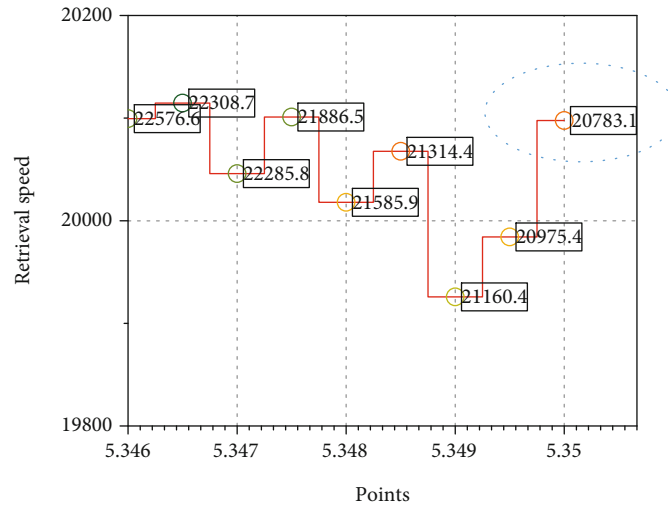
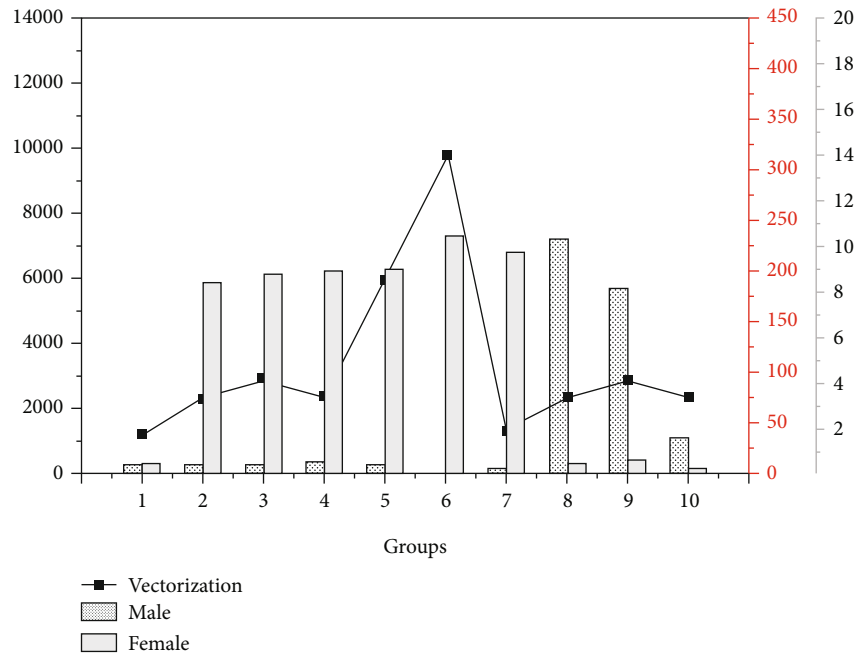Figure 6: Text class information retrieval speed.

Figure 7: Vectorization of event names.

According to the lexical distribution probability within the topic, the corresponding content range of the topic can be summarized. According to the proportion of documents on different topics, the topics that account for a larger proportion can be regarded as the main topics of the document. Three-dimensional "document-topic-vocabulary" constitutes a necessary condition for the proportion of document generation vectors. When inputting the original text, the LDA topic model adopts the basic word bag representation and transforms each input document content into a corresponding word frequency vector. In LDA modeling Internet event data, it is necessary to model the text content of all events after word segmentation. In order to effectively express the characteristics of events, the characteristic text

content of events generally includes the name of events, the venue of events, and the brief introduction of events. These data indicators reflect the content of events in theme characteristics, regional tendentiousness, and overall characteristics. After text preprocessing such as word segmentation, the input original event content text is changed into a content feature entry matrix for modeling. As shown in the event-theme combination content based on LDA theme model in Figure 8, when modeling the event-theme through LDA model, the characteristic contents of m events, R is input, and the number of possible themes T of a given event is k. After training the model code, the event-theme matrix and theme-content vocabulary matrix can be finally obtained.
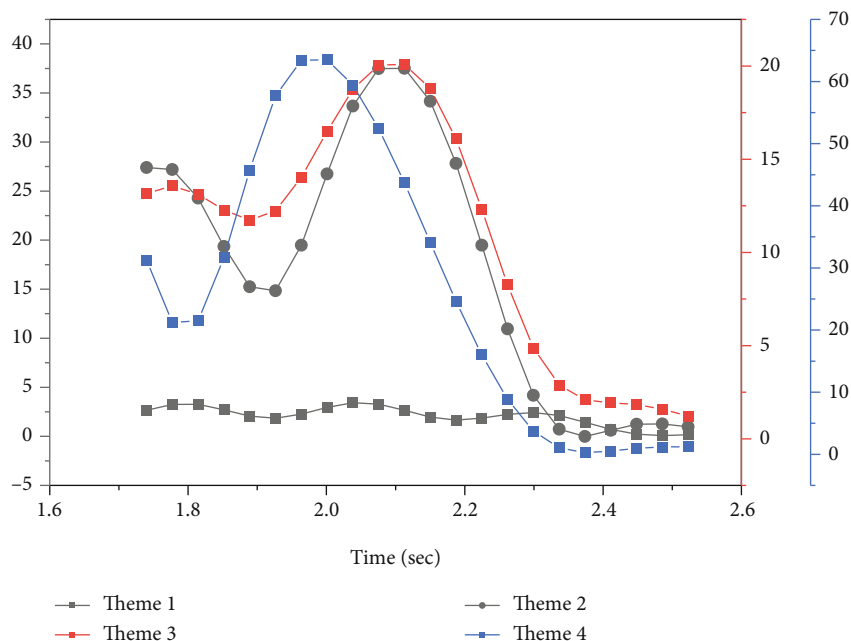
FIGURE 8: Theme combination of theme model.

The basic assumption of TF-IDF model is that a corpus contains multiple documents. If a word in a document appears many times in the document and is marked with the TF value at the same time, but the word appears less in the whole corpus and is marked with IDF value, then the larger TF × IDF mark value of this word may be the keyword or subject word of the document to which it belongs. According to the rules of the TF-IDF algorithm, the corresponding TF-IDF value of words in each event name can be calculated. By observing the TF-IDF vector values of words in the event corpus, we can find that "Shanghai" is the key theme of "Shanghai International Marathon Events." "Cross-country Running" and "Challenge" are the key themes of "Beijing International Cross-country Running Challenge." It also shows that TF-IDF algorithm can effectively separate the feature themes of each event in the event text corpus and endow the event with a certain vector dimension. In practice, the improved model is no different from the basic Word2vec model in algorithm principle, but the selected input data set is changed from the event feature text data in the Internet to the sequence data of users participating in the event. Then the output content changes from the vector value of constructing characteristic vocabulary to the vector value of each event in the input event sequence set. This transformation is more suitable for constructing corresponding spatial vectors for event entities and calculating the similarity of event vectors. At the same time, in the use of data, it can avoid using a single event feature text data, which cannot effectively verify the recommendation accuracy of the recommendation model and turn to using the user competition sequence data in the Internet. Finally, it can compare and analyze the performance of the recommendation models when using different data sets.

After the above description, it can be seen that the Internet data sets of sports events have the characteristics of large number and many types, and a single content-based recommendation algorithm modeling method is easy to achieve better recommendation results on specific Internet event data input samples. However, from the perspective of user selection of entities, it lacks universal applicability. For example, the LDA topic model requires a high number of feature texts in the input data set, and the selection of the number of subjects will affect the numerical size of the final topic vector space of each entity. The event recommendation model based on TF-IDF can construct the spatial vector of each event according to the feature text of the event. However, if the feature text description of sample events is too little or different language description methods are adopted, the phenomenon of vector deviation in feature space will easily occur, which will make the events with similar types or properties have a long spatial distance in vector space and affect the final recommendation efficiency. Similarly, in the event recommendation modeling based on Word2vec sequence model, the insufficient sample size of users' entry records or uncertain entry types in the network will easily affect the generation of event space vectors. If a single user only participates in the same race for many years or professional runners participate in the race with too wide type or region and large span, it is easy to have the phenomenon that two marathon events with low correlation are close in vector space. Therefore, in order to reduce the influence of single algorithm and input samples on marathon event recommendation, the author considers combining the TF-IDF model with the Word2vec sequence model, so that they can fuse and calculate the event similarity matrix after the vectorization of events and propose a comprehensive event recommendation model and compare the performance difference between the comprehensive recommendation model and other single algorithm models under the test sample data, so as to improve the performance and accuracy of

event recommendation. At the same time, when the comprehensive model carries out the event vectorization step, it can directly call the event vector matrix generated by a single TF-IDF model and Word2vec sequence model, so as to save computing resources and improve the fault-tolerant space of the comprehensive model.

5. Conclusion

Taking the marathon event data in combustion network as an example, this paper discusses the algorithm selection and model construction of event recommendation model and determines the content-based recommendation technology and three key algorithms commonly used to achieve the goal of event recommendation model construction. On the basis of the above, this paper further studies the construction framework of event recommendation model under three key algorithms, focusing on the steps of event vectorization and event similarity calculation, and makes experimental comparison on the collected marathon event data set. From the results, the constructed event recommendation model has a good performance in the marathon event recommendation, which verifies the feasibility of content-based recommendation technology in the event information recommendation. This can effectively meet the needs of the existing people for marathon event recommendation and also provide technical support and theoretical basis for the research of building an effective Internet data processing mechanism and event recommendation model of sports events. This paper studies the Internet data of current sports events, which has the characteristics of huge quantity, various types, fragmentation, and low correlation, and builds a general event data collection framework and collection method. Taking the popular marathon event among the masses as an example, this paper discusses and analyzes the similar and unique characteristics of its network information. Through the collection and statistics of marathon data, the characteristics of data diversity are verified, which provides basic data support for the construction of event recommendation model.

In the future, it gives the possibility of integration and development of sports event information and related information technology from the perspective of technology, which provides a broader idea for sports informationization research and enriches the technical means of sports research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P. Haje, A. Arystanbaeva, Z. Oralbaeva, and Z. Kupenova, "The role and importance of accounting information system in the context of digitalization," *Asian Journal of Social Sciences and Humanities*, vol. 5, no. 1, pp. 64–73, 2019.

[2] M. McKee, M. C. I. van Schalkwyk, and D. Stuckler, "The second information revolution: digitalization brings opportunities and concerns for public health," *European Journal of Public Health*, vol. 29, Supplement 3, pp. 3–6, 2019.

[3] Y. Shrayberg and K. Y. Volkova, "Features of copyright transformation in the information environment in the age of digitalization," *Scientific and Technical Information Processing*, vol. 48, no. 1, pp. 30–37, 2021.

[4] T. Pistunova, V. Slabko, K. Averina, I. Romanchenko, O. Babakina, and O. Vidomenko, "Providing the minds of the out-of-the-box digitalization as the basic component of the information culture of the teacher," *Laplage Em Revista*, vol. 7, no. 2, pp. 691–696, 2021.

[5] T. Rudakova and M. Semikolenova, "Evolution of accounting information model under conditions of changing economic formations and digitalization of the economy," *Electronic Markets*, vol. 12, no. 1, pp. 108–115, 2018.

[6] A. Rafisovich Gapsalamov, E. Munirovich Akhmetshin, R. Rashatovich Sharipov, V. Lvovich Vasilev, and T. Nikolaevna Bochkareva, "Approaches to information security in educational processes in the context of digitalization," *TEM Journal*, vol. 9, pp. 708–715, 2020.

[7] J. Tuzii, "Healthcare information technology in Italy, critiques and suggestions for European digitalization," *Pharmaceuticals, Policy and Law*, vol. 19, no. 3-4, pp. 161–176, 2018.

[8] E. Hofmann, H. Sternberg, H. Chen, A. Pflaum, and G. Prockl, "Supply Chain Management and Industry 4.0: conducting research in the digital age," *International Journal of Physical Distribution & Logistics Management*, vol. 49, no. 10, pp. 945–955, 2019.

[9] I. G. Utomo and G. S. Darma, "Measuring optimization of digital military programs: an innovation of information and communication systems in Industrial Digitalization 4.0," *International Research Journal of Engineering, IT and Scientific Research*, vol. 6, no. 2, pp. 39–46, 2020.

[10] S. D. Karakozov and N. I. Ryzhova, "Information and education systems in the context of digitalization of education," *Journal of Siberian Federal University*, vol. 12, no. 9, pp. 1635–1647, 2019.

[11] S. A. Dyatlov, O. S. Lobanov, and W. B. Zhou, "The management of regional information space in the conditions of digital economy," *Economy of Region*, vol. 14, no. 4, pp. 1194–1206, 2018.

[12] S. Treber and G. Lanza, "Transparency in global production networks: improving disruption management by increased information exchange," *Procedia CIRP*, vol. 72, pp. 898–903, 2018.

[13] Z. Wei and L. Sun, "How to leverage manufacturing digitalization for green process innovation: an information processing perspective," *Industrial Management and Data Systems*, vol. 121, no. 5, pp. 1026–1044, 2021.

[14] A. V. Nesterov, "Digitalization of society and the economy: systematization of personal data in information systems," *Scientific and Technical Information Processing*, vol. 47, no. 2, pp. 133–138, 2020.

[15] J. Y. Baek, A. Jo, S. S. Yu, T. B. Kim, and H. J. Oh, "Development and application of evaluation factors for the digitalization of intangible cultural heritage record information resources," *Journal of Korean Society of Archives and Records Management*, vol. 19, no. 3, pp. 123–153, 2019.

[16] T. M. Lappi, K. Aaltonen, and J. Kujala, "Project governance and portfolio management in government digitalization," *Transforming Government: People, Process and Policy*, vol. 13, no. 2, pp. 159–196, 2019.

[17] S. N. Bobylev, O. Y. Chereshnya, M. Kulmala et al., "Indicators for digitalization of sustainable development goals in PEEX program," *Geography, Environment, Sustainability*, vol. 11, no. 1, pp. 145–156, 2018.

[18] M. L. Martín-Peña, J.-M. Sánchez-López, and E. Díaz-Garrido, "Servitization and digitalization in manufacturing: the influence on firm performance," *Journal of Business & Industrial Marketing*, vol. 35, no. 3, pp. 564–574, 2019.

[19] M. U. Hulla, M. Hammer, H. Karre, and C. Ramsauer, "A case study based digitalization training for learning factories," *Procedia Manufacturing*, vol. 31, pp. 169–174, 2019.

[20] J. Björkdahl, "Strategies for digitalization in manufacturing firms," *California Management Review*, vol. 62, no. 4, pp. 17–36, 2020.

[21] M. Stoyanova, "Good practices and recommendations for success in construction digitalization," *TEM Journal*, vol. 9, no. 1, pp. 42–47, 2020.