English Literature Appreciation Teaching Resources Retrieval System Based on Mutual Information Entropy

Xiaojun Jiang[1(✉)] and Haiquan Chi[2]

[1] School of Translation Studies, Xian Fanyi University, Xi'an 710105, China
seanrain2022@163.com

[2] Liao Ning Provice Zhangwu Senior High School, Zhangwu 123200, China

Abstract. With the development of globalization and the increase of international communication, the appreciation of English literature has become an important part of English learning for many learners. English literature appreciation aims to help students understand and appreciate the artistic value, cultural connotation and language expression of English literary works. The imbalance of resource allocation is the main reason why the application host can not accurately retrieve teaching information. In view of the above problems, a retrieval system for English literature appreciation teaching resources based on mutual information entropy is designed. According to the calculation result of mutual information entropy, the generalized entropy function is solved, and then based on this, the processing of teaching resources is completed, so as to achieve the acquisition of teaching resources based on mutual information entropy. Improve the retrieval architecture, and through the definition of resource retrieval methods, realize the application of various technical means, and complete the design of English literature appreciation teaching resource retrieval system based on mutual information entropy. The experimental results show that the application of the above retrieval system can make the maximum balance of resource allocation reach 95%, which solves the problem that the application host cannot accurately retrieve teaching information due to the imbalance of resource allocation.

Keywords: Mutual Information Entropy · English Literature Appreciation · Teaching Resources · Resource Retrieval · Generalized Entropy Function · Retrieval Architecture · Ratio Balance

1 Introduction

English literature refers to literary works written in English. English works in American literature can be counted as part of English literature, but they are usually treated as an independent and important discipline; The same is true of Irish literature. Although there are many repetitions between English literature and English literature, they include literature from other regions and languages in Great Britain, so they are two different

G. Gui et al. (Eds.): eLEOT 2023, LNICST 543, pp. 156–171, 2024.
https://doi.org/10.1007/978-3-031-51465-4_11

concepts. In academic circles, "English literature" is often analyzed and criticized in departments or projects devoted to "English research". The reason is that the former colonies of England have developed their own English literature, and the English used in these places is different from each other [1]. So English literature has developed various branches with the changes of English in the world. Literary appreciation is to obtain aesthetic pleasure and spiritual satisfaction through the interpretation of the language symbols of the works. It is divided into three stages: perception, taste, and understanding. It has differences and consistency, certainty and uncertainty. It is the reader's psychological activities such as imagination, association, emotion, thinking, and recreation based on the understanding of literary works to meet aesthetic needs, to pursue the readability and interest of theoretical works.

The construction of educational resources is to integrate educational and teaching resources by means of information technology, and finally build an interactive and multimedia shared resource warehouse. The purpose of building an educational resource database is to serve education. Therefore, we should fully consider the needs of education in terms of both content and function, so that students, teachers and other educators can easily and timely obtain the information they need and have usability. On the basis of understanding the needs of users, we must carry out anti demand analysis, that is, combining with the actual situation, scientifically analyze and express the demand information provided by users from a more professional perspective. At present, the construction of basic education resources exists: there are many ways to construct education resources, but the two main ways are: 1. Directly purchase the existing commercial education resources products in the market; 2. Organize school teachers to develop educational resources by themselves.With the continuous expansion of data on the Internet, there are a lot of valuable education information, and online education resources have gradually become an important source of education resource database construction. At present, there are many kinds of educational resources on the Internet, which provide learners with a wide range of choices and broad space for development.

For the construction of English literature appreciation teaching project, the vertical retrieval system based on Lucene and the retrieval system based on Grassberger estimate judge the unit cumulative amount of teaching resource information according to the two parameters of browsing times and authoritative sources, and analyze and read XML documents related to teaching tasks with the help of the open source software package Dom4j. However, the above two types of systems cannot guarantee the balanced allocation of teaching resources, so it is easy to make the application host unable to retrieve accurate teaching information. In information theory, mutual information entropy is the average amount of information contained in each received message, also known as information entropy, source entropy, average self information amount [2]. Here, messages represent events, samples, or characteristics from a distribution or data flow. In the information world, the higher the entropy, the more information can be transmitted. The lower the entropy, the less information can be transmitted. Literature [3] designed a density-based clustering algorithm based on MapReduce based on weighted grid and information entropy. Based on the spatial distribution of data points, an adaptive partitioning strategy (ADG) was proposed to carry out adaptive partitioning of grids. A weighted grid construction strategy (NE) is designed, and a density

computing strategy (WGIE) is designed to calculate the density of the grid based on weighted grid and information entropy. A core cluster computing algorithm based on MapReduce (COMCORE-MR) is proposed to compute the core cluster of clustering algorithm in parallel. Literature [4] Visualizes citation navigation using Litmaps, the ultimate scientific discovery platform. It provides an interface for discovering scientific literature, exploring areas of research, and discovering articles that are highly relevant to the map. Litmaps provides quick-start options to import articles from the reference manager, keyword search, ORCID ID, DOI, or using torrent articles. This paper uses the research strategies of keyword search and Open Educational Resources (OER).To solve the above problems, a new retrieval system for English literature appreciation teaching resources is designed based on the principle of mutual information entropy. Based on the principle of mutual information entropy, a complete generalized entropy function is defined, and the necessary search conditions are combined to realize the on-demand processing of teaching resources. Generalized correlation function is used as the classification feature of information entropy, and support vector machine classifier is used to realize the classification retrieval of English literature appreciation teaching resources. By improving the connection form of the search structure and combining the necessary search conditions, the design of the English literature appreciation teaching resource retrieval system is completed.

2 Acquisition of Teaching Resources Based on Mutual Information Entropy

Obtaining teaching resources is a necessary link to build a retrieval system for English literature appreciation teaching resources. Supported by the principle of mutual information entropy, a complete generalized entropy function is defined and necessary retrieval conditions are combined to achieve on-demand processing of teaching resources.

2.1 Calculation of Mutual Information Entropy

Mutual information entropy is an important concept in information theory, which can be used to describe the statistical correlation between two systems, or the extent to which one system contains information about another system. The dependence between two variables can be measured by calculating the information entropy of two variables and the gap between their joint entropy, which can be used to express the amount of information shared between two variables.

Generally speaking, mutual information entropy can represent the received random variable $\tilde{p}$ Post acquired variables about teaching resources $\dot{O}$ Information quantity of: it can measure the degree of the reduction of the uncertainty of the other when one of the two variables is known, indicating the degree of statistical constraints between two random variables. For example, if $\dot{O}$ and $\tilde{p}$ Independent of each other, then the random variable is received $\dot{O}$ Will not get any information about $\tilde{p}$ And their mutual information entropy is 0. Discuss another extreme case, if $\dot{O}$ and $\tilde{p}$ One to one correspondence, they are mutually deterministic functions, without any uncertainty, and share all the information transmitted. In this case, mutual information entropy and $\dot{O}$ (or $\tilde{p}$) The uncertainties

provided separately are the same, and the mutual information entropy is equal to $\dot{O}$ Entropy of ($\tilde{p}$ Entropy).When $\dot{O}$ and $\tilde{p}$ It is also applicable to this situation when it is the same random variable.

Therefore, mutual information entropy can be used to measure the correlation between specific classes when they are divided. When the amount of information is large, it indicates that there is a large correlation between the two, and vice versa.

The calculation formula of information entropy is:

$$\tilde{p} = \sum_{\delta=1}^{+\infty} \beta \dot{O} \cdot \left[\frac{(\chi - 1)^2}{\bar{I}} \right] \tag{1}$$

Among them, δ Represents the information transmission coefficient of teaching resources, β It indicates the probability of the retrieval of English literature appreciation teaching resources, χ It represents the evaluation variable of teaching resource information, $\bar{I}$ It refers to the unit accumulation of English literature appreciation teaching resource information.

The amount of information contained in a random event is only related to the probability of the event. The lower the probability of an event occurring, the greater the amount of information contained in the received information when the event actually occurs. That is, when the probability of an event occurring is infinitely close to 0, the amount of information corresponding to its occurrence will be very large, otherwise the amount of information will be very small. The mutual information entropy can be obtained by calculating the mathematical expectation of the amount of information.

The calculation result of mutual information entropy is:

$$\partial_{\tilde{p}} = \log \left| \hat{i} \right| \frac{1}{\alpha \cdot \tilde{p}} \tag{2}$$

$\hat{i}$ Represents the unit expectation parameter of the information entropy index.

Mutual information entropy can not only describe the linear correlation between variables, but also effectively characterize the nonlinear correlation. The mutual information entropy feature in information theory has a high accuracy rate when retrieving English literature appreciation teaching resources in a low SNR environment, so it is feasible to use it as a classification feature.

2.2 Generalized Entropy Function

The teaching resource detection system based on mutual information entropy function performs Wigner bispectrum analysis on the real target echo or deception jamming signal, calculates their DSWB, and uses singular value decomposition to achieve dimensionality reduction, and then uses the equidistant algorithm to obtain the dimensionality reduction with the LFM information WB, the information entropy of the data after singular value decomposition and the joint entropy between the two. The generalized correlation function is calculated as the classification feature of teaching resource information, and the classification recognition of deception jamming signal is realized by SVM classifier.

The specific function operation steps are as follows:

The retrieval system receives teaching resource information and mutual information entropy parameters.

Wigner bispectrum analysis is carried out on the teaching resource information to obtain the three-dimensional WB retrieval data.

DSWB is calculated to obtain two-dimensional DSWB data of time retrieval frequency, and singular value decomposition is used to reduce the dimension of mutual information entropy.

The information entropy of the data after dimensionality reduction and singular value decomposition, as well as the joint entropy between them, is calculated by using the equidistance algorithm and the linear resource information WB.

The generalized correlation function is calculated as the classification feature of information entropy, and the SVM classifier is used to realize the classification and retrieval of English literature appreciation teaching resources.

The process of defining the generalized entropy function of English literature appreciation teaching resource information is shown in Fig. 1.

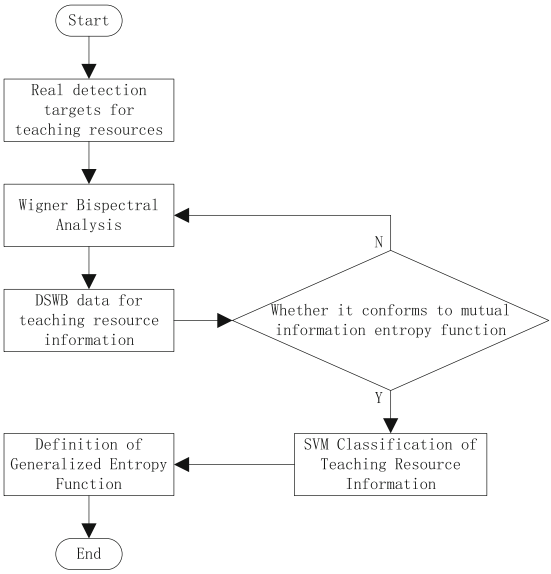

Fig. 1. Flow chart of definition of generalized entropy function of teaching resource information

The solution to the generalized entropy function of English literature appreciation teaching resources information satisfies the following expression:

$$U = \min_{\substack{e=1 \\ r=1}} |\dot{y}|^{(e-r)^2} + \partial_{\tilde{p}} \sum_{-\infty}^{+\infty} (u_{\max} - u_{\min})^2 \tag{3}$$

where, e、r Represents two randomly selected entropy increasing vectors, and $e \neq r$ The inequality value condition of is always true, $\dot{y}$ It represents the real-time retrieval

characteristics of English literature appreciation teaching resources, $u_{\max}$ Represents the maximum value result of the joint entropy parameter, $u_{\min}$ Represents the minimum value of the joint entropy parameter.

For a given number of teaching resource information, the larger the grid is divided, the more points will be in the grid, and the calculation of average probability will be more accurate, but the joint probability distribution obtained will change more smoothly, and the mutual information entropy will be lower; The smaller the lattice is, the greater the possibility that the joint probability distribution will change in a short distance, the more intense the change fluctuation, and the greater the mutual information entropy. Therefore, it is very important to select the appropriate lattice size for accurate estimation of mutual information entropy.

2.3 Processing of Teaching Resources

Since the query text input by the client on the query interface is natural language, the system must carry out a series of processing on the query text to automatically extract the key features of the query conditions and query objectives from the query English literature appreciation teaching resources. The process is as follows:

Query text → word segmentation → removing invalid words → extracting keywords → expanding synonyms.

Match the subject term after the extended synonym with the content description field in the database to find the records that meet the conditions.

For query text segmentation, the maximum positive matching algorithm is used.Relevant files used in word segmentation include: word segmentation dictionary (lastdictionary. Txt) and word segmentation function (fcresultl).

After the invalid words are removed, there are still some words in the string, which are meaningless to describe the characteristics of the query target. Therefore, it is necessary to extract words that can describe the characteristics of the query target from the remaining strings.

Using formula (3), the invalid words of English literature appreciation teaching resources information based on mutual information entropy can be expressed as:

$$\breve{W} = \sqrt{\dot{R}\frac{1}{\gamma}}\left(\frac{E'^2 - E_0^2}{U}\right)^{-1} \tag{4}$$

$\dot{R}$ Represent the descriptive characteristics of English literature appreciation teaching resources information, γ Indicates the evaluation parameters of invalid words of resource information, E' Represents the forward matching vector of resource information, E_0 Indicates the initial value of the teaching resource information matching parameter.

Because there are a large number of synonyms and synonyms in the information of teaching resources, it is difficult for users to list all the words expressing the same concept when querying, so it is easy to miss the inspection and the recall rate of the system is not high [5].

The result of solving the processing conditions of English literature appreciation teaching resources is as follows:

$$Q = (-\phi) \int_{\gamma=1}^{+\infty} \breve{W} \left(1 + \frac{s_1 + s_2 + \cdots + s_n}{q}\right)^2 \tag{5}$$

ϕ Indicates the calibration parameters of the transmission direction of English literature appreciation teaching resources information, γ Represents the feature extraction coefficient of teaching resource information retrieval, s_1、s_2、$\cdots$、s_n express n Unequal teaching resource information string definition vector, whose values belong to $(0, +\infty)$ Value range of, q Represent resource information description parameters based on mutual information entropy.

The user's description of the query target in Chinese natural language on the query interface is presented in text form on the query interface. The retrieval system calls this descriptive text query text. The system should automatically extract the key features of the query conditions and query targets from the query text, and then find the records that meet the conditions in the database[6]. However, computers cannot directly understand natural language, which requires the processing of query text. Natural language word segmentation is a necessary part of query information processing.

3 Design Scheme of English Literature Appreciation Teaching Resources Retrieval System

On the basis of the principle of mutual information entropy, the design of the retrieval system for English literature appreciation teaching resources is completed by improving the connection form of the retrieval architecture and combining the necessary retrieval conditions.

3.1 Retrieval Architecture

The construction of retrieval architecture is based on metadata. The so-called metadata is a kind of structured data about the information or data of English literature appreciation teaching resources. It is a structured description of information resources, and it is machine understandable information for a specific application to describe resource attributes[7]. The structural characteristics of metadata can more accurately describe the semantics of resources, so that web data can be transformed from machine readable to machine understandable.

There are many metadata standards. RDF standard metadata can make the data compiled by web developers more consistent and effective in documents, and its main role is to describe and obtain metadata. RDF provides a basic dismissal for encoding, exchanging and reusing metadata on the web.

The commonly used RDF data model is the triple representation, and its basic object types are: Resource, Property, and Statement. Resources are identified by the unique resource identifier URI. Attributes are used to describe the common characteristics of

resources. Declarations are composed of a resource, a property and the value of the property. Each declaration is a triple of "object, attribute and value", which is equivalent to the subject, verb and object in the sentence.

The complete teaching resource retrieval architecture is shown in Fig. 2.

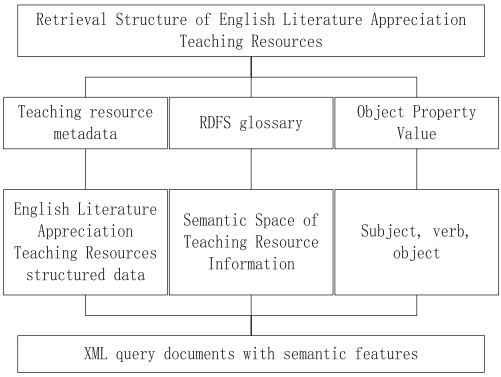

Fig. 2. Retrieval structure of English literature appreciation teaching resources

The RDF data model provides an abstract and conceptual framework for defining and using data. The data described according to the RDF specification is expressed in XML code to generate XML documents with semantic characteristics.

RDFS forms a complete semantic space. It provides the rules for using properties, defines a domain dictionary, organizes the dictionary with a type hierarchy, and provides a richer vocabulary to describe classes and class attributes.

Table 1 shows the composition of the RDFS vocabulary of the retrieval schema.

Table 1. RDFS vocabulary of teaching resource retrieval architecture

Metalinguistic object	First level project object	Secondary project objects
Information category item	rdfs: Resource、rdfs:Class 、rdfs:Property	Rdfs、subClassof and rdfs:subPropertyof
Resource Properties	rdfs:domain、rdfs:range	RDFS、DAML + OIL and OWL
Constraint effect	Apollo、OILEd、OntoEdit、OntoSaurus	WebODE、Protege 、OKBC

In order to achieve fast and accurate automatic word segmentation and part of speech tagging, it is necessary to consider the elimination of segmentation ambiguity, the recognition of unknown words and the elimination of multi category parts of speech in each link. In consideration of the above reasons, as well as the rapidity and practicality required by the system, the resource retrieval system has established a vocabulary dictionary based

on mutual information entropy. The dictionary is saved in the dic folder under the system directory with the file name lastcitian.txt.

3.2 Resource Retrieval Method

The retrieval method of English literature appreciation teaching resources with mutual information entropy as the core is as follows: first, refer to the resource type of the target information source, and with the help of a vocabulary dictionary, build the ontology of related fields; The second step is to obtain the resource information in the information source, and refer to the corresponding ontology built to make semantic annotation of the obtained data and store it in the metabase; The third step is to extract relevant concepts from the user's query statements, process the concept semantics with the support of ontology technology, and find qualified semantic information documents from the metabase; Finally, the retrieved results are processed and returned to the user [8] (Fig. 3).

The general retrieval model is shown in the following figure.

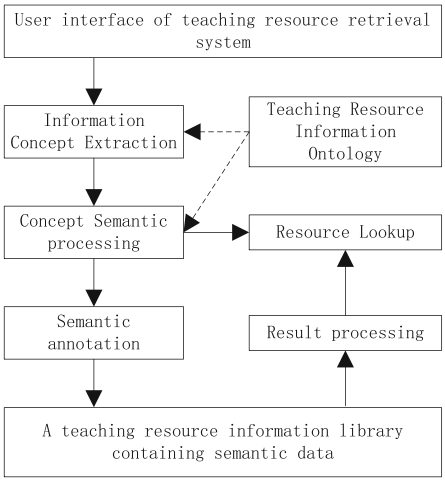

Fig. 3. Teaching resource retrieval model

Ontology is the key to realize the retrieval of teaching resources;The document library contains semantic information after semantic annotation; The feature items matched by the system are no longer keywords but domain concepts in the ontology; Conceptual semantic processing is the semantic expansion of user retrieval conditions under the guidance of ontology.

The following principles should be followed in the process of constructing the retrieval system of English literature appreciation teaching resources based on mutual information entropy:

(1) Analyze the discipline characteristics of resources, and construct the corresponding ontology of learning resources and knowledge.
(2) Create a metadata database containing semantic descriptions for learning resources.

(3) Word segmentation is required to convert the query statement entered by the user into a keyword set.
(4) Semantic expansion is needed to improve the recall and precision of retrieval.
(5) The retrieval result set should be processed to meet the needs of users.

Using formula (5), the resource retrieval expression followed by the system host is derived as follows:

$$A = \left. \frac{\prod_{a=1}^{+\infty} Q\left(\hat{d}\right)}{f \times |\Delta D|} \right|_{f \neq 0} \tag{6}$$

Among them, a Represents the return coefficient of the teaching resource search results, $\hat{d}$ Metadata definition vector representing English literature appreciation teaching resources, ΔD Represents the unit cumulative amount of teaching resource information in the metabase, f Represents the ontology retrieval coefficient.

Referring to the conceptual structure of knowledge point ontology, the corresponding semantic description of concepts is added with RDF (Resource Description Framework) language, and stored in the metadata base in the form of XML documents.Based on the concept structure of knowledge point ontology, the corresponding concept semantic information of teaching resources is expanded by using automatic semantic reasoning technology.

3.3 Key Technology Realization

The resource database building module in the English literature appreciation teaching resource retrieval system based on mutual information entropy allows users to upload learning resources to enrich the resource database, and simply describe and label resources according to the hierarchical relationship of knowledge points. In traditional retrieval systems, announcers usually use keywords to annotate resources. This standard method is simple and efficient, and is widely used; In the semantic web, such annotation information is no longer the keyword of natural language, but the concept, data attribute and value defined in the ontology.

regulations $\dot{g}$ It represents the semantic annotation features of English literature appreciation teaching resources, and the solution result is:

$$\dot{g} = \phi \times \sqrt{1 - \left| \frac{1}{A} \times \hat{h} \right|^2} \tag{7}$$

ϕ Indicates the description parameters of teaching resources, $\hat{h}$ Data attribute assignment vector representing teaching resource information.

There are two methods of semantic annotation: manual annotation and machine annotation. Among them, the manual annotation means that the standard direction is determined by the annotation personnel and the annotation items describing the resources are determined by themselves according to a certain process and under the guidance of the

ontology. This method is strictly constrained by the ontology model, and the workload of annotation personnel is large, and the efficiency of annotation completion is low. Therefore, many researchers are committed to the automatic (or semi-automatic) technology of semantic annotation [9]. This semantic annotation process has both the simplicity and efficiency of traditional keyword based annotation and the structural normalization of ontology based annotation, which is an efficient way of semantic annotation.

The calculation formula for the result of the semantic standard of English literature appreciation teaching resources is:

$$K = \vec{j} \cdot \left(1 - \frac{\dot{g}}{\varphi}\right)^{\iota} \tag{8}$$

$\vec{j}$ Real time coding vector representing semantic information of teaching resources, φ Represents automatic dimensioning parameters, ι The key value coefficient representing the information of English literature appreciation teaching resources.

Semantic reasoning has different meanings in natural language and machine language. The semantic reasoning in the teaching resource retrieval system is based on people's understanding of objective things, which is the deepening of understanding and the transfer of semantic expression; Semantic reasoning in machine language does not extend conceptual relationships as it does in natural language. It can only process formal symbolic expressions. It deals with logical relationships between symbolic objects [10]. In the semantic retrieval system, the symbolic object of machine language represents the concept of natural language, and the ontology structure defines the semantic relationship between concepts. The semantic reasoning of symbolic objects is completed using ontology technology. In a word, semantic reasoning is to enable computers to understand and understand the conceptual structure and metadata information of domain ontology, and complete the transformation from one concept to another with defined logical rules.

Using formula (8), the semantic derivation expression of English literature appreciation teaching resources is:

$$L = \frac{\prod_{-\infty}^{+\infty} \left|\lambda \cdot \tilde{H}\right|^2}{K - 1} \tag{9}$$

λ Represent the logical authentication parameters of the retrieval system for English literature appreciation teaching resource information, $\tilde{H}$ The transfer behavior vector representing the teaching posture information under the condition of mutual information entropy.

In the retrieval system, semantic reasoning is expressed in the form that when the user enters a query statement, the system returns the reasoned result set to the user one by one, in order to meet the different needs of the user. The main function of online semantic reasoning is to expand the relevant semantics based on the query statements entered by the user. Its life cycle occurs in the session phase of interaction with the user. This online reasoning relationship is also called conditional extended semantic reasoning. The feature of this retrieval method is that it can expand the connotation and extension of the query conditions entered by users, which will inevitably reduce the workload of users in multiple retrieval. After the user submits the initial query statement, after the

semantic expansion of the query statement, he or she will get a set of query statements with high relevance to the query statement. This set of query statements obtained through online semantic reasoning will greatly improve the recall and precision of information retrieval.

On the basis of formula (9), the execution conditions of key retrieval technologies can be expressed as:

$$X = \frac{1}{\dot{b}^2} \sum_{z=1}^{+\infty} v' \sqrt{\frac{L}{\psi \times \xi}} \tag{10}$$

Among them, $\dot{b}$ Represent the semantic reasoning features of teaching resource information based on mutual information entropy, z Query parameters representing English literature appreciation teaching resource information, ψ Represents the real-time query coefficient of the resource information to be retrieved, ξ It represents the semantic expansion coefficient of English literature appreciation teaching resource information.

Under the support of mutual information entropy, the system host should also avoid repeating the value of semantic annotation parameters in the process of executing the retrieval instructions in order to achieve on-demand retrieval of English literature appreciation teaching resources.

4 Example Analysis

This paper takes the English literature appreciation teaching resource retrieval system based on mutual information entropy, the vertical retrieval system based on Lucene, the estimation retrieval system based on Grassberger, the literature [3] system and the literature [4] system as the experimental objects.According to the ratio of teaching resources information under the influence of different methods, the retrieval ability of host components to teaching information was analyzed.

4.1 Experimental Environment

Use the equipment components shown in Table 2 to build the operating environment of the teaching resource retrieval system.

In order to ensure the absolute fairness of the experimental results, in addition to the different experimental methods, the connection forms of other equipment components in the experimental group and the control group are always consistent.

4.2 Principle Description

For English literature appreciation teaching resources, in the process of implementing information retrieval, the balance of resource allocation affects the precise retrieval ability of host components to information parameters. Without considering other interference conditions, the higher the balance of resource allocation, the stronger the host components' ability to accurately retrieve information parameters; On the contrary, if the balance of resource allocation is relatively low, it means that the host components have relatively weak ability to accurately retrieve information parameters.

The specific implementation process of this experiment is as follows.

Table 2. Operation environment of retrieval system

Project	Component Name	Equipment model
1	Resource information sharing chip	BASE version XC7K325T kit
2	Teaching Resource Master Microprocessor	S3C2416XH-40 main control CPU
3	Data information storage	BGA-676 SOC CORTEX-A9
4	Resource sharing host	S3C2410AL-20 equipment
5	Information encryption component	Stc89c52RC programming motherboard
6	Information trigger	DS3231MZ + SOIC-8 clock chip

- Connect the application equipment at all levels in Table 2 on demand to provide a stable transmission environment for English literature appreciation teaching resource information.
- Input the executive program of the English literature appreciation teaching resource retrieval system based on mutual information entropy into the application host, record the numerical changes of the resource allocation balance under the action of the system, and the results are experimental group variables.
- Input the executive program of Lucene based vertical retrieval system into the application host, record the numerical changes of resource allocation balance under the action of the system, and the results are the control (1) group variables.
- Input the executive program of the retrieval system based on the Grassberger estimation into the application host, record the numerical changes of the resource allocation balance under the action of the system, and the results are the control (2) group variables.
- The document [3] system executive program is input into the application host to record the numerical change of resource allocation balance under the action of the system, and the result is the control (3) group variable.
- The document [4] system executive program is input into the application host to record the numerical change of resource allocation balance under the action of the system, and the result is the control (4) group variable.
- Integrate the variable data obtained and summarize the experimental rules.

4.3 Discussion of Results

The following figure reflects the changes of experimental values of resource allocation balance in the experimental group, control (1) group, control (2) group control (3) group and control (4) group (Fig. 4).

Experimental group: With the extension of the experimental time, the balance of teaching resources in the experimental group kept increasing, and by the end of the 100min experiment, its maximum value reached 95.0%.

Control (1) group: The numerical change trend of the balance of teaching resources allocation in control (1) group is the same as that in the experimental group, but its average level is relatively low. By the end of the 100min experiment, its maximum value

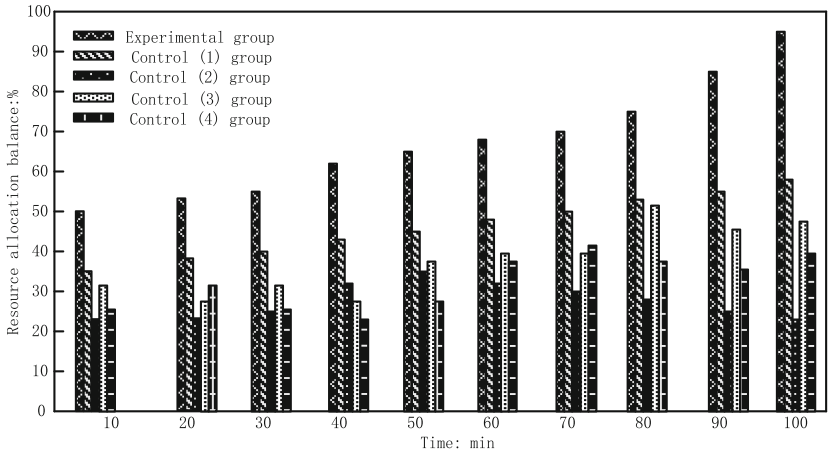

Fig. 4. Balance of resource allocation

can only reach 67.9%, which is 27.1% lower than the maximum value of the experimental group.

Control (2): Control (2) The balance of teaching resources allocation maintained the trend of increasing first and then decreasing. At the 50th minute, the maximum value was 35.0%, which was 32.9% lower than the maximum value of control (1) group, and 60.0% lower than the maximum value of the experimental group.

Control (3): The value of the system teaching resource allocation balance in literature [3] has a large trend of change, and the maximum value is 51%.

Control (4): The change trend of the value of the system teaching resource allocation balance in reference [4] is basically the same as that in control (3), and the maximum value is 43%.

To sum up, the conclusion of this experiment is:

- The application of the retrieval system based on Grassberger estimation can only keep the balance of teaching resources at a low numerical level, so the application of this system cannot solve the problem of unbalanced resource allocation.
- Although the application of Lucene based vertical retrieval system has properly improved the numerical level of the balance of teaching resources, it still does not meet the actual application needs.
- The application of the literature [3] system and the literature [4] system has appropriately improved the numerical level of teaching resource balance, but it still cannot meet the needs of practical application.
- Compared with the vertical retrieval system based on Lucene and the retrieval system based on Grassberger estimation, the application of the retrieval system of English literature appreciation teaching resources based on mutual information entropy has greatly improved the numerical level of the balance of teaching resources allocation, which can better solve the problem of unbalanced resource allocation, thus ensuring the accurate retrieval ability of the application host for teaching information.

5 Conclusion

The retrieval system of English literature appreciation teaching resources based on mutual information entropy mainly has the following application characteristics:

(1) The user's query text is segmented and part of speech tagged using the self used word segmentation function. Remove function words and default words from the query text after word segmentation, extract nouns, verbs and idioms, obtain the subject content of the target media teaching resources required by the user terminal, and expand the extracted subject content.
(2) It supports natural language query and establishes a friendly interface of human-computer interactive multimedia resource retrieval system. Users are allowed to input query text in natural language, and some text features such as file type, file size, etc. can be selected from the drop-down box to reduce the query range.
(3) The user terminal can intuitively observe the semantic characteristics of the search results and related teaching resources on the result output interface, and can enter the webpage where the media is located through the hyperlink on the interface.

In addition, since the system host searches English literature appreciation teaching resources indiscriminately, the actual value range of teaching resource information should be expanded as far as possible when defining the expression of mutual information entropy. The analysis of user query text should be further deepened. Conduct semantic analysis on user query statements, add weight to the obtained subject words, further determine the retrieval focus, and improve the retrieval accuracy. This is necessary for the further development of the system, but it is also the most difficult part to upgrade and realize. With the expansion of the database scale, when the query conditions are nested too much, the retrieval speed is relatively slow. The database retrieval algorithm should be further optimized to accelerate the speed and improve the practicability of the system.The system has some limitations in the semantic understanding and inference of user query, and can not accurately capture the user's intention and demand. In future studies, by introducing more advanced natural language processing technology, expanding data coverage, establishing user feedback mechanism and providing diversified search methods, the retrieval system of English literature appreciation teaching resources based on mutual information entropy can be further improved, and its retrieval accuracy, comprehensiveness and user experience can be improved.

References

1. Aoyagi, S., Kamochi, K., Miisho, A., et al.: Interpretation of TOF-SIMS data based on information entropy of spectra. Surf. Interface Anal. **54**(4), 356–362 (2022)
2. Jiang, S., Ding, J., Zhang, L.: A personalized recommendation algorithm based on weighted information entropy and particle swarm optimization. Mob. Inf. Syst. **2021**(4), 1–9 (2021)
3. Yu, X., Zeng, F., Mwakapesa, D.S., et al.: DBWGIE-MR: A density-based clustering algorithm by using the weighted grid and information entropy based on MapReduce. J. Intell. Fuzzy Syst. **40**(6), 10781–10796 (2021)
4. Sinhababu, A., Chakravarty, R., Kaur, A.,et al.: Visual citation navigation of open education resources using Litmaps. Library Hi Tech News **39**(5), 7–11 (2022)

5. Gao, M.: Smart campus teaching system based on ZigBee wireless sensor network. Alex. Eng. J. **61**(4), 2625–2635 (2022)
6. Rublev, V.S, Kondakov, M.D.: Automated teaching system "sets" (research for organizing the 1st part of the project). Modeling Anal. Inform. Syst. **28**(1), 90–103 (2021)
7. Zhao, H., Guo, L.: Design of intelligent computer aided network teaching system based on web. Comput.-Aided Design Appli. **19**(S1), 12–23 (2021)
8. Kardoyo, E.: Development of e-learning management model for teaching system at the police academy. Turkish J. Comput. Mathem. Educ. (TURCOMAT) **12**(5), 188–196 (2021)
9. Yinping, Z., Gencheng, W.: Research on security structure retrieval simulation of sensitive information in mobile internet. Comput. Simulat. **39**(09), 451–455 (2022)
10. Liu, K.: On the construction of teachers' professional quality-oriented english practice teaching system—exemplified with the english major of sichuan university of arts and science. Theory Pract. Lang. Stud. **11**(4), 390–395 (2021)