Research on Database Language Query Method Based on Cloud Computing Platform

Shao Gong[1](✉), Caofang Long[2], and Weijia Jin[3]

[1] School of Humanities and Communication, University of Sanya, Sanya 572022, China
1116544511@yeah.net

[2] School of Information and Intelligence Engineering, University of Sanya, Sanya 572022, China

[3] Zhejiang Tongji Vocational College of Science and Technology, Hangzhou 311231, China

Abstract. Some database language query methods have the problems of time cost and resource occupation. Under this background, a database language query method based on cloud computing platform is designed. Convert the formally described information into data that can be stored, extract the semantic Web hierarchy, and establish relationships with real objects, identify the potential association features of database words and sentences, restore the original orthogonal semantic structure space, build ontology annotation model based on cloud computing platform, transplant HanLP's neural network dependency syntax analysis tool, and design language query methods. Test results: The average time cost and resource utilization of the database language query method designed this time are 1720 ms and 42.36% respectively, which shows that the database language query method designed this time is more effective under the technical support of the cloud computing platform.

Keywords: Cloud Computing Platform · Database Language Query · Semantic Level · Potential Association · Dimension Model

1 Introduction

Most of the existing language query methods are non-temporal and are limited to accessing a single database state [1–3]. When measuring the similarity of two short texts or two words, it is impossible to directly edit the distance. Therefore, the database language query method needs to be optimized [4].

Literature [5] proposed a database language query process analysis method based on data mining. It will identify the query target and query conditions through database semantic analysis based on the query statement segmentation array, build a semantic dependency tree, divide it into collection blocks, and convert the natural language query into SQL statements through comprehensive conversion method. On this basis, use data mining technology to locate user demand data in the database, The location data is extracted and fed back to the user to realize the analysis of the database language query

B. Wang et al. (Eds.): ICMTEL 2023, LNICST 534, pp. 273–286, 2024.
https://doi.org/10.1007/978-3-031-50577-5_20

process. The experimental results show that the method can provide accurate database services for users. Literature [6] proposes a database keyword query method based on pattern graph. This method designs a combined network structure, which can effectively avoid redundant operations caused by redundant structures between candidate networks in traditional methods; At the same time, an improved candidate network generation strategy is proposed, which can avoid generating redundant candidate networks and reduce the traversal range, thus improving efficiency; Finally, based on the merging network, a merging network execution algorithm is designed to further improve the query efficiency. The experimental results show that this method can ensure that the query results are not missing.

Although the above methods improve the query effect of database language to a certain extent, there are problems of long query time and high resource utilization. Therefore, this paper proposes a database language query method based on cloud computing platform.

2 Extracting the Semantic Web Hierarchy

The data in the semantic web has a clear definition, and there are associations between the data. Therefore, the semantic web has great advantages in retrieval, automatic processing, reuse, integration, etc. The semantic network connects multiple entities through semantic relationships to form a directional knowledge network from the starting point to the end point. The relationship model uses the relationship table as the data structure, behavior tuple attributes, and field attributes to form a two-dimensional structure. Semantic web can not only describe video, image, digital resources and other content in the page, but also describe abstract resources such as time, behavior, space, and events, and declare the relationship between resources in the description. Semantic web has the following characteristics: the access to information resources is no longer limited to local, and it is more convenient to obtain a wide range of information resources through the network. Although the semantic network knowledge representation is intuitive, when there are many nodes, a strict knowledge network is formed, but as the number of nodes increases, the knowledge network develops horizontally, which is cumbersome and difficult to search and query. The number of groups develops vertically and the cardinality is huge, and even an excellent query algorithm inevitably requires a high time complexity. The Semantic Web adopts a formal description method based on document structure, which can facilitate the integrated processing of data. The information is described in a form that can be accurately understood by the computer, so as to realize the intelligent processing of the information. Convert formally described information into storable data and establish relationships with real-world objects. The basic architecture of the semantic web hierarchy is shown in Fig. 1:

The semantic network knowledge representation itself has the problem of combination explosion, while the chain structure in common data structures can effectively avoid the exponential increase problem caused by queries and searches. Therefore, the semantic relationship of the semantic network can be added to the field attributes of the sub relational model by combining the horizontal and vertical aspects of the semantic network, and then the pointer field can be added to the parent relational model to point to

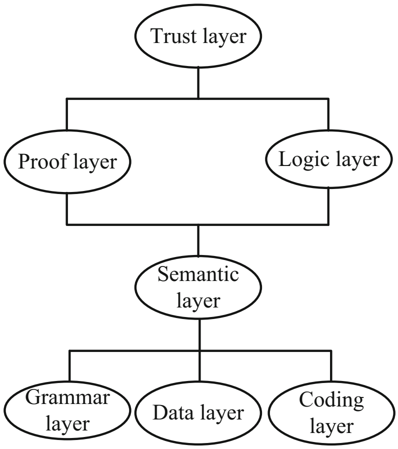

Fig. 1. Basic Architecture of the semantic web hierarchy

the sub relational model, Finally, a semantic network knowledge representation method based on relational model is formed. As the bottom layer of the semantic web, it has two functions. Unicode provides a unified double byte character encoding standard for the semantic web. Each character has a unique identifier corresponding to the Unicode rm code. At present, Unicode can support languages of all known countries. XML Schema mainly describes the structure of the document to limit the semantic representation of the document. Because words are the smallest semantic unit that can move independently. In today's natural language query methods, such as search engines and question answering systems, two kinds of texts are mainly queried, including English text and Chinese character text. As a uniform resource identifier, URI can semantically identify network resources in the Web, thereby providing identifiers for distinguishing these network resources. Syntax layer, namely XML + NS + XMLs layer. XML does not provide semantic interpretation of documents, but only provides syntax for the Semantic Web to achieve document structuring. The data layer is the RDF/RDFs layer. As a key technology in the Semantic Web, the data layer provides a basic description of resources and their relationships. RDF represents resources at the semantic level, which can represent the attributes of resources and the relationship between resources. Therefore, in the database language query method, the difficulty of Chinese word segmentation is higher than that of English word segmentation. Even in some special query methods, it is necessary to combine Chinese text with English text for hybrid query. And the word segmentation algorithm based on human understanding knowledge is usually based on computer pattern recognition, so as to imitate human understanding of special Chinese characters, and finally can recognize the effect of keywords. RDF can define the concepts, attributes and relationships of objects, and use XML as a syntax to provide simple semantic representation functions. It uses URIs to describe the concepts, attributes, relationships, views, etc. of objects, so as to realize the interoperability of applications. However, it cannot deeply describe the meaning of concepts or attributes and the relationships between terms. The value of point type represents a point or undefined value in the European space, and the value of points type represents a set of finite points, so the definition of the abstract type point and the carrying set of points is as follows:

$$\delta_1 = W^2 \cup \{\alpha \perp \beta\} \tag{1}$$

$$\delta_2 = \{L \subset (\alpha, \beta)\} \tag{2}$$

In formulas (1) and (2), W represents real bearing coefficient, L represents finite point set, and α, β represents two adjacent mapping values respectively. In addition, RDFs can define resource attributes, types, elements and concepts, and constrain the combination of resources and concepts according to constraints, so as to ensure that conflicts between constraints are detected. The basic idea of this segmentation algorithm is usually to perform word segmentation, semantic analysis and syntactic analysis at the same time to facilitate the processing of ambiguous words. Due to the particularity and complexity of Chinese text, the theoretical research and technical development of word segmentation algorithm based on natural language is still at the basic stage, but the vast number of scientific researchers have already found its huge commercial value and potential scientific research value. In the hierarchical structure of the Semantic Web, the syntax layer, the data layer and the semantic layer are the core parts of the Semantic Web, which realize the semantic representation of Web resources. Among them, XML realizes the structuring of documents. RDF uses triples (objects, attributes, values) to describe resources. Using this structure is conducive to automatic processing by machines. The basic idea of the word segmentation matching algorithm based on probability statistics is that the probability or frequency of adjacent words can better reflect the credibility of the composed words. RDFs describe the concepts and attributes of resources, and provide semantic relationships between concepts and attributes according to the conceptual hierarchy of ontology. The logic layer, the proof layer and the trust layer realize the logic reasoning function, and provide authentication and trust mechanisms. The digital signature technology is mainly used to trust and authorize the description, reasoning and certification of resources. As the upper structure of the semantic Web, it ensures the security of semantic Web operations and content to the maximum extent.

3 Identify Potential Associated Features of Database Words and Sentences

The database implementation structure adds a spatiotemporal processing layer on top of the traditional relational database management system, and the data operations are completed through the spatiotemporal processing layer without any modification to the underlying database management system kernel. At the same time, this keyword-based text processing method is mainly based on word frequency information, and the similarity of two texts depends on the number of common words they have, so it is impossible to distinguish the semantic ambiguity of natural language. The spatiotemporal processing layer is responsible for the translation between the database language and SQL, and the optimization of spatiotemporal queries. All data requests must be processed through the spatiotemporal layer. The SQL converted from spatiotemporal queries is very complex, which is not conducive to query optimization of the underlying relational database management system. Due to the existence of a large number of synonyms and polysemy in language, the accurate expression of semantics depends not only on the proper use of the vocabulary itself, but also on the definition of the word meaning by the context. If the limitation of context is ignored and only isolated keywords are used to represent the content of the text, the accuracy and integrity of the query results will be affected. In addition, the space-time layer will become a bottleneck for application development, because all requests must first be converted to standard SQL through the space-time processing layer. The space-time operation efficiency of the prototype system of this architecture will mainly depend on the space-time processing layer. Use statistical computing methods to analyze large text sets to extract and represent the semantics of words. Because the context provides a set of interrelationships and constraints on the things in it, it largely determines the semantic relevance between words. A value of type line represents a set of continuous curve curves on the plane, defined as follows:

$$l : [0, 1] \rightarrow W^2 \tag{3}$$

In formula (3), l represents curves mapping space. Because the object relational database management system provides the extension function of user-defined data types, new data types and spatio-temporal operations can be extended on the basis of the object relational database management system, and spatio-temporal indexes can also be extended to the DBMS kernel through UDR and other technologies. Therefore, this potential meaning is the sum of all the contextual information of a word. The starting point of latent semantic analysis is that there is a certain relationship between words in the text, that is, there is a certain latent semantic structure. But it also has some problems. Although the underlying database management can use the extended spatiotemporal index to speed up the query, its support for spatiotemporal query optimization is limited to this. The underlying database management system still uses the query optimization rules of relational databases to process spatiotemporal queries, which is obviously not suitable for spatiotemporal queries. This latent semantic structure is implicit in the contextual usage patterns of words in the text. Therefore, the method of statistical calculation is used to analyze a large number of texts to find this potential semantic structure. It does not need a certain semantic code, but only depends on the relationship between things in the context. It uses semantic structures to express words and texts, so as to eliminate the correlation between words and simplify the text vector. The constant types point, points, lines, and regions in the category SPATIAL are spatial data types provided by the spatial database. The constant type instant in the category TIME is a basic time type, which can be regarded as a real number or a time class with several operations attached. Real numbers are briefly used in the study.If the closure of a set E coincides with itself, then E is a regular closed set. The purpose of this regularization process is then to take into account that the regions should be regular. On the basis of formula (3), the boundary constraints of E are obtained:

$$D = \frac{\phi^2}{\left\{ L \subseteq W^2 | \exists | \phi - 1 | \right\}} \tag{4}$$

In formula (4), ϕ represents interval set. Assume that the implied meaning hidden in words (that is, these semantic dimensions of the potential semantic space) can better depict the true meaning of the text. Different word groups are used in different text sets to express the meaning of some semantic dimensions. The discrete types in the database construct a range that can act on all types in the categories BASE and TIME, thus generating range (int), range (real), range (string), range (bool), and range (instant).The type constructs that moving can act on all types in the categories BASE and SPATIAL, resulting in new types. For example, moving(points) represents the moving point type, and moving(real) represents the real number that changes at any time, which can be used as the return of some space-time operations type. The recognition of this semantic space representation is hindered [5, 8]. Through proper data processing, the purpose of restoring the original orthogonal semantic structure space and the original semantic dimension is achieved. In the language space structure, texts and words are organized and stored according to their semantic relevance, and synonyms scattered in different texts are adjacent in space. Among them, Intime represents the binary ordered pair of time and value, which is defined as:

$$G_{\eta-1} = \frac{\{L \cup W^2\}}{D_\eta} \tag{5}$$

In formula (5), η represents a closed sign quad. In addition, this type constructs that the intime can act on all types in the category BASE and SPATIAL, thus generating new types. For example, the intime (points) type is a binary group, representing the location of the space-time point object at a certain time. The dimension of semantic space is reduced and transformed to eliminate the "noise" in semantic expression (rare or unimportant usage meaning of words). The meaning of words is the weighted average of multiple meanings of words. The potential semantic structure is used to express the term and text, and the term and text are mapped into the same K-dimension semantic space, which is expressed in the form of K factors. Therefore, the meaning of vector reflects not the frequency and distribution of simple terms, but the strengthened semantic relationship.hf is constructed to act on TEMPORAL types, resulting in composite spatiotemporal types that can support past and future motion. For example, hf(mpoints, mhnes) represents a space-time object whose motion before the current moment is a moving point, and is expected to be a moving line afterward. While maintaining most of the original information, it overcomes the phenomenon of polysemy, synonyms and word dependence produced by traditional vector space representation methods. At the same time, the similarity analysis in the new semantic space has a better effect than using the original feature vector, because it is based on the semantic layer and not only the lexical layer.

4 Construction of Ontology Annotation Model Based on Cloud Computing Platform

In the cloud computing platform, one of the most important problems of cloud service ontology annotation is to maintain the consistency between the known cloud service documents and the ontology library that stores annotations. Consistency refers to maintaining the correctness between annotations and language expressions in cloud service documents. Database is the basis for computer to understand natural language, which makes it possible to realize the universality of interfaces. According to the applicable scope of knowledge, this paper divides the database into two parts: general database and special database. The general database is composed of a word segmentation dictionary, a general database dictionary and a synonym word forest, and is not affected by the application field. In a value oriented database, the primary code, that is, the value, is used to represent real world entities.For the representation of entities, a fundamental problem is the update of the master key. Updating the primary key results in a mutation of entity continuity, and care should be taken to update all tuples and sets associated with that entity. Therefore, all changes to the annotated cloud service document (e.g., addition, deletion of cloud service information) must be reflected. Otherwise, if the change information of the source file is not recorded in the ontology library, the cloud service ontology annotation will be inconsistent. The synonym Cilin is a semantic database, which is mainly used for synonym identification, but in order to ensure the accuracy of conversion, this paper retains the thesaurus as a supplement. Object-oriented data models allow the real world to be clearly identified through the use of object identities, which are distinct from values in a database [9]. In a database, each object is assigned a unique object ID by the system, and the relationship between object IDs and objects is fixed. Unlike values, object identifiers can be generated, deleted, but not changed. The special database focuses on the database objects and consists of a special thesaurus, a synonym thesaurus, an entity database, a domain name database, a compound concept database, and an enumeration value database. If cloud service ontology annotation supports complex ontology, it is necessary to maintain the consistency and consistency of ontology annotation between the constantly changing cloud service documents and all ontologies in the process of cloud service ontology annotation. If there are two objects whose property values are the same but whose object IDs are different, the database system also considers them different. The user can send a retrieval request to the cloud service semantic search engine, and the cloud service semantic search engine retrieves the content annotated by the cloud service ontology according to the ontology, and finally responds to the cloud computing-based database language query result in the annotation library through the semantic search engine user. Under the support of the cloud computing platform, the main steps of ontology annotation are optimized and described, as shown in Fig. 2:

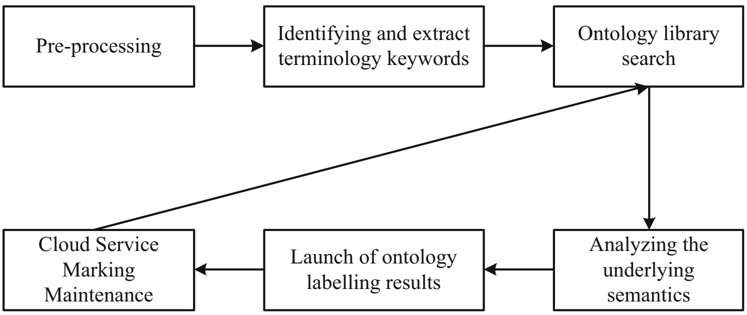

Fig. 2. The main steps of ontology labeling

Since words in language are the smallest unit of natural language processing, it is necessary to preprocess query statements. The main steps of lexical analysis are: word segmentation, part of speech tagging, relational data semantic tagging. Due to the lack of support for natural language interface in existing word segmentation devices, a secondary development is carried out on the basis of existing research results. In the process of word segmentation, the semantics of relational data are taken into account, so that words and database objects can be associated. The key point is the processing between the first part and the second part. The evolution part in the second part is the core. Cloud service, ontology and annotation are combined with the evolution log to evolve the cloud service ontology annotation, and the final result of the evolution of cloud service ontology annotation is returned to the cloud service ontology annotation library in the first part, forming an iterative process between the cloud service ontology annotation library and the ontology annotation evolution. Transplant HanLP's neural network dependency parser, which uses decision-based analysis to obtain word dependencies. The dependency parser believes that whether a sentence is reasonable or not is closely related to its probability of occurrence, and the probability of a sentence is the probability that each word is combined in order, and a sentence is recorded as T, which consists of a series of words, $h_1, h_2, ..., h_n$ in order, then the probability calculation method of a sentence is:

$$P(T) = \begin{cases} P(h_1, h_2, ..., h_n) \\ P(h_1)P(h_2|h_1)...P(h_n|h_1, ...h_{n-1}) \end{cases} \tag{6}$$

In formula (6), refers to the th word. The parser will eventually output a dependency tree that annotates the semantics of related data, parse it by combining rules, and eliminate the ambiguity generated in the process of data semantic annotation according to semantically independent collection blocks, so as to ensure that words have unique semantic information as much as possible. Finally, a consistent cloud service ontology annotation is generated to ensure the correctness of the user's cloud computing based database language query in the first part. Ontology is the fourth layer of the seven tier semantic Web architecture. It is a modeling method for the semantic description of cloud services. Ontology uses a formal way to define the terms that have gained consensus in the field, clarify the relationship between terms, and use the hierarchy and relationship between concepts and attributes to express the semantic relationship between concepts and attributes. Cloud service ontology annotation process is to apply ontology to label cloud service resources based on concepts, attributes Attribute to realize semantic reasoning function. The main task of the structured statement generation module is to map the intermediate structure to SQL, complete the identification and conversion of query targets and query conditions, and generate complete SQL query commands. In order to describe the inconsistency of semantic annotation of cloud services, we define consistency constraints and inconsistency constraints respectively, and describe an annotation model. Among them, the consistency constraint ensures the consistency of the underlying ontology between semantic annotation entities. The semantic annotation constraint is to ensure that the ontology model conforms to the definition of the annotation model. The specific expression formula of the ontology annotation model is:

$$A = \frac{(E, S, D, W, L)^{\mu}}{\sum (D \subseteq T)} \tag{7}$$

In formula (7), S represents the resource set. According to the syntax rules of SQL, this paper classifies and parses the query targets and query conditions, and proposes the corresponding rationality verification scheme, which is more likely to ensure the correctness of the conversion. This module will output a complete SQL query command. The cloud service ontology annotation implementation adds semantic description information to the resources to be tagged based on ontology, making the resources to be tagged from machine readable to machine understandable, thus realizing the semantic Web function. Cloud service ontology annotation can establish a connection between fuzzy natural language and ontology based formal language. This process mainly involves inserting labels into documents. These labels represent the connection between text fragments and ontology elements (attributes, concepts, relationships and instances).If the cloud service document to be annotated is a structured document, the structure of the cloud service document to be annotated needs to be considered, and the annotation is performed according to the structural characteristics of the cloud service document to be annotated. Generally, the structured decomposition method is used for ontology annotation. For example, the structural information in a word document includes: title, hierarchy, style and so on.

5 Design Language Query Methods

The database language query based on cloud computing applies the retrieval conditions described by natural language to analyze and process in a semantic way, realizing the expansion of retrieval at the semantic level, and thus improving the recall and precision of the database language query process based on cloud computing. The process of database analyzing query statements expressed in descriptive language and determining reasonable and effective execution strategies and steps for them is called query optimization. Query optimization is an important part of query processing, especially for relational databases. The essence of query conditions is to limit the given database information, and the purpose of analysis is to express it in a standard structured form, so as to obtain the data records that users want. Generally speaking, in NLQ, in addition to the query target and query verb, it can be considered as a component of the query condition. Database language query based on cloud computing involves two aspects: retrieval purpose representation and semantic expansion constraints. The cloud service retrieval request input by the user usually has a certain retrieval purpose, and the application expression for the cloud service retrieval purpose needs to be processed at the semantic level, so as to ensure the consistency between the semantic expression between the machine understanding and the user's cloud service retrieval requirement. According to the structure type of the heterogeneous database, the number of bidirectional middleware is defined, as follows:

$$\zeta = \frac{\gamma(\gamma - 1)}{2} \tag{8}$$

In Formula (8), γ represents the required cost of heterogeneous databases. In addition, query conditions in the cloud computing platform can be divided into single-layer conditions and nested conditions. Nested conditions refer to conditions that imply sub queries in the condition expression. Single-layer conditions are non nested conditions and no sub conditions are included in the condition expression. Single layer conditions can be divided into simple conditions and compound conditions. There is only one condition expression in the simple conditions, and the compound conditions are connected by multiple simple conditions through logical symbols.

The basic principle of database language query based on cloud computing: the user puts forward a cloud service retrieval request, and the human-machine interface analyzes the lexical rules of the user's cloud service retrieval request, and converts the obtained user search words into concepts through word segmentation. Based on reasoning and expansion at the semantic level, other concept sets related to concepts are extracted, and the concept sets are retrieved and optimized. In SQL, the aggregate function returns the calculated result of the column. Composite query targets have no explicit relational data semantics, and can generally be obtained by combining, transforming, or calculating simple query targets. For example, Age is a composite query target, which is obtained by subtracting the birth year from the current system year. Sort according to semantic distance and semantic similarity, and push the information that meets the user's cloud service retrieval conditions to the retrieval result set after sorting. When the query method uses the head word to predict its context, the network layer that the optimized query target must pass through is described based on the neural network:

$$V = \frac{1}{y}\sum\sum \log \partial\left(\partial_{x,y}\middle|K_{xy}\right) \tag{9}$$

In formula (11), ∂ represents the size of the word bag, y represents the input time, and x represents the unique heat vector of the head word. In imperative sentences, the query target often appears at the end of the sentence as a noun phrase structure. Therefore, when processing query statements, we use reverse order to scan forward from the end of the sentence. The query statement is limited in length, and the rule set is small and unambiguous. In this paper, an improved reverse maximum length matching algorithm is used to store the matched target phrase into the target phrase set, and the entities in the target phrase set are stored into the entity set. The conditional DELETE statement will definitely query, while the non conditional DELETE statement is equivalent to SELECT * FROM TABLE first, and then delete. Therefore, the core of database manipulation statement is query, and the key to improve query efficiency is query optimization. When the cloud service resource base continues to grow, it is necessary to create a semantic index for the cloud service resource base to generate a semantic index base, thereby forming an orderly organization and management of cloud service resources, which is helpful to improve the database language query based on cloud computing.

The specific process of language query is shown in Fig. 3.

6 Simulation Test

6.1 Test Preparation

The specific configuration of the test environment is: CPU: 2*Quad-Core GHz AMD Opteron(tm), Memory: 4GB Fully Buffered Dimm ECC Registered SDRAM, Hard Disk: 2* Seagate (ST9146802SS) 146GB 10000 RPM, Operating System: Linux RedHat,

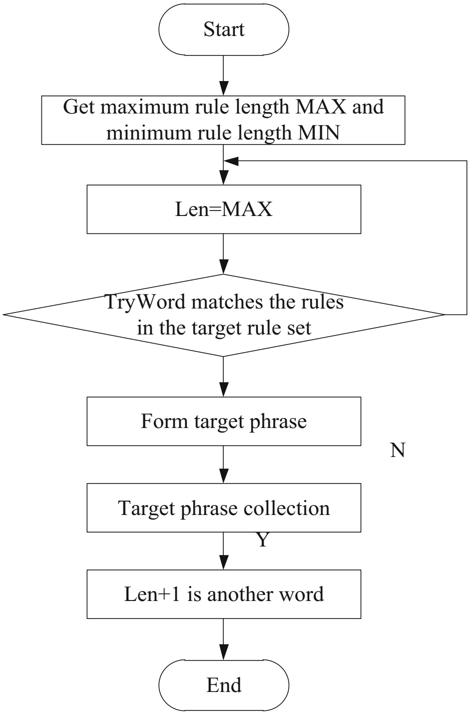

Fig. 3. Language query flow chart

Filesystem: EXT3 page/block size of 4096 bytes, Compiler: GCC. Since the system is a restricted natural language interface, users should support in-scope questions as much as possible.

6.2 Test Results

Select literature [5] method, literature [6] method and the database language query method designed this time to compare the effect. Test the time cost and resource utilization of the three database language query methods under the same number of tuples, as shown in Figs. 4 and 5:

It can be seen from Figs. 4 and 5 that under the condition of the same number of tuples, the average time cost and resource utilization of the database language query method designed this time are: 1720ms and 42.36% respectively; The average time cost and resource utilization of the method in literature [5] are 2340 ms and 59.87% respectively, and the average time cost and resource utilization of the method in literature [6] are 2285 ms and 56.44% respectively.

In order to further verify the effectiveness of the database language query method designed this time, take the query accuracy rate as an experimental index, and compare the methods of literature [5] and literature [6] with the database language query method designed this time. The results are shown in Fig. 6.

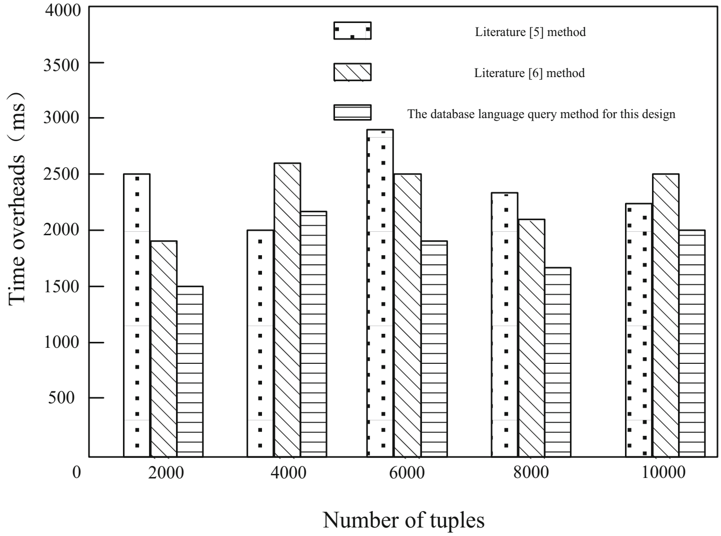

Fig. 4. Time overhead (ms)

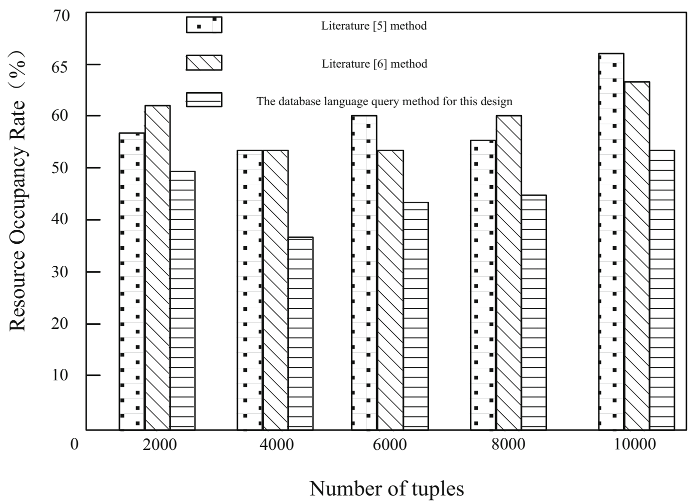

Fig. 5. Resource occupancy rate (%)

It can be seen from Fig. 6 that the maximum query accuracy of the database language query method designed this time is 86.1%, and the maximum query accuracy of the document [5] method and the document [6] method are 73% and 73.5%, respectively. It can be seen that the query effect of this method is better.

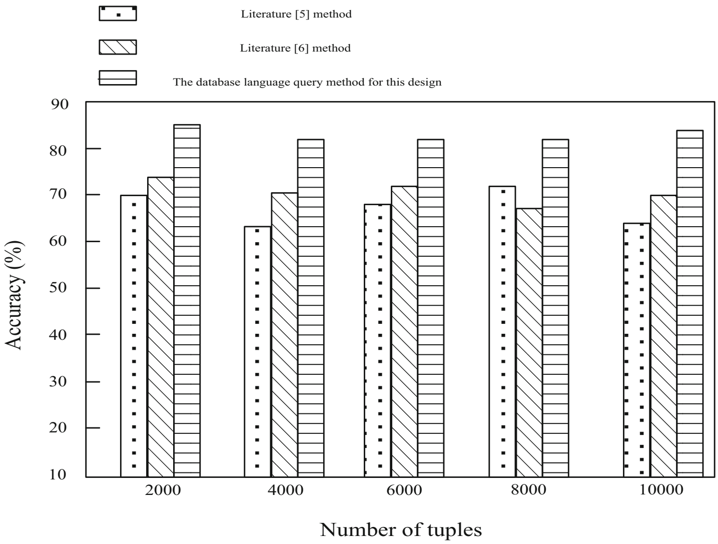

Fig. 6. Query accuracy rate (%)

7 Conclusion

In order to solve the problems of time cost and resource utilization of traditional database language query methods, this paper designs a database language query method based on cloud computing platform. This paper focuses on the semantic annotation method of relational data, and realizes the association between words and database objects in the way of data semantic coverage. In addition, the tagging information of the conditional value of the question is added to the database language query method, and the NL2SQL task of Tab1eQA dataset that may not directly contain the conditional value required by the SQL statement is decomposed into two tasks: question parsing and text matching. In the neural network of question parsing, the separator determined by the column type is fused into the input of RoBERTA encoder to extract the column features. No additional calculation is required by the column attention mechanism. The text matching problem is solved by combining the editing distance and semantic dictionary. The experimental results show that the time cost and average resource utilization of the database language query method designed this time are low, which indicates that the application effect of this method is good.

Aknowledgement. This article is the 2022 Hainan Provincial Philosophy and Social Science Planning Project "Research on the International Language Environment Construction of Hainan Free Trade Port" [HNSK (YB) 22-126] and the 2022 Hainan Provincial Higher Education Scientific Research Project "Hainan Dialect Diversity Protection Research" [Hnkyzc2022-10] Phased research results.

References

1. Yan, K., Chen, H., Fu, D.J., et al.: Bibliometric visualization analysis related to remote sensing cloud computing platforms. J. Remote Sensi. **6**(2), 310–323 (2022)
2. Zhou, X., Zhang, Y., Zhong, Z., et al.: Analysis on the resource management model based on heterogeneous cloud computing platform. Office Autom. **27**(12), 31–33 (2022)
3. Cao, J., Huang, T., Chen, G., et al.: Research on technology of generating multi-table SQL query statement by natural language. J. Front. Comput. Sci. Technol. **14**(7), 1133–1141 (2020)
4. Pan, X., Xu, S., Cai, X., et al.: Survey on deep learning based natural language interface to database. J. Comput. Res. Dev. **58**(9), 1925–1950 (2021)
5. Hou, Y.: Query process analysis of database language based on data mining. Microcomput. Appl. **37**(02), 49–52 (2021)
6. Fei, Y., Ding, G., Teng, Y., et al.: Efficient relational database keyword search method based on schema graph. Appl. Res. Comput. **36**(03), 838–843+860 (2019)
7. Liu, B., Wang, X., Liu, P., et al.: KGDB: knowledge graph database system with unified model and query language. J. Softw. **32**(3), 781–804 (2021)
8. Jiang, S., Cao, L.: Research on the secret homomorphism retrieval method of multiple keywords in privacy database. Comput. Simul. **39**(4), 408–412 (2022)
9. Liu, J., Du, J., Xie, H., et al.: A concurrent OLAP query performance prediction approach for distributed database. Comput. Meas. Control **30**(10), 209–215+221 (2022)