July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

Application of Cloud Computing and Speech Recognition Technology in Computer Teaching

Xu Lili
Zhejiang Institute of Economics and Trade,Hangzhou, China.
E-mail: 37820893@qq.com

Article Info
Volume 83
Page Number: 1365- 1375
Publication Issue:
July-August 2020

Article History
Article Received:*06 June 2020*
Revised: *29 June 2020*
Accepted: *14July 2020*
Publication: *25July 2020*

Abstract

With the development of information technology, the degree of informatization of Computer teaching is gradually increasing. The new teaching technique and infomation segmentation method of cloud computing and skill recognition technology in Computer teaching are the difficult points in skill recognition technology. The key research was carried out in this paper, a kind of new computer segmentation method was proposed; then, combined with the advantages of probability statistics, character string matching and semantic understanding method, a cascaded hierarchical computer word segmentation algorithm (CCWS) was implemented; finally, the best segmentation results were obtained. Experiments show that the recall rate and accuracy rate of this method are 98.59% and 98.89% respectively, which are one or two percentage points higher than the normal method. After applying the algorithm to the cloud computing platform, the speed of computer programming has been greatly improved.

Keywords:*cloud computing; computer programming; computer teaching*

1.INTRODUCTION

The current society has stepped into the era of informatization. Computer is extensively applied in the daily life by relying on its advantages in data processing, however, it can't think like a human being, and this makes it difficult for learners to rely on relative natural methods to pass the messages to computers.And then, this factor becomes a key factor limiting the application of computers in more scenarios . In order to deal with this issue, the man-machine dialogue system is generated. This system uses skill recognition technology and natural language understanding technology to imitate the way of human communication and tries to start natural dialogues with people, so as to achieve the purpose of human-computer natural interaction. However, the relatively big difficulty for dealing with large-scale corpora, low natural language understanding and skill recognition rate and other problems have some restrictions for achieving this goal. In the context of the continuous development of machine learning and statistical analysis techniques, some of these problems begin to be effectively solved. Nowadays, based on the actual situation both in China and abroad, some simplified methods have been appropriately applied to the study of man-machine systems based on the objective of reducing the amount of calculation and improving the accuracy of the system, such as restricting the field of application and limiting the words used in conversation, etc.. Nowadays, the large-scale computing and mass storage have gradually become a factor limiting the human-machine dialogue system to some extent.

Multimedia Computer Assisted Instruction (MCAI) and Multimedia Distance Education (MDE) have become important symbols of modern educational technology. In the process of computer-aided instruction (CAI) activities, the presentation of text, graphics and animation information, accompanied by narration or reading aloud, will make the teaching better, which was also proposed by Treicheer in 1967. The conclusions of research on learning and memory. How to apply the mature speech recognition technology intelligently to the teaching system is one of the topics that current educational technicians and computer-aided teaching researchers should care about.Through speech recognition technology, students can better understand and memorize the knowledge taught by the teacher, and through the feedback of the students, conduct targeted review and guide, and help students to improve more efficiently.

2.CHOICE OF CLOUD COMPUTING AND SPEECH RECOGNITION

In the case of that the use freedom degree and openness of the Internet are getting higher and higher, the amount of data involved becomes very massive, and it is still growing rapidly. In this context, many Internet companies and research departments have taken effective measures to improve the data processing capability as much as possible, so as to get the data information that meets their own needs in the vast Internet information.

The cloud computing developed in recent years belongs to the business model that the distributed computing, parallel computing and grid computing are fully associated, it spreads the extremely massive computing and mass storage to a number of low-cost devices, and then allocates the storage space and scalable computing resources in accordance with the specific needs. In addition to the concept at the commercial level, the technologies used for building large high performance server cluster in the cloud computing have also become the focus of research. For example, on the basis of MapReduce distributed computing model, Google has become a leader in the field of computer by connecting with BigTable distributed storage system, GFS Google file system and so on, furthermore, it has also built the computer centers for more than one million servers in more than 200 regions around the world and has provided the search engine, Google earth and other related applications. However, due to technical secrecy, Google does not carry out the open source processing to the specific implementation details of MapReduce, so it is difficult for people to form a clear understanding to it . In this case, on the premise of that Apache fully draws the three types of technology, an open source project focused on DFS and MapReduce is developed and built. Today, Hadoop has gradually become the standard for implementing the cloud computing applications and research in academia and industry. Hadoop is the basic construction of the distributed system, which is interconnected with a large number of common computers, thereby turning them into clusters with the same processing and storage capabilities of supercomputers, and keeping the transparent development.

Speech recognition is an interdisciplinary subject. Speech recognition is gradually becoming the key technology of human-machine interface in information technology. The combination of speechrecognition technology and speech synthesis technology enables people to get rid of the keyboard and operate through voice commands. The application of voice technology has become a competitive emerging high-tech industry.

3.METHODOLOGY

3.1Cascaded Hierarchical Computer Segmentation (CCWS) Algorithm

Today, many voice segmentation methods and optimization and improvement methods are implemented, and some need to be optimized. How to fully relate the advantages of the current approach and quickly identify the sentence divergence and unregistered words within a given period of time has gradually become the key to subsequent research. There is a simple and fast advantage in the voice matching method, but such methods are limited by the size of the dictionary . For the computer voice segmentation method obtained based on probability statistics, in terms of the sentence structure, words are constructed by all kinds of continuous words. Therefore, in the sentence, if the appearance

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

frequency of the words associated in one place is relatively high, the probability that this part of the words makes up a word is also correspondingly large, so that the co-occurrence probability of a word can be regarded as an index for defining the words. Till now, such approaches have gradually become different types of statistical models. And meanwhile, their advantages are that they can start the word segmentation processing without dictionaries by virtue of statistical information of words. The probability and statistics have a solid theoretical foundation of mathematics, which can automatically eliminate the ambiguity and identify new words according to the statistical information, thereby effectively improving the shortcomings of the method of character string matching word segmentation. However, the method based on statistical word segmentation is also not perfect. There is a problem in the word segmentation process that a word that its occurrence probability is big but actually it isn't a word is taken as a word, and the amount of necessary statistical information is also relatively large.

Combined with the above several main word segmentation methods, a computer words segmentation algorithm of the interrelated cascaded word segmentation (Cascaded computer segmentation, CCWS) is proposed in this paper. Probability statistics, character matching and semantic understanding methods show the advantages of various methods, and improve the accuracy of word segmentation as far as possible. In CCWS, the first is the segmentation of atoms; the second is to gain a relatively small coarse result set containing the optimal result as much as possible on the premise of full segmentation and carry out the unknown word recognition; the third is to tag parts of speech; and finally, the results with maximum probability are selected as the optimal word segmentation results in accordance with the part of speech combination.

3.2 CCWS Algorithm Steps

The process of CCWS algorithm includes atom segmentation, full segmentation and segmentation candidate and so on. Atomic segmentation is the first step of this algorithm, its purpose is to get all the atomic words; then is to carry out the atomic segmentation to get the corresponding minimum independent computer characters; next is to obtain all possible combinations of the above adjacent atomic words, this is called the full segmentation.

The purpose of the segmentation candidate period is to obtain a coarse result set. In addition to covering the optimal result, the set is relatively small. Considering the speed and recall rate, the forward maximum matching, the reverse maximum matching and the N-Gram statistical word segmentation results are regarded as the rough result set. Generally speaking, the optimal result exists in the three types of participles. To some extent, this method can be regarded as a compromise method of single method and total cut method. Combining advantages and making up for deficiencies can not only obtain relatively high recall rate, but also ensure the word segmentation speed.

The first is the maximum matching word segmentation result, the algorithm is: for each word of an input sentence, the first is to start from the first word to find out the longest word appeared in the dictionary behind, and then to repeat this action in the remaining character string, until all the words are found out. The time complexity of this algorithm is O (n2), where n is the length of a sentence. Because the position information of each word has been obtained in the front full segmentation according to the line numbers with the same starting word number of the two-dimensional table, and the order is arranged according to its length, and then it is convenient to find the positive (reverse) longest word. Next, the location information is taken as a word and cut out. This process is repeated in the remaining part, until all the words are cut out. The time complexity of this algorithm is reduced to O (n). After completing the segmentation candidate, in order to ensure the accuracy of the part of speech,

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

the part of speech tagging is carried out, and then the unknown words are identified. Finally, the optimal word segmentation results are obtained.

In the reverse maximum matching segmentation results, the algorithm idea is: for each word of an input sentence, the first is to start from the last word to find out the longest word appeared in the dictionary in the front, then to repeat this action in the remaining character string, until all the words are found out. The idea of this algorithm is almost the same as that of the forward maximum matching method, and here is not necessary to mention again.

For 2-Gram statistical method word segmentation results, the word string $W=(w_1,w_2,...,w_N)$ is a possible word segmentation sequence in the computer string $C=(c_1,c_2,...,c_m)$, $w_i$is a word, P (w_i) represents the probability of the appearance of w_i. According to Bayesian formula, there is:

$$P(W|C)=\frac{P(W,C)}{P(C)}=\frac{P(W)P(C|W)}{P(C)} \qquad (1)$$

in formula (1), P(C) is the probability of computer string; it is a constant, which doesn't need to be considered. The probability of recovering from a word string to a computer string is P(C | W) = 1 (there is only one way). Therefore, the goal of this paper is to determine the P(W) largest segmentation result set, namely:

$$P(W)=P(w_1,w_2,...w_N)\approx P(w_1)\prod_{i=2}^{N}P(w_i|w_1,w_2,...w_{i-1}) \quad (2)$$

High-order models can describe the formulation of a sentence well, but it also has drawbacks. Because the corpus size used is limited, and many words rarely appear in the corpus or even do not appear, there is a result: the higher-order model can only give a reasonable probability of those character strings that are approximate to the training data, which can't obtain a reasonable probability of those character strings that are not approximate to the training data, that is, the so-called data sparse problem. In addition, the high-order model's storage space is large, the amount of calculation is large, which will be difficult to achieve. And in the practical applications, when n = 2 or n = 3, the good results can be achieved. According to make a comprehensive consideration, this paper uses a second-order model, namely, assuming that all words are affected only by the front word,then it is obtained from equation (2):

$$P(W)=P(w_1,w_2,...w_N)=P(w_1)\prod_{i=2}^{N}P(w_i|w_{i-1}) \quad (3)$$

Assuming w_0 is the starting logo word, then $P(w_1)=P(w_1|w_{i-1})$, and then:

$$P(W)=P(w_1)\prod_{i=2}^{N}P(w_i|w_{i-1})=\prod_{i=1}^{N}P(w_i|w_{i-1}) \quad (4)$$

Then, the problem is transformed into the shortest path problem for solving finite acyclic graphs. Here, the classical Dijkstra algorithm is adopted to solve such problems, and the word sequence appearing in the shortest path is taken as the 2-Gram word segmentation results. Finally, the positive maximum matching word segmentation result, the reverse maximum matching segmentation result and the 2-Gram word segmentation result are used as the rough result set.

3.3 MapReduce Process of CCWS Algorithm under Cloud Computing;

On the basis of Hadoop cloud computing platform, the CCWS computer word segmentation algorithm proposed in third chapters is deployed, and the system architecture diagram is shown as follows.

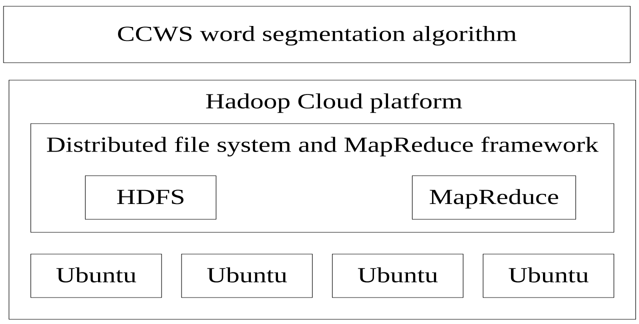

Figure 1. Architecture of computer word segmentation based on cloud computing platform

In the MapReduce process of CCWS, MapReduce tasks are completed by a JobTracker and multiple

TaskTracke such two types of node control. JobTracker's main task is to schedule and manage TaskTracker, which generally runs on NameNode. After JobTracker assigns Map tasks and Reduce tasks to idle TaskTracker, tasks are executed by TaskTracker in parallel. Meantime, JobTracker also oversees the running situation of these tasks. If there is an exception to a TaskTracker, JobTracker will assign this task to another suitable TaskTracker.The word segmentation code is submitted to the JobTracker in the NameNode. JobTracker commands MapReduce works, decides which files should be processed, and divides big files into several blocks according to the actual situation;JobTracker assigns different tasks for the blocks in the first step to deal with the task, at the same time, it also assigns the calculation nodes for each task, TaskTracker is independently responsible for the implementation of the specific tasks; TaskTracker reads the file data from HDFS according to the Map function. In the Map phase, TaskTtracker directly writes the output results to the local disk, and then tells JobTracker the location of these results. Then, JobTracker tells the Reduce task of TaskTracker to take these intermediate results in the corresponding nodes. In the Reduce phase, TaskTracke reads the Data and executes the reduce task according to the location information notified by JobTracke. And finally, TaskTracke writes the result of the Reduce phase to the distributed file system.

3.4Application of Speech Recognition Technology in Computer Teaching

Selecting the recognition unit is the first step in speech recognition research. The speech recognition unit has three types of words , syllables and phonemes, and which one is selected depends on the specific research task. The word unit is widely used in small and medium vocabulary speech recognition systems, but it is not suitable for large vocabulary systems. The reason is that the model library is too large, the training model tasks are heavy, the model matching algorithm is complex, and it is difficult to meet the real-time requirements. The syllable unit is more common in Chinese speech recognition, mainly because Chinese is a monosyllabic structure, while English is a multi-syllable, and although there are about 1,300 syllables in Chinese, if there are no tones, there are about 408 unvoiced syllables. . Therefore, for medium and large vocabulary Chinese speech recognition systems, it is basically feasible to use syllables as the recognition unit. The phoneme unit was previously seen in the study of English speech recognition, but the medium and large vocabulary Chinese speech recognition system is also being used more and more. The reason is that the Chinese syllable consists only of initials (including 22 with zero initials) and finals (28 total), and the acoustic characteristics of the vowels vary greatly.

In practical applications, the initials are often formed according to the different final vowels, which increases the number of models, but improves the distinguishing ability of confusing syllables. Due to the influence of coordinated pronunciation, the phoneme unit is unstable, so how to obtain a stable phoneme unit remains to be studied. A fundamental problem with speech recognition is the rational selection of features. The purpose of feature parameter extraction is to analyze and process the speech signal, remove redundant information unrelated to speech recognition, obtain important information that affects speech recognition, and compress the speech signal. In practical applications, the compression ratio of the speech signal is between 10 and 100. The speech signal contains a large variety of different information, which information is extracted, and which method is used for extraction. It is necessary to comprehensively examine various factors such as cost, performance, response time, and calculation amount.The non-specific person speech recognition system generally focuses on extracting the feature parameters reflecting the semantics and removing the speaker's personal information as much as possible. However, the specific person's speech recognition system hopes to include the speaker's personal information as much as possible while extracting the feature parameters reflecting the semantics.

4.RESULT ANALYSIS AND DISCUSSION

4.1 Experimental Result

This experiment focuses on testing the efficiency of the computer word segmentation algorithm CCWS based on cloud computing proposed in this paper, which achieves the word segmentation accuracy and recall test, and here, the focus is the word segmentation rate. The experiment was implemented on the cloud computing platform based on Hadoop, and the test was carried out under the standard of 2 nodes, 3 nodes and 4 nodes. The specific test data was divided into 5 types, and all of them were continuous computer paragraphs. The data size was 52.3M, 156.9M, 313.8M, 627.6M and 12_5_5.2M respectively. Based on the purpose of removing the influence of accidental factors, and making the result closer to the real value, 4 tests were carried out for each group of test data, and finally, the average value of the 4 operation results was taken as the final result. The following table shows the time of the test data running on the data node. Due to the limitation of space, the word segmentation time results on 3 Data Node and the word segmentation time results on 4 Data Node are not listed.

For the same number of nodes test, as the data size increases, the word segmentation rate also increases, which shows that cloud computing platform has obvious advantages for large-scale data processing. However, when the data size is 52.3M, the word segmentation time does not decrease with the increase of data nodes, but remains basically unchanged. This is because the concept of slicing stored data in HDFS is adopted to divide the large-scale data into different pieces with fixed size and then distribute them to different nodes, and the data in each piece will be stored in the same data node. By default, the size of each slice is 64M. So when the amount of data is 52.3M and less than 64M, Hadoop will only start the MapReduce task in the data node that stores the slice. Therefore, the word segmentation time is basically the same. Similarly, when the amount of data is 156.9M, the word segmentation time of 3 nodes and 4 nodes is basically the same, because 156.9M is divided into 3 fragments.

Table.1 The segmentation time on 2 DataNode

File size (M)	52.3	156.9	313.8	627.6	1255.2
Time 1 (s)	62	121	211	379	675
Time 2 (s)	68	119	229	350	680
Time 3 (s)	64	120	230	346	685
Time 4 (s)	63	124	209	344	682
Mean time (s)	64.25	121.00	219.75	354.75	680.50
Average rate (M/s)	0.81	1.30	1.43	1.77	1.84

For the same size of data, as the number of cloud platform nodes increases, the time spent in word segmentation decreases. Because the cloud computing platform stores the big data in different nodes in a decentralized way, MapReduce algorithm is used for distributed computing in the process of word segmentation, and this is equivalent to that each node only processes its local data, which greatly increases the processing parallelism and improves the system segmentation rate, at the same time, this also reflects the characteristic that the cost of mobile computing is lower than the cost of moving data. In other words, if the operation is performed next to the data, the efficiency is higher. When the data is particularly large, the effect will be more obvious, so that the network congestion can be reduced and the system throughput can be improved. Of course, the word segmentation time does not decrease with the increase of the number of nodes. This is because there are still some synergistic processes among the nodes. The master node distributes the tasks to each data node, and the data

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

node will regularly report its own running status during the processing, finally, the results are told to the main nodes, and a certain amount of time consumption will also generate during this period.

4.2Speech Recognition Test Results and Comparison

After the acquisition, the voice signal number must be pre-processed by filtering, A/D conversion, pre-emphasis and endpoint detection before it can enter the practical applications of recognition, synthesis and enhancement. There are two purposes for filtering: one is to suppress all components in the input signal whose frequency exceeds to prevent aliasing interference; the other is to suppress 50 Hz power line interference. Therefore, the filter should be a bandpass filter. The A/D conversion converts a voice analog signal into a digital signal. In the A/D conversion, the signal is quantized, and the difference between the quantized signal value and the original signal value is a quantization error, which is also called quantization noise. The purpose of the pre-emphasis processing is to raise the high-frequency portion, flatten the spectrum of the signal, and maintain the entire frequency band from low frequency to high frequency. The spectrum can be obtained with the same signal-to-noise ratio for spectrum analysis. Endpoint detection is the determination of the start and end points of a speech from a segment of speech containing speech. Effective endpoint detection not only reduces processing time, but also eliminates noise interference from silent segments. There are currently two main types of methods: time domain feature methods and frequency domain feature methods. The time domain feature method uses the voice volume and the zero-crossing rate for endpoint detection, and the calculation amount is small, but the air-frequency will cause misjudgment, and different volume calculations will also result in different detection results. The frequency domain feature method is to perform speech detection using the variation of the spectrum of the sound and the detection of the entropy, and the calculation amount is large.

In software teaching, if the students cannot perform good exercises in the classroom and cannot solve the problems well, the learning effect must not be ideal. The efficiency of the improved auxiliary industrial design is compared with that before the improvement, and the comparison effect is shown in Figure 2.

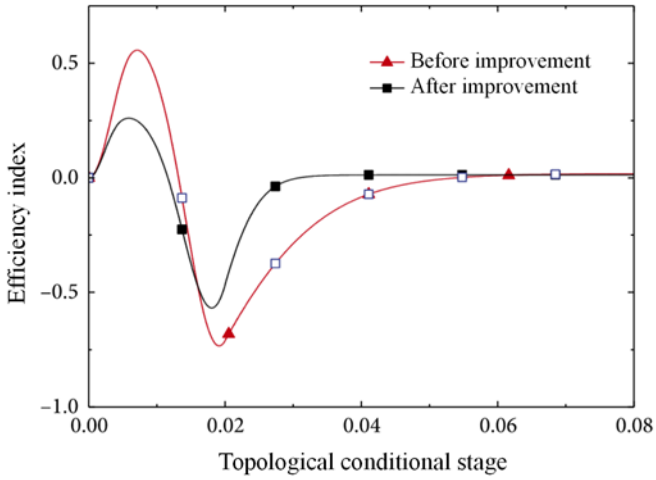

Figure 2. Comparison before and after improvement

In the process of training, we counted the above data information. At present, the mainstream algorithms are BP neural network and J48 decision tree. Among them, the correct rate of training set of BP neural network is 99.0093%, and the correct rate of training set of J48 decision tree is 99.2902%; we can see that the correct rate in training set On the other hand, the accuracy of the mainstream algorithm is very high, reaching the level of two 9s. Although there is a certain gap between the high requirements of the intelligent algorithm 5, 9 for the special industry of speech recognition, more than 9 The accuracy rate is already high enough. The correct rate used in the test set consistent with the working environment is 96.8354% and 95.7278%, respectively, which has far exceeded the accuracy requirement of one 9 in speech recognition. But on the other hand, it is interesting to note that the difference between the two is very large. The former model took 4.51 seconds to build, while the latter reached 0.1 seconds in sub-seconds. Especially in the case that the correct rate of the J48 decision tree is almost the

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

same as the correct rate of the BP neural network, this time-consuming advantage is very obvious.

The frequency at which people speak is below 10 kHz. According to the Shannon sampling theorem, in order to include the information of the desired word in the sampled data of the speech signal, the sampling frequency of the computer should be more than twice the highest speech frequency contained in the speech signal to be recorded. Generally, the signal is divided into several blocks, and each block of the signal is called a frame. In order to ensure that important information that may fall on the edge of the frame is not lost, the frames should be overlapped. For example, when using a sampling rate of 20 kH*, the standard one frame is 10 ms and contains 200 samples. Acoustic waveforms can be acquired by voice input devices such as microphones, as shown in 10ms. Although the waveforms of these sounds contain information about the words required, it is not possible to visually observe the waveforms with the naked eye. Therefore, it is necessary to extract characteristic information from the sampled data that can help identify the words. In speech recognition, linear predictive coding techniques are commonly used to extract speech features. The basic idea of linear predictive coding is that there is a correlation between the sample points of the speech signal, and the current and future sample point values can be predicted by the linear combination of several sample points in the past. The linear prediction coefficient is uniquely determined by minimizing the mean square error between the predicted signal and the actual signal. As a characteristic parameter of speech signal, speech linear prediction coefficient has been widely used in various fields of speech processing.

Vector quantization (VQ) technology is a data compression and coding technique developed in the late 1970s. The vector-quantized feature vector can also be used as an input observation symbol in the late hidden Markov model. In scalar quantization, the entire dynamic range is divided into several cells, and there is a representative value between each cell. For an input scalar signal, the value falling into the cell when quantizing uses this representative value Gotto. Since the semaphore at this time is a one-dimensional scalar, it is called scalar quantization. The concept of vector quantization is to quantify vectors by changing the scalar to a one-dimensional vector from the perspective of a linear space. Like scalar quantization, vector quantization divides the vector space into several small regions. Each small region searches for a representative vector. The vector that falls into the small region during quantization is replaced by this representative vector. The basic principle of vector quantization is to combine several scalar data into a vector (or a feature vector extracted from a frame of speech data) to give overall quantization in a multi-dimensional space, so that the amount of data can be compressed with less loss of information.

4.3 Voice Segmentation Rate Contrast

The acoustic model is usually generated by training the acquired speech features using a training algorithm. The input speech features are matched and compared with the acoustic model at the time of recognition to obtain the best recognition result. The acoustic model is the underlying model of the recognition system and is the most critical part of the speech recognition system. The purpose of the acoustic model is to provide an efficient way to calculate the distance between the feature vector sequence of speech and each pronunciation template. The design of the acoustic model is closely related to the language pronunciation features. Acoustic model unit size and word pronunciation model, semi-syllable model or phoneme model have a great influence on the amount of speech training data, system recognition rate, and flexibility. The size of the recognition unit must be determined according to the characteristics of different languages and the size of the recognition system vocabulary.

The voice segmentation rate of the computer segmentation algorithm CCWS based on cloud computing can reach 3.05M/s (3123KB/s) under 4 nodes. The current segmentation rates of better computer word segmentation algorithm were compared, as shown in the following table:

Table.2 Segmentation speed of different computer word segmentation algorithms

word segmentation algorithm	ICTCLAS 3.0	IK Analyzer 3.2.8	CCWS	CCWS based on Cloud Computing
Word segmentation rate(KB/s)	996	1508	1200	3123

Experimental results environmental description: ICTCLAS3.0 word segmentation rate is provided by the authorities, IK Analyzer3.2.8 and CCWS stand-alone operating environment is the same, and CCWS based on cloud computing runs on four nodes. In the table, ICTCLAS is a computer lexical analysis system developed by the Institute of Computing Technology, computer Academy of Sciences. It is one of the existing computer lexical analysis systems with good overall performance, and the ICTCLAS3.0 stand-alone word segmentation speed is 996KB/s. IK Analyzer is a computer word segmentation toolkit developed by open source Lucene. IK Analyzer 3.2.8 single word segmentation speed reaches 1_508KB/s, and the official data is 1600KB/s. The result shows that the word segmentation rate of computer word segmentation algorithm CCWS based on cloud computing running on 4 nodes is 2.6 times of CCWS, 1.9_5 times of that of IK Analyzer3.2.8 stand-alone and 3.14 times of that of ICTCLAS3.0 stand-alone. It can be seen that the computer word segmentation algorithm based on Hadoop cloud computing can greatly improve the rate of computer word segmentation.The acoustic model is usually generated by training the acquired speech features using a training algorithm. The input speech features are matched and compared with the acoustic model (pattern) at the time of recognition to obtain the best recognition result. The acoustic model is the underlying model of the recognition system and is the most critical part of the speech recognition system. The purpose of the acoustic model is to provide an efficient way to calculate the distance between the feature vector sequence of speech and each pronunciation template. The design of the acoustic model is closely related to the language pronunciation features. The acoustic model unit size (word pronunciation model, semi-syllable model, or phoneme model) has a large impact on the amount of speech training data, system recognition rate, and flexibility. The size of the recognition unit must be determined according to the characteristics of different languages and the size of the recognition system vocabulary.

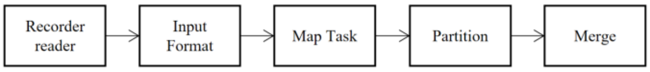

Figure 3. Map phase of Hadoop's data analysis process

According to the above experimental results and the mainstream data processing framework, we can see that the main influencing factors of data processing speed are in the Map process. The Map process is divided into four phases, namely Read phase, Map phase, Collect phase and Combine. stage. First, the RecordReader reads the data into the data analysis process. The InputFormat automatically selects the existing data preprocessing method in the framework according to the data format to clean the original data. Second, MapTask pre-stores the buffers of the data and groups and sorts them according to the corresponding key-value pairs. In the MapReduce calculation, sometimes the final output data needs to be divided into different files. For example, according to the province, the data of the same province needs to be put into one file; if the gender is divided, the data of the same gender is needed. Put it in a file. We know that the final output data comes from the Reducer task. Then, if you want to get multiple files, it means that the same number of Reducer tasks are running. The data of the Reducer task comes from the Mapper task, which means that the Mapper task divides the data and assigns different data to different Reducer tasks. The

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

process of dividing data by the Mapper task is called Partition. The class responsible for implementing the partitioning of data is called a Partitioner.Then, in order to facilitate the final invocation and parsing of the data, the main work of the Collect phase is to partition and sort the data again. Finally, the data about the previous partition and sort completion is merged according to the uniform key-value pair order. In this process, data is always running in memory, affecting speed. In addition to the obvious read and write speed of memory, grouping and sorting of key-value pairs is one of the most important factors.

5.CONCLUSIONS

Building a state network is a word-level network that is developed into a phoneme network and then expanded into a state network. Make this network big enough to include any text path. But the bigger the network, the harder it is to achieve better recognition accuracy. Therefore, according to the needs of the actual task, the network size and structure should be reasonably selected.The speech recognition process is actually searching for an optimal path in the state network. The probability that the voice corresponds to this path is the largest. This is called "decoding". The path search algorithm is a dynamic plan pruning algorithm called Viterbi algorithm for finding the global optimal path. The decoding in speech recognition is to convert the sound signal into text or similar control signal. The root is to find the most likely word sequence problem matching the input speech signal. This is a search process. If the system is based on HMM, the speech recognition system. That is a search in a huge picture, and the search is based on the scoring and language model probabilities of the speech signal in the acoustic model. It is the acoustic and linguistic model mentioned earlier.

To simplify the discussion of these models, they are classified according to the state of the data being processed. Some systems can process data in batch mode, and some systems can stream data that continuously flows into the system. There are also systems that can handle both types of data at the same time. Batch processing is ideal for calculations that require access to a full set of records. For example, when calculating totals and averages, the data set must be treated as a whole, not as a collection of multiple records. These operations require that the data maintain its state as the calculation progresses. Tasks that need to process large amounts of data are usually best suited for processing with batch operations. Whether processing a data set directly from a persistent storage device, or first loading a data set into memory, the batch system takes into account the amount of data in the design process and provides sufficient processing resources. Batch processing is often used to analyze historical data because it performs extremely well in responding to large amounts of persistent data. The processing of large amounts of data takes a lot of time, so batch processing is not suitable for situations where processing time is high. The stream processing system calculates the data that enters the system at any time. This is a completely different approach than the batch mode. The stream processing method does not need to perform operations on the entire data set, but performs operations on each data item transmitted through the system.

We believe this approach is more suitable for our speech recognition and cloud computing in computer teaching applications, from the perspective of the use of the results.In the era of rapid development of information technology, how to obtain information and transmit information are two issues that deserve attention, especially in teaching. At present, many technologies have been introduced into Computerteaching. And how to fully apply these technologies is the key. In this paper, the skill recognition technology in Computer teaching was studied under the background of cloud computing, a new computer word segmentation method was proposed, and a cascaded hierarchical computer word segmentation algorithm (CCWS) was implemented. The idea of this algorithm was to first get the smallest coarse result set that covered the

July-August 2020
ISSN: 0193-4120 Page No. 1365- 1375

optimal result. Then, by stratified filtering, the ambiguities that couldn't be resolved by the upper layers were left to the lower layers, besides, each layer adopted the corresponding strategies according to the specific problems encountered in the word segmentation process, and the optimal word segmentation results were obtained. Finally, the algorithm was verified by experiments. The results show that the recall rate and accuracy rate of this method are 98.59% and 98.89% respectively. After applying this algorithm to the cloud computing platform, the speed of word segmentation has been greatly improved.

REFERENCES

1. Xing D N, Zhao Q L, Song Z G. A Study of Service Model Innovation for Logistics Information Platform Based on Cloud Ecosystem[J]. Journal of Business Economics, 2016.
2. Ding J, Xiong C, Liu H. Construction of a digital learning environment based on cloud computing[J]. British Journal of Educational Technology, 2015, 46(6):1367–1377.
3. Yan H, Hu H Y. Design and Realization of Innovation and Entrepreneurship Service Platform for Undergraduates Based on Big Data[J]. Applied Mechanics & Materials, 2013, 411-414:394-397.
4. Jing N. Research on Mobile Learning Platform Construction in Higher Vocational Colleges Based on Cloud Computing[C]// International Conference on Computational Intelligence & Security. 2015.
5. Wang X Y, Chen J C, Xiao-Yong D U. Survey on OLTP Application Oriented Data Distribution in Cloud Computing[J]. Chinese Journal of Computers, 2016.
6. Ming X, Guo J. Research on Influence of Cloud Environment on Traditional Network Security[C]// 2018:012080.
7. Zhang T, Ying S, Geng J. Database resource integration of shared cloud platform based on RAC architecture[J]. Neural Computing and Applications, 2018:1-12.
8. Zhang W. Research on Data Mining of the Internet of Things Based on Cloud Computing Platform[C]// Iop Conference Series: Earth & Environmental Science. 2018.
9. Yuan C, Peng L, Zhang Y. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing[J]. Future Generation Computer Systems, 2018, 88:279-283.
10. Zhang Wei, Jia Yuhui, Zhang Zhinan. A Customizable Cloud Computing Method for Speech Recognition[J]. Journal of Ocean University of China(Natural Science Edition), 2014, 44(01): 112-117.
11. Xing Y, Chen W. Design of Speech Recognition Robot Based on MCU[J]. 2012.
12. Mushangwe H. Computer-Aided Assessment of Tone Production: A Case of Zimbabwean Students Learning Chinese as a Foreign Language.[J]. Iranian Journal of Language Teaching Research, 2014, 2.
13. Li J, Wu Z, Li R, et al. Multi-modal Multi-scale Speech Expression Evaluation in Computer-Assisted Language Learning[J]. 2018.
14. Kirillov S N, Dmitriev V T. A complex algorithm for objective evaluation of the decoded speech signal quality under the action of acoustic interference[J]. Tr Spiiran, 2018, 1(56):34–55.
15. Yuan C, Peng L, Zhang Y. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing[J]. Future Generation Computer Systems, 2018, 88:279-283.