Hindawi
Scientific Programming
Volume 2022, Article ID 4055809, 13 pages
https://doi.org/10.1155/2022/4055809

*Research Article*

# Machine Translation of Scheduling Joint Optimization Algorithm in Japanese Passive Statistics

**Changsheng Liu**

*School of Foreign Studies, Anhui Sanlian University, Hefei, Anhui 230601, China*

Correspondence should be addressed to Changsheng Liu; liucs907@mail.slu.edu.cn

Received 18 June 2022; Revised 22 August 2022; Accepted 30 August 2022; Published 23 September 2022

Academic Editor: Juan Vicente Capella Hernandez



Machine translation is different from written translation. How to improve the performance of machine translation has been a research hotspot in current research on machine translation. In this paper, based on the semantic analysis and research of Japanese passive, a joint optimization algorithm of scheduling has been proposed, and the machine translation of Japanese passive has been studied. At present, machine translation is more and more widely used. Machine translation has solved many vocabulary problems, and it can complete a large amount of translation work and save a lot of manual translation time. While improving the translation speed, in the process of Japanese passive translation, it is also found that direct machine translation shows many shortcomings, and the quality of passive translation is not particularly ideal, exposing the basic problems of machine translation, such as semantic errors, syntactic errors, unclear and rigid expressions, and messy structures. In response to the problems above, this paper has improved the machine translation model for scheduling joint optimization algorithms. The paper has proposed several optimization algorithms and used resource awareness and computing power scheduling algorithms to conduct experimental analysis of translation performance. Finally, it is found that, among the two scheduling optimization algorithms, the resource-aware scheduling algorithm has better performance. With the same data, the resource-aware scheduling algorithm has saved 15.5% of the time compared with the computing power scheduling algorithm, and the accuracy of Japanese passive translation was 6%, 5%, and 21% higher than the computing power scheduling algorithm under different data volumes. Not only has the time taken been shortened, but the translation accuracy has also been improved.

## 1. Introduction

At present, machine translation is used more and more times in translation work, and it has even become the basic processing method for many translators or those who need to translate materials. Aiming at the problems exposed by Japanese passive dynamics in the process of machine translation, this paper has proposed a statistical machine translation model, constructed a parallel corpus, and then proposed a joint scheduling optimization algorithm to optimize the internal system of machine translation and improve its translation accuracy. The channel model, scheduling model, interference model, and service flow model have been analyzed. Column generation algorithm and maximal clique algorithm, system computing power scheduling algorithm, and resource-aware scheduling algorithm have been proposed to optimize machine translation, and the performance of machine translation has been experimentally tested according to the proposed algorithm model. Experimental results have shown that the resource-aware scheduling algorithm is more efficient in allocating resources within the system. Compared with the computing power scheduling algorithm, in the case of 6G data volume, the translation efficiency of the resource-aware scheduling algorithm was 15.5% higher, and the translation accuracy for passive was 21% higher. Using the column generation algorithm and the maximal clique algorithm to make an experimental analysis on the optimization problem of Japanese passive in the statistical machine model, it is found that the operation time of the maximal clique optimization algorithm was all less than 1, which is much smaller than that of the column generation algorithm. The role of scheduling joint optimization algorithm in network maintenance and operation optimization is becoming more and more obvious, and its optimization role has been applied to various fields. Cai et al. studied a new type of UAV safety communication system. Two drones were used in the system. Among them, one drone communicated with multiple users on the ground via orthogonal time division multiple access, while another drone in the area intercepted eavesdroppers on the ground to protect the communications of the desired users. They developed a new joint optimization algorithm to deal with the transformed problem. To further improve the secrecy rate performance, they also extended the proposed algorithm to the case of multi-UAV jamming. The simulation results showed that the performance of the proposed joint optimization algorithm is significantly better than the traditional algorithm [1]. In order to formulate production scheduling plans and preventive maintenance plans, Fei and Huimin studied the joint optimization problem of these two plans in the intermittent discrete re-entry machining environment. Batch processors can process multiple jobs at the same time, whereas discrete processors can only deal with jobs one after the other, with many re-entrant processes. By measuring the information and adjusting the level of individual variation, the scope of the search space has been effectively narrowed [2]. Zhao and Mili found that industrial CR networks were planned to operate in a wide spectrum, and they had high energy consumption due to frequency switching, while other wireless technologies did not have this problem. A striking feature of this switching cost was that it depended on the width between the two frequency bands. Considering the different energy consumption generated when CR devices switch to different frequency bands, the joint frequency allocation and scheduling problem of multihop industrial CR networks with a single transceiver has been established. Therefore, a polynomial time heuristic algorithm has been proposed to solve the energy consumption problem caused by channel switching. Simulation results showed that the performance of our heuristic algorithm was very close to the results of CPLEX optimization software implementing integer linear programming [3]. Al-Abbasi and Aggarwal considered a representative system architecture of a Content Delivery Network (CDN), given multiple parallel streams/links between each server and edge routers. They determined, for each client request, a subset of servers to stream video, and one parallel stream from each selected server. To achieve this scheduling, a two-phase probabilistic scheduling has been proposed. On top of using playback time, an optimization problem has been formulated to jointly optimize all requested convex combinations of average pause duration and average video quality. Among them, two-stage probabilistic scheduling, video quality selection, bandwidth partitioning between parallel streams, and auxiliary bound parameters can be selected. This nonconvex problem has been solved using an efficient iterative algorithm. Experimental results have showed a significant improvement in QoE metrics for cloud-based videos compared to the considered baselines [4].

With the continuous development of internationalization, language communication is particularly important, and there are more and more translations between different languages. Song said that there were still many shortcomings in the Chinese-Japanese machine translation method, the corpus information processing was not deep enough, and the translation process lacked rich language knowledge support. In particular, the recognition accuracy of Japanese characters was not high. Based on machine learning technology, combined with image feature retrieval technology, a Japanese character recognition model was constructed, which was used as the object of algorithm recognition. Image features were extended by generating a brightness enhancement function using a bilateral grid [5]. In order to solve the problems of small scale, slow speed, and incomplete field of traditional parallel corpus machine translation, Zheng and Zhu constructed a Japanese translation teaching corpus based on a bilingual nonparallel data model and used this corpus to train a Japanese translation teaching machine translation model to obtain more accurate information good auxiliary effect. In the construction process, for nonparallel corpora, they used a translation retrieval framework based on word graph representation to extract parallel sentence pairs from the corpus and then built a translation retrieval model based on bilingual nonparallel data [6]. Mangeot-Nagata has argued that French-Japanese bilingually aligned corpora and machine translation systems were logically equally rare. They have built a high-quality and broad-coverage dictionary available on the web. In order to update these data, whose vocabulary may be very old, existing electronic resources, such as Wikipedia, or Japanese-English electronic resources can be reused, and the resulting resources can be made available online for review and correction by voluntary contributors. First, Japanese bilingual dictionaries (printed or electronic) and their historical evolution were taken stock of. Then, the resource to be built was described. The next part involved the conversion of the three sources, followed by several error corrections for French and Japanese. Finally, the resource was published on a website built around the Jibiki platform, allowing online viewing and editing of articles, and also providing a French-Japanese bilingual corpus and active reading mode [7]. The abovementioned literature has been relatively thorough in research on scheduling joint optimization algorithm and statistical machine translation and has certain guidance for the following research.

This paper mainly studies the related problems of machine translation of scheduling joint optimization algorithm in Japanese passive statistics. The statistical machine translation model is different from the traditional machine translation model. It is based on the corpus, conducts mathematical derivation and analysis of the survey to be translated, and has integrated sentence features to achieve accurate translation. Building a reasonable optimization model is conducive to improving the quality of Japanese passive translation, improving the utilization of internal resource information of the system, and reducing the cost of system maintenance and operation. Using the scheduling joint optimization algorithm to optimize machine translation can effectively reduce translation time, improve translation efficiency, and take into account the quality of translation. The combination of corpus and big data can reduce the complexity of translation work and provide important technical support for machine translation. This paper proposes a statistical machine translation model that is different from traditional machine translation models to enrich the development of linguistics.

The structure of this paper is mainly to briefly introduce the scheduling joint optimization algorithm and, at the same time, briefly introduce the Japanese passive classification and statistical machine translation model. In the specific experimental process, the scheduling algorithm is applied to Japanese passive statistics for simulation experiments. The experimental conclusions of this paper are obtained.

## 2. Machine Translation Method of Scheduling Joint Optimization Algorithm in Japanese Passive Statistics

*2.1. Passive Classification of Japanese.* There are many ways to classify Japanese passives, which can be classified from multiple perspectives. Classification according to the sentence structure of Japanese articles can be divided into direct passive and indirect passive [8]. Based on the emotional tendencies of sentences, the direct passive can be divided into sentient passive and ruthless passive. According to the difference of the subject in the sentence, it can be divided into the passive that expresses people and the passive that expresses objects. According to whether there is a beneficial relationship in the sentence, it can be divided into the passive dynamic of interest and the passive dynamic of non-interested interest. Finally, according to the composition structure of the sentence, it can be divided into direct and indirect passive.

*2.2. Statistical Machine Translation Model.* Statistical machine translation is a type of machine translation, and it is also a method with better performance in machine translation in nonlimited fields. Statistical machine translation is a method based on parallel corpora. It has used mathematical derivation to perform statistical analysis on the content of the corpus, built a statistical translation model, and then used this model for translation [9]. Statistical machine translation methods have attracted more and more attention due to their rigorous mathematical derivation, good model consistency, automatic learning, and strong robustness. The early word-based machine translation has transitioned to phrase-based translation and is incorporating syntactic information to further improve translation accuracy. Figure 1 shows the workflow of the statistical machine translation model.

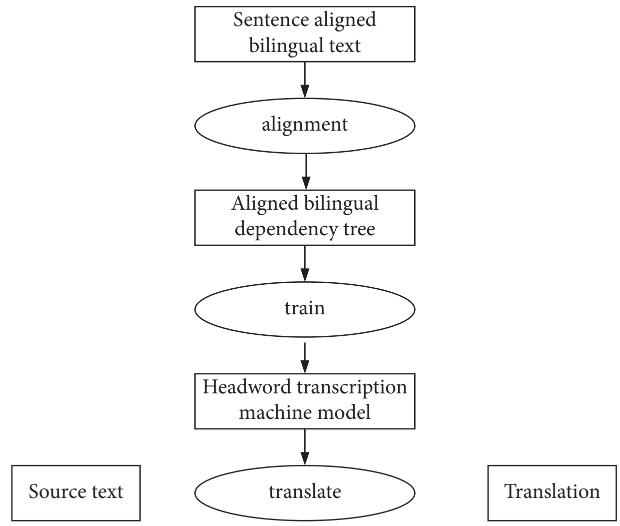

Figure 1: Workflow of a statistical machine translation model.

*2.2.1. Noise Channel Model.* The noise channel model is assumed to be obtained after channel coding the noise generated during the transmission of information in the language. The properties of the source language and the target language can be known through the channel model. According to the calculation rules of the known noise channel, the probability of the target language obtained in the source language is obtained, which is beneficial to translation, and the best translation result can be obtained, which can be obtained according to the channel model:

$$\tilde{r} = \arg\max_{r\in r*} w(r \mid k). \quad (1)$$

Using Bayesian formula, given the value $k$, formula (1) can also be converted to

$$\tilde{r} = \arg\max_{r\in r*} w(r \mid k) = \arg\max_{r\in r*} \frac{w(k \mid r)w(r)}{w(k)} = \arg\max_{r\in r*} w(k \mid r)w(r). \quad (2)$$

The statistical machine translation model consists of two parts, the translation model and the language model of the representation. Among them, the translation model mainly reflects the correspondence between the vocabulary of the source language and the target language, and the language model represents the characteristics of the language itself [10]. The translation model guarantees the meaning of the translation, and the language model guarantees the smoothness of the translation. Machine translation can be regarded as a process of information transmission, and machine translation is interpreted according to the source-channel model. The noise channel model is a suitable method for the vast majority of language models.

According to the translation model proposed by experts, the calculation formula of translation probability is

$$w(k \mid r) = w(k, t \mid r). \quad (3)$$

In the translation model, the relationship between the passive and the sentence can be expressed as

$$w(k \mid r) = \prod_{(i,j)\in t} w\left(k_i, r_j\right). \quad (4)$$

In formula (4), $i,j$ represents a link in the sentence that is dynamically aligned in the word. The translation probability here is between words and not between positions. If a position is added to the entry, the formula is as follows:

$$w(k \mid r) = \prod_{(i.j)\in t} w\left(k_i, r_j\right) w(i, j, K, L). \tag{5}$$

Among them, $K$ and $L$ are the sentence lengths of the source and target languages, respectively, and $i$ and $j$ represent links in the sentence.

In statistical machine translation models, the word alignment with the highest probability can be expressed as

$$\tilde{t} = \arg\max_{t} w_{\beta}(k, t \mid r). \tag{6}$$

A noise model is not suitable for all machine translations. In the noise model, its important role is the word alignment function. In the current machine translation, the parallel corpus word alignment method is mostly used [11]. The more parallel corpora in the system model, the higher the requirements for system translation. However, the accuracy of word alignment is still lacking. The accuracy of Chinese machine translation lags behind other languages, and the error rate of word alignment technology has always been high.

*2.2.2. Discriminant Model.* Unlike noisy channel models, discriminative models are primarily based on data analysis. It does not use Bayesian formulas or assumptions but divides and translates passive sentences or words in the form of feature functions.

The characteristic function in the model framework is expressed as

$$g_n(r, k), n = 1, \cdots, N. \tag{7}$$

The formula obtained after parameterisation is

$$W(r \mid k) = \frac{\exp\left[\sum_{n=1}^{N} \partial_n g_n(r, k)\right]}{\sum_{r*} \exp\left[\sum\limits_{n=1}^{N} \partial_n g_n(r, k)\right]}. \tag{8}$$

According to the weight of the feature function, the estimated value of the model is used as a parameter set, the system is given a source passive sentence, and the decision criterion for the best translation is obtained. The decision criterion for the best translation is to find the solution with the largest characteristic function.

$$\tilde{r} = \arg\max_{r*} \{w(r*, k)\} = \arg\max_{r*} \left\{\sum_{n=1}^{N} \partial_n g_n(r*, k)\right\}. \tag{9}$$

According to the joint optimization method, if the objective evaluation criterion is used as the optimization goal, the result is

$$\tilde{\partial} = \arg\max_{\partial *} \left\{ \sum \int \frac{1}{E_{\int}} \sum_{e=1}^{\int^{E}} F\left( f_{\int, e}, p_{\int} \right) \right\}. \tag{10}$$

Finally, the model is converged and decoded. In the Japanese passive statistical machine translation model, the corresponding passive parameters and the parameters to be translated are trained to form the search results with the maximum probability. If different types of passives are decoded, the optimal path for translation is finally formed [12].

### 2.3. Scheduling Joint Optimization Algorithm

*2.3.1. Channel Model.* The channel model uses a pattern to represent the network topology. The network topology consists of router rendezvous points, collections with gateway functions, and collections of node wireless links. Within transmission range, the wireless link can receive the node's transmission, which indicates that the link is valid. Figure 2 is a schematic diagram of the channel model.

Assuming that the transmission range between nodes in the model is the same, there are both wireless links and reverse links between the two nodes in the link during the transmission process, and the channels used are mainly single-channel and single-transceiver. To simplify the calculation, assuming that all routing nodes use the same power spectral density $Q$ to send signals, the power propagation gain model is

$$Y_{i,j} = M \cdot \log_2\left(1 + \frac{R_{i,j} \cdot G}{\lambda}\right). \tag{11}$$

$M$ is the available bandwidth of channel $Y$ and $\lambda$ is the noise power spectral density.

*2.3.2. Interference Model.* Interference between wireless links can be divided into primary interference and secondary interference. For any two different edges $(i, j)$ and $(k, l)$ in $E$, if there are $i \neq k$, $i \neq l$, $j \neq k$, $j \neq l$, then the fundamental interference does not exist in $(i, j)$ and $(k, l)$ do not exist; that is, the fundamental interference limits multiple transmissions, multiple receptions, and simultaneous transmission and reception. The two interference models are introduced separately as follows [13].

For secondary interference, it is usually described by protocol interference model or physical interference model. At some time $t$, receiving node $j$ treats all other sending nodes except sending node $i$ as interfering nodes. The signal-to-interference-noise ratio at the receiving node is calculated by the interfering node and compared with the threshold to judge whether it can be successfully received. In the physical interference model, the minimum signal-to-interference-to-noise ratio that ensures correct reception by the receiver needs to be defined first as the decision threshold.

In the protocol interference model, the interference between nodes or links is defined as a state parameter with a value of 0/1. Each node has the same transmit power, so all nodes have the same maximum transmission distance. For the interference distance, $\theta$ is the interference factor, usually $1 \leq \theta \leq 3$. All nodes within the maximum transmission distance can communicate directly, and all sending nodes

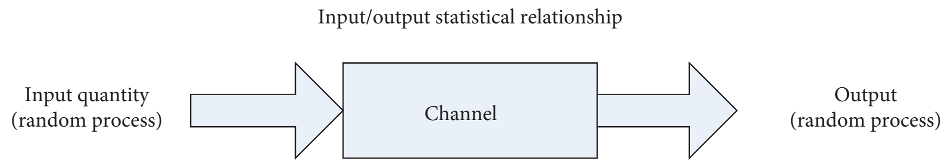


Figure 2: Schematic diagram of the channel model.

within the interference distance will interfere with the node's reception. Under the protocol interference model, two links, $(i, j)$ and $(k, l)$, interfere with each other when one of the following two situations occurs:

(1) $(i, j) \cap (k, l) \neq \varphi$; that is, the two wireless links have a common node.

(2) $j \leq Ir$ or $dk, j \leq Ir$; that is, the interference range of the sending node of one link includes the receiving node of the other link.

If two links interfere with each other, they cannot be simultaneously active for data transmission. Figure 3 can be used to illustrate the limitations of mutual interference between links under the protocol interference model. If node 1 transmits data to node 2, other nodes that would interfere with node 2 cannot send data. Because node 2 is within the interference range of node 3 and node 5, respectively, neither node 3 nor node 5 can transmit data. Consider another scenario, if node 3 sends data to node 4, because node 6 is not within the interference range of node 3, node 5 can use channel 1 to send data to node 6 at the same time. In wireless networks, when the path fading index $\beta$ is greater than 2, the protocol interference model can be basically equivalent to the physical interference model.

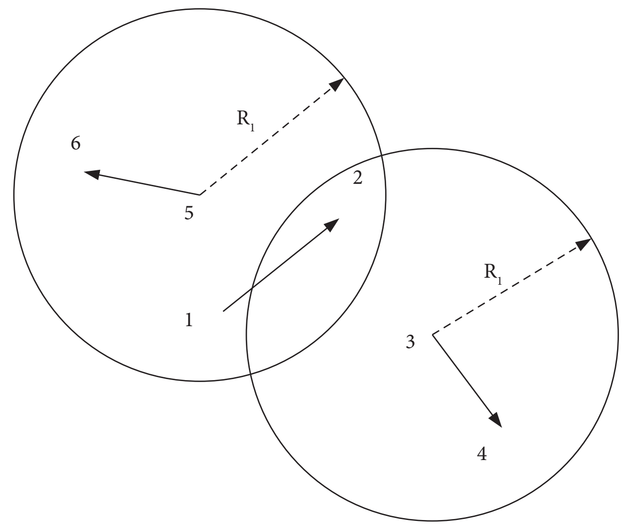


Figure 3: Example of interlink interference.

*2.3.3. Scheduling Model.* Any two links can be activated to transmit data at the same time only if they do not interfere with each other; otherwise, the corresponding data transmission will fail. The set of links that can be activated simultaneously at any given time is called Concurrent Transmission mode (CTS). A concurrent transfer mode, in which no further concurrent links can be added, is called a maximally concurrent set. The scheduling model diagram is shown in Figure 4.

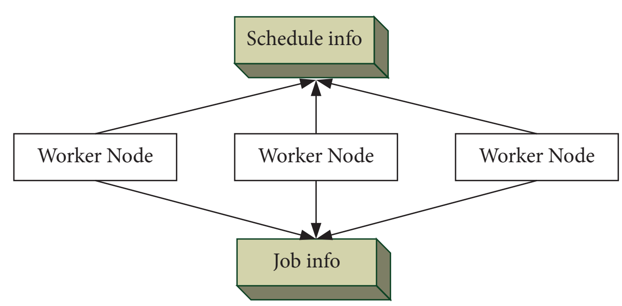


Figure 4: Decentralized scheduling model.

According to the transmission modes described above, all concurrent modes are set as a set, which grows exponentially with the number of links [14]. In the model, variables were introduced and disabled to reduce the calculation error of the model, and whether the link in the model China is activated during the transmission process was observed. If the aggregate link is 1 during transmission, it means that the link is active, and if it is 0, it means it is not active. If two links interfere with each other but cannot be activated at the same time, where $K$ is used to represent the interfering link set, then the result is

$$m_{i,j}^{n} + m_{k,\lambda}^{n} \leq 1\Lambda(i, j) \in E, \Lambda(k, \lambda) \in Y_{(i,j)}. \quad (12)$$

It is assumed that, in a scheduling period, the system allocates a dedicated time segment $Ts$ for each different transmission mode, and each transmission mode can only transmit data in its own dedicated time segment [15]. The length of the time slice may be 0. If, in a certain scheduling period, the time segment allocated by the system to a certain transmission mode is 0, it means that all links corresponding to this transmission mode cannot be activated to transmit data in this scheduling period.

The activation test is performed on the abovementioned traditional transmission mode, and the intrascheduling period of the link is adjusted to obtain the total activation time. The formula is

$$\theta_{i,j} = \sum_{n\in N} W_n \cdot m_{i,j}^{n}, \Lambda(i, j) \in E. \quad (13)$$

*2.3.4. Business Flow Model.* To optimize the throughput of data flowing from the router node to the gateway node, the optimization problem of the downstream data flow can be obtained by simply modifying the algorithm. Furthermore, although in practical systems, the transmission interval is usually divided into time slots, and information is transmitted in the form of data packets, if the duration of the time slot is small enough, and a transmission interval can contain multiple time slots, the data transmitted in the form of data packets can be regarded as a bit stream [16].

tThe multisession (Multisession) business flow model was used to describe the multipath routing of business data flow through the network. The business data flow sent from the source node to the gateway node was called a session (session) and is identified by its source node. In a multihop network, the source node often needs to pass through multiple intermediate nodes to transmit data to the destination node [17]. In addition, if each session adopts a single path, it will not be able to take advantage of the load balancing advantage of the multihop network, so we assume that the data traffic flow of the same session can reach the destination node through multipath routing. The data flow of the data flow sent by the source node $c$ to the gateway node is the data flow of the link $(i, j)$, where $U_c$ represents the bandwidth requirement sent by the node $c$ to the gateway node.

(1) The amount of data flowing from source node $c$ to the gateway and the amount of data of service flows related to $c$, the sum of which must be equal to the total bandwidth requirement $U_c$ of node $c$:

$$\sum_{t\in Q,(c,t)\in E} k_{v,t}^{c} = U_c, \Lambda_c \in V_r. \tag{14}$$

(2) To avoid loops, any traffic flow on the link to source node $c$ related to $c$ must be equal to 0:

$$k_{i,v}^{c} = 0, \Lambda(i,c) \in E, \Lambda_c \in U_r. \tag{15}$$

(3) At the gateway node, since the gateway node is connected to the wired network and does not provide relays for the flow of other nodes, all traffic flows out of the gateway node must be equal to 0:

$$k_{i,j}^{c} = 0, \Lambda(i,j) \in E, \Lambda i \in U_g. \tag{16}$$

(4) At the relay node, the traffic flow of all nonlocal nodes must satisfy the flow conservation law; that is, the total amount of business flows flowing into the other node of the relay node must be equal to the total amount of business flows flowing out of the other node of the relay node:

$$\sum_{i\in U_r,(i,j)\in E} k_{i,j}^{c} - \sum_{h\in U_r\,(j,h)\in F} k_{j,h}^{c} = 0. \tag{17}$$

(5) The sum of all traffic flows through a link $(i, j)$ cannot exceed the physical capacity of the link multiplied by the total activation time of the link in the scheduling period:

$$\sum_{c\in U_r} k_{i,j}^{c} \le \beta_{i,j} \cdot A_{i,j}, \Lambda(i,j) \in E. \tag{18}$$

The total throughput of the system is defined as the ratio of the total uplink data of all router nodes in a certain time interval (e.g., a scheduling period) to the total transmission time, such as system activation time, allocated to each relevant node for transmitting these data [18]. Obviously, given the uplink traffic (i.e., bandwidth requirements) to each router node, if the system activation time is shorter, the total system throughput will be larger; that is, the system performance will be better. Therefore, the optimization goal is to minimize the system activation time under the premise of meeting the bandwidth requirements of all nodes in the network. Therefore, the optimization problem can be formulated as

$$\min \sum_{c\in C} R_c. \tag{19}$$

*2.3.5. Column Generation Algorithm.* The column generation algorithm is a decomposition technique used to solve large linear programming and integer linear programming problems with many variables [19]. The column generation algorithm decomposes a linear programming problem into a restricted master problem and a slave problem (the pricing problem). The column here refers to the column vector of the constraint matrix of the linear programming problem. Initially, the master problem contains only a subset of the column space, and then the slave problems are solved to determine whether the master problem can be expanded by adding columns to obtain a better solution and by iterating between master and slave problems until the master problem contains all the columns for which the optimal solution is obtained. A brief flowchart of the column generation algorithm is shown in Figure 5.

For the linear optimization problem studied in this chapter, a column in the column generation algorithm corresponds to a certain concurrent transmission mode $s$, so the restricted main optimization problem is formulated as follows:

$$\min \sum_{c\in C_0} R_c. \tag{20}$$

Here the concurrent transmission mode set $S0$ is a subset of the set $S$. In order to simplify the initialization of the concurrent transmission mode set $S0$, it can be selected to include $L$ (number of links) concurrent transmission modes at the beginning, and each concurrent transmission mode only contains a corresponding link. In this way, by solving the restricted main optimization problem, the optimal solution corresponding to the concurrent transmission set $S0$ and the dual variable $P_{i,\,j}$ of the main optimization problem can be obtained. Then, it needs to be determined whether the main optimization problem can be reoptimized by adding a new concurrency mode to $S0$. This is equivalent to a concurrent transmission mode $s$, whose corresponding activation time $Ts$ has a negative cost reduction for the existing concurrent transmission mode set. Let RCs be the reduced cost of concurrent transmission mode $s$. According to the duality property of linear programming, there is

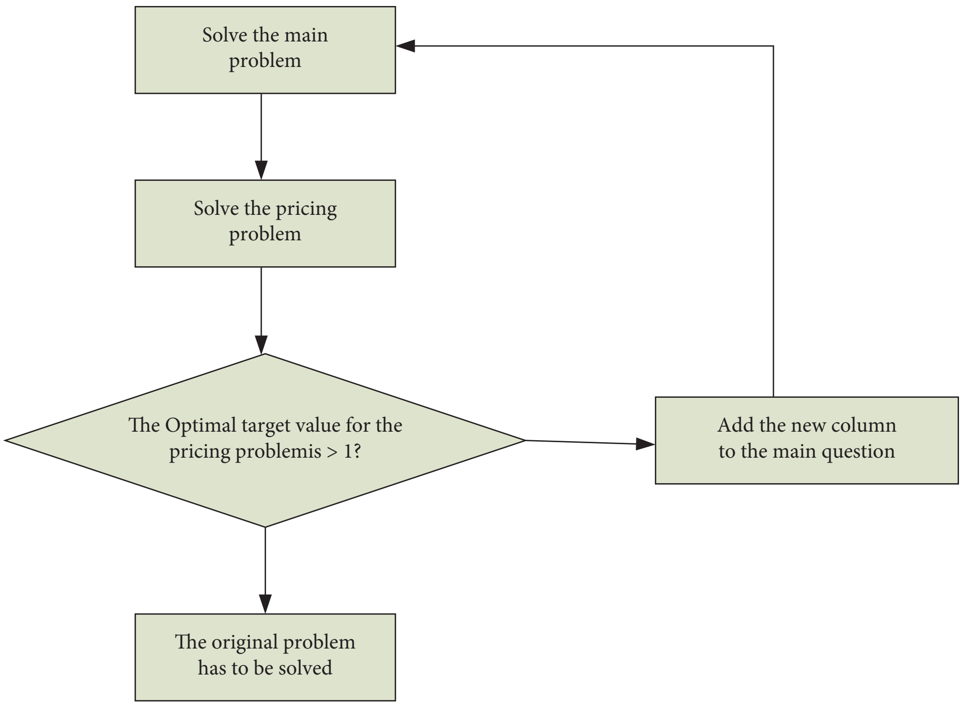


FIGURE 5: Flowchart of the column generation algorithm.

$$WH_c = 1 - \sum_{(i,j)\in E} Q_{i,j} k_{i,j}^c H_{i,j}. \tag{21}$$

Obviously, to make the cost value smaller, the maximum value is required, and as long as the value is negative, it means that the original optimization problem can be further optimized after adding the corresponding concurrent transmission mode. Thus, the optimization problem (pricing problem) can be expressed as

$$\max \sum_{(i,j)\in E} Q_{i,j} k_{i,j}^c H_{i,j}. \tag{22}$$

*2.3.6. Maximal Clique Optimization Algorithm.* The optimal result can be obtained by using the column generation algorithm, but when it is applied to this optimization problem, the optimization solution of the pricing problem is a 0-1 integer programming, which is still a NP-hard problem. Therefore, the operation time will be too long, when the network size is slightly enlarged [20]. In graph theory, a clique refers to a subset of the vertex set of a graph so that its derived subgraphs are all complete graphs. If a clique is not a subset of any other clique, the clique is called maximal group. The clique with the largest number of vertices in a graph is called the maximal clique of the graph. According to the problem, an optimization algorithm based on maximal clique search has been proposed, which has greatly improved the efficiency of translation optimization. In vectorless graphs, the maximal clique optimization algorithm is widely used and is a research hotspot in search algorithms. If the data in the process of statistical translation is too complicated, it can be processed by dividing and conquering and parallel computing methods according to the characteristics of different types of data information of Japanese passive.

Optimization studies were performed on statistical machine translation models. The details are as follows. According to the previous optimization conditions, combined with the wireless link model, the wireless link interference graph was constructed. The complementary graph was obtained from the link interference graph, which is an undirected graph, and then the improved BK algorithm was used to search for the complementary graph. The maximal clique of the graph, and all maximally concurrent transfer modes in the system were updated. The transfer mode was substituted into the optimization problem of the statistical machine translation model to solve. It should be pointed out here that the subset of any maximal transmission mode is not listed, and the activation of a subset can be replaced by the activation of a maximal transmission mode that includes the subset, which will not affect the optimized results. Every nonmaximal concurrent transfer mode must be a subset of some maximal concurrent transfer mode. For the constraint expression in the original problem, using a maximal concurrent mode to replace one of its subconcurrent modes means that the constraints are satisfied in the original linear programming problem and have a larger feasible region, so adding any subconcurrent transmission mode will not increase the feasible region of the original optimization problem. Therefore, the full enumeration of all concurrent modes can be replaced by all the maximum concurrent modes without reducing the feasible region of the optimization problem. In this way, the variables of 0-1 integers in the original optimization problem can be eliminated. The mixed-integer linear programming of the original optimization problem is converted into a general linear programming problem to solve so that the number of variables in the optimization model is greatly reduced, and the optimization operation efficiency is greatly improved.

*2.4. System Computing Capacity Scheduling Algorithm (CS).* Capacity Scheduler is a scheduling algorithm based on multiuser environment developed by Yahoo. Its general design idea is to divide multiple queues in the entire cloud computing system, and these queues are also independent of each other [6]. Compared with the FIFO job scheduling method, the computing power scheduling algorithm has made up for the shortcomings of its low resource utilization. According to the nature of each queue, system resources are allocated to each queue with a certain strategy, and the upper and lower limits of resource allocation can be intelligently set according to the classification and nature of the queue. For the usage of resources in the queue, the scheduling algorithm can set the usage by itself. This is to prevent system resources from being occupied for a long time and system resource utilization being affected. In the process of machine translation in Japanese passive statistics, the computing power scheduling algorithm was used. When the computing nodes of various language types are idle, the system always allocates computing resources to the queue with the lowest proportion of resources first, which can reduce the number of processes starved. In a queue, the scheduler always schedules jobs in the queue according to the first-in, first-out policy. However, it is also possible to set priorities for jobs in the queue, and schedule them according to job priorities. At the same time, in the computing power scheduling algorithm, the resources of the system are flexible; that is, when the resources of a certain queue have remaining after meeting its own needs, these remaining resources can be allocated to other queues for use.

The characteristics of the computing power scheduling algorithm are as follows:

(1) Capacity guarantee: The scheduler will set the minimum lower limit and the maximum upper limit of the resource ratio for each queue according to the job characteristics of each queue so as to ensure the normal execution of the jobs of each queue.

(2) Flexibility: The resources of the system are flexible. When resources in the system need to be moved, these resources will be returned to the queue from the queue that borrows resources to ensure the normal execution of tasks in the queue.

(3) Support priority: The scheduler schedules jobs in the queue according to the time-critical order of tasks in the system and executes tasks with earlier submission times first. Statistical machine translation methods have attracted more and more attention due to their rigorous mathematical derivation, good model consistency, automatic learning, and strong robustness.

(4) Multiple tenancy: The queues on the entire system and the jobs in each queue run independently in parallel, and the system can set certain constraints. The purpose of this is to prevent a job, a user, or a queue from occupying system resources for a long time, resulting in slower system response and unbalanced and wasteful resources.

(5) Support resource-intensive jobs: The system's resource requirements for jobs are incrementally reserved. When a task requires more resources, the system will meet the needs of the task by incremental reservation according to the task requirements. By limiting the queue and user resources in the system, the computing power scheduling algorithm logically divides the entire system cluster into several subclusters with relatively independent resources. These subclusters actually share the resources in the larger cluster, so they can share resources, improve resource utilization, and reduce operation and maintenance costs. However, currently the only resource type supported is memory.

All in all, the computing power scheduling algorithm can make up for the shortcomings of low computer utilization and divide different types of passive Japanese into subclusters of independent resources by restricting data queues and user resources in the system. By utilizing resource sharing in large clusters, resource utilization has been improved and operation and maintenance costs have been reduced.

*2.5. System Resource Aware Scheduling Algorithm.* Although the computing power scheduling algorithm has the advantages of reducing the low utilization rate of the system and allocating resources to the system queue, it ignores the competition between different resources [21]. Due to the continuous increase of the processing power of the CPU and the disk, the processing power of the node and the disk access speed show a great difference. If the processing power in the system is reduced, it will greatly slow down the system's machine translation work for Japanese passive statistics and waste system resources [22]. Figure 6 shows the algorithm flowchart.

Some experts have proposed a resource-aware scheduling algorithm for processing Japanese passive statistics of machine translation and improving work efficiency. This algorithm is mainly used to solve the resource competition relationship between different types of dynamic Japanese translations in the system in a heterogeneous environment [23]. The main step of the algorithm is to schedule two different types of passive dynamics[24]. The node load information is sent to the system, and the system determines the classification after analysis and then selects the appropriate task for scheduling. This can not only shorten the response time of the system, but also improve the utilization of resources [25, 26].

## 3. Experiment and Destruction of Machine Translation of Scheduling Joint Optimization Algorithm in Japanese Passive Statistics

*3.1. Experimental Destruction of Scheduling Algorithms.* According to the actual situation of Japanese passive type, this paper has analyzed the statistical machine translation model for computing power and resource-aware scheduling

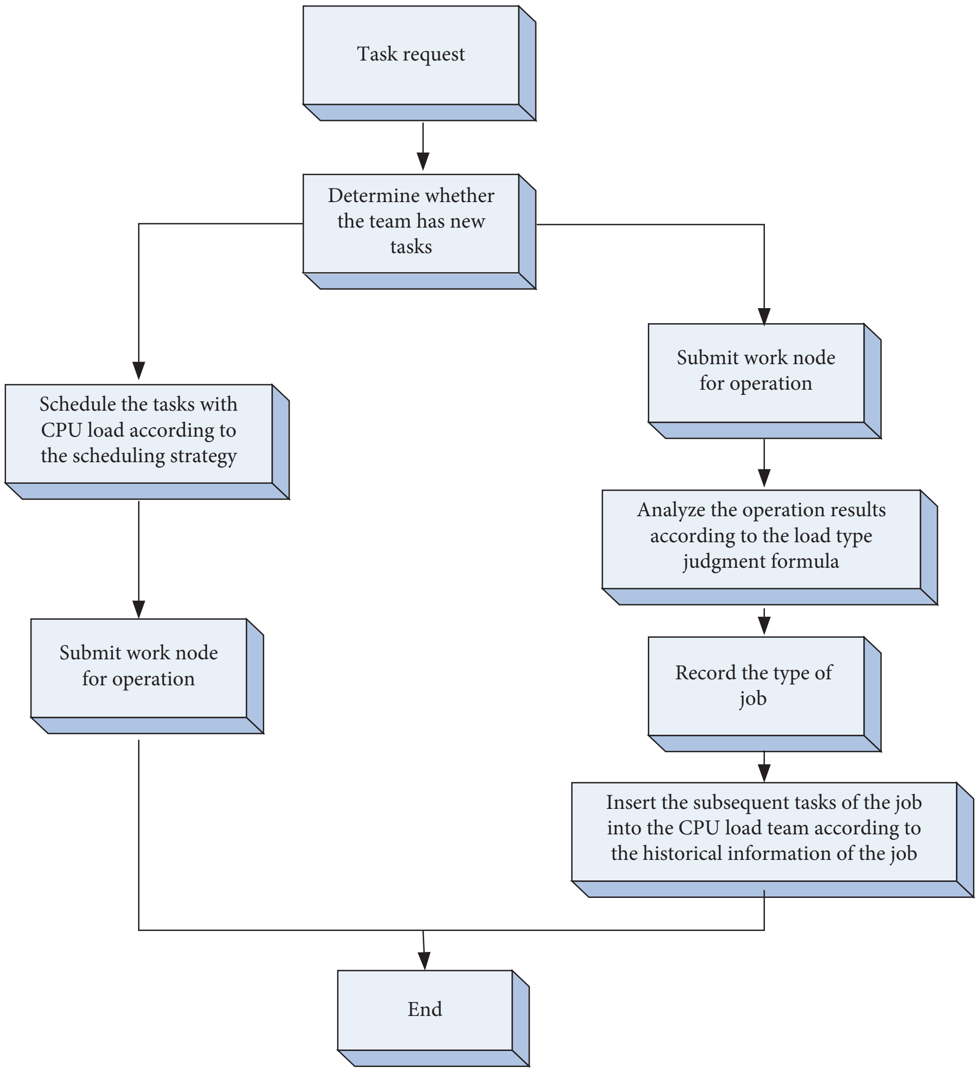


Figure 6: Scheduling algorithm flow.

algorithm experiments. The two scheduling algorithms were placed in a resource pool at the same time, the size and time of the output data were compared, and the performance of the two optimal scheduling algorithms was analyzed [27]. The computer configuration required in the experiment is shown in Table 1.

As shown in Figure 7, when all queues of the computing power scheduling algorithm (CS) share a resource pool, the performance of the computing power scheduling algorithm (CS) is not as good as that of the resource-aware scheduling algorithm. When the input data volume is 6 GB, the calculation time of the two algorithms differs by about 18 s. Compared with the computing power scheduling algorithm, the time of the resource-aware scheduling algorithm is reduced by 15.5%. Although the computing power scheduling algorithm has multiple work queues, it is still a shared resource pool in terms of system resources. Therefore, each queue does not have a separate resource pool, which will cause queue resource competition and reduce scheduling efficiency. As the amount of Japanese passive data increases, the efficiency of the workload-aware scheduling algorithm increases. This is because with the increase of the number of jobs, the total job processing time ratio of job type judgment and job insertion operation continues to decrease, and jobs of different load types make the operating load of the system reach maximum. The resource utilization of the system is greatly improved, thereby reducing the running time of the job.

In the case of the same amount of data, the time spent processing the data has been reduced, which means that the efficiency with which Japanese is dynamically translated has been improved [28]. When the total amount of data is the same, the comparison results of the accuracy of Japanese dynamic machine translation after the two scheduling algorithms are shown in Figure 8.

Among them, the total amount of data given by the middle system of the two optimization algorithms each accounts for 50% of the total amount of data. In order to ensure the authenticity and accuracy of the experimental data, multiple experiments were carried out on the experimental objects of the lock in the same environment, and the final average of the completion time was obtained.

Table 1: Cluster configuration.

| Node type | CPU configuration | Memory configuration (G) | Disk read and write rate (M/s) |
|---|---|---|---|
| Master | 2-Core i3 2.3 G | 6 | 190 |
| Slave 1 | 4-Core i5 3.1 G | 8 | 190 |
| Slave 2 | 4-Core i5 3.1 G | 8 | 190 |
| Slave 3 | 4-Core i5 3.1 G | 4 | 190 |
| Slave 4 | 2-Core i5 3.1 G | 4 | 190 |

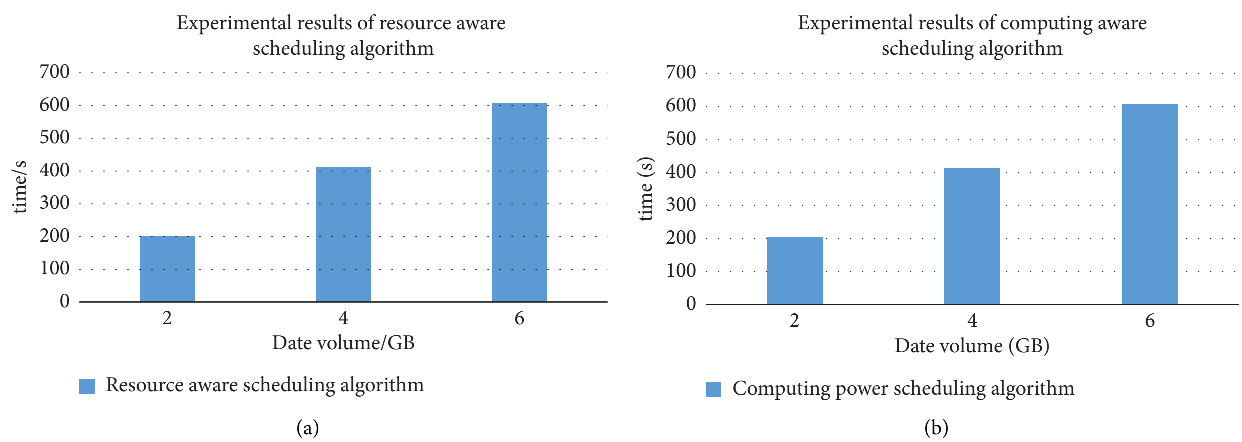


Figure 7: Comparison of experimental results between resource perception and computing power scheduling algorithms.

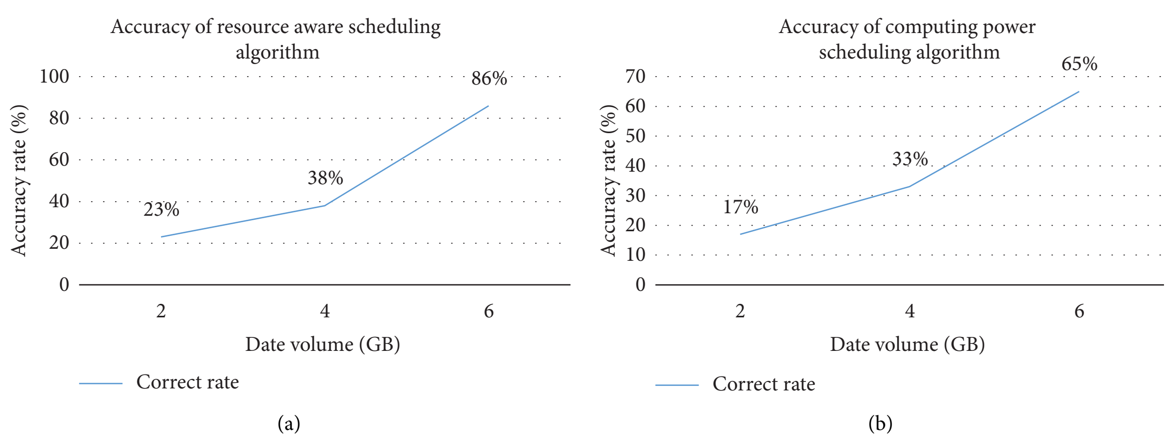


Figure 8: Comparison of the accuracy of Japanese dynamic machine translation under the optimization of two algorithms.

The experimental results have shown that the resource-aware scheduling algorithm of the computing power scheduling algorithm was more accurate than the computing power scheduling algorithm in the Japanese passive machine translation. When the total amount of data is 2 G, 4 G, and 6 G, the accuracy rates of the resource-aware scheduling algorithm are 23%, 38%, and 86%, respectively. Under the premise of the same amount of data, it is 6%, 5%, and 21% higher than the computing power scheduling algorithm, respectively. It can be seen that the optimization ability of the resource-aware scheduling algorithm is relatively good, so it is suitable for the optimization calculation of machine translation in Japanese passive statistics. In addition, it also includes gradient descent algorithm, Newton's method, and other algorithms with optimization properties, which has certain significance for the analysis of the experimental object.

*3.2. Machine Translation Simulation Experiment of Scheduling Joint Optimization Algorithm in Japanese Passive Statistics.* According to the passive characteristics of Japanese and the exponential increase in the number of decision variables in the optimization problem in the statistical machine translation model, when the optimization problem reaches a certain scale, it is quite difficult to find the optimal result due to the limitation of calculation conditions [29]. This also accords with the characteristic that this

Table 2: Experimental results of optimization of 10 network nodes in a statistical machine translation model.

| Simulation serial number | Column generation algorithm | | Maximum group algorithm | |
|---|---|---|---|---|
| | Operation time (s) | Optimization result (time slot) | Operation time (s) | Optimization result (time slot) |
| 1 | 173.99 | 94 | 0.102 | 94 |
| 2 | 203.80 | 89 | 0.080 | 89 |
| 3 | 195.17 | 66 | 0.081 | 66 |
| 4 | 198.56 | 85 | 0.090 | 85 |
| 5 | 171.23 | 74 | 0.080 | 74 |
| 6 | 188.65 | 74 | 0.080 | 74 |
| 7 | 206.43 | 90 | 0.075 | 90 |
| 8 | 88.948 | 41 | 0.110 | 41 |
| 9 | 208.946 | 75 | 0.080 | 75 |
| 10 | 98.412 | 62 | 0.080 | 62 |

optimization problem is NP-hard in nature. In this paper, the column generation algorithm and the maximal clique algorithm proposed in the above optimization algorithm have been used to make an experimental analysis on the optimization problem of Japanese passive in the statistical machine model. The application of this model has a certain complexity, so this paper optimizes the statistical machine translation model and simplifies the experimental process. The experimental results of optimization of 10 network nodes in the statistical machine translation model are shown in Table 2.

It can be seen from Table 2 that the two algorithms have the same optimization results when the demand is different. When the optimization result is 94, the operation time of the column generation algorithm is 173.99 s, and the operation time of the largest group algorithm is 0.102 s; but when the optimization result is 62, the operation time of the column generation algorithm is 98.412 s, and the operation time of the largest group algorithm is 0.080 s. The operation time of the maximal clique optimization algorithm is all less than 1, which is much smaller than that of the column generation algorithm. This has fully proved the correctness and effectiveness of the maximal clique algorithm for the optimization model of machine translation in Japanese passive statistics.

## 4. Discussion

This paper has mainly studied the machine translation of the joint optimization algorithm for scheduling in Japanese passive statistics. The article has analyzed the channel model, scheduling model, interference model, and business flow model in detail and proposed a column generation algorithm, a maximal clique algorithm, a system computing power scheduling algorithm, and a resource-aware scheduling algorithm to optimize machine translation. In the experimental process, this paper simulates and analyzes statistical machine translation in combination with the experimental damage of the scheduling algorithm and adopts a shared resource platform to compare the output data size and time of the two scheduling algorithms and analyze the performance of the algorithm, and combined with the passive characteristics of Japanese, the results of optimization of 10 network nodes in the statistical machine translation model are selected, and the conclusions of this paper are finally drawn. Based on the proposed algorithm model, the experiment on the performance of machine translation has found that the resource-aware scheduling algorithm was more efficient in allocating resources within the system. These technical principles have provided a theoretical basis for the machine translation work in Japanese passive statistics. However, since there are not many algorithms combined in this paper, there will be insufficient data when performing machine translation of Japanese passive statistics, so more scholars are needed to discuss and study.

## 5. Conclusions

This paper has improved the scheduling joint optimization algorithm of the machine translation model and focused on the analysis of the resource-aware scheduling algorithm, the computing power scheduling algorithm, the column generation algorithm, and the maximal clique algorithm for the optimization of Japanese passive in statistical machine models. The experimental results show that the accuracy of the resource-aware scheduling algorithm for dynamic translation optimization of Japanese was greater than that of the computing power scheduling algorithm. When the data volume is 2 G, it is 6% higher; when it is 4 G, it is 5% higher; and when it is 6 G, it is 21% higher. At the same time, the time spent has been reduced by 15.5%. It can be seen that the optimization effect of the resource-aware scheduling algorithm is better. According to the passive characteristics of Japanese and the optimization conditions in the machine translation model, this paper has used the column generation algorithm and the maximal clique algorithm to conduct experimental analysis in the network nodes to find the optimal optimization algorithm. The experimental results show that when the data is the same, the optimization results of the two algorithms are the same. However, the operation time of the maximal clique optimization algorithm is much smaller than that of the column generation algorithm, which fully proves the correctness and effectiveness of the maximal clique algorithm for the optimization model of machine translation in Japanese passive statistics. Machine translation is a focus in the future. This paper combines Japanese passive statistics with a large number of scheduling optimization algorithms, which brings a more meaningful reference direction for the development of machine translation.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The author declares that there are no potential conflicts of interest regarding the publication this study.

## Acknowledgments

This study was supported by the Science Foundation of Educational Commission of Anhui Province of China (Grant no. SK2021A0798).

## References

[1] Y. Cai, F. Cui, Q. Shi, M. Zhao, and G. Y. Li, "Dual-UAV-enabled secure communications: joint trajectory design and user Scheduling;Joint trajectory design and user scheduling [J]," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 1972–1985, 2018.

[2] Y. Fei, M. A. Huimin, and S. O. Business, "Joint optimization of batch-discrete re-enter production scheduling and preventive maintenance[J]," *Computer Integrated Manufacturing Systems, CIMS*, vol. 25, no. 1, pp. 44–52, 2019.

[3] J. Zhao and L. Mili, "Robust uk filter for power system dynamic state estimation with unknown noise statistics," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1215–1224, 2019.

[4] A. O. Al-Abbasi and V. Aggarwal, "VidCloud: joint stall and quality optimization for video streaming over cloud[J]," *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, vol. 5, no. 4, pp. 1–32, 2020.

[5] G. Song, "Accuracy analysis of Japanese machine translation based on machine learning and image feature retrieval," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2109–2120, 2021.

[6] G. Zheng and J. Zhu, "Japanese translation teaching corpus based on bilingual non parallel data model[J]," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 5, pp. 1–11, 2020.

[7] M. Mangeot-Nagata, "Erratum to collaborative construction of a good quality, broad coverage and copyright free Japanese-French dictionary," *International Journal of Lexicography*, vol. 31, no. 1, p. 113, 2018.

[8] J. Wang, C. Liu, and M. C. Zhou, "Improved bacterial foraging algorithm for cell formation and product scheduling considering learning and forgetting factors in cellular manufacturing systems," *IEEE Systems Journal*, vol. 14, no. 2, pp. 3047–3056, 2020.

[9] H. Khani and H. E. Z. Farag, "Joint arbitrage and operating reserve scheduling of energy storage through optimal adaptive allocation of the state of charge," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 4, pp. 1705–1717, 2019.

[10] L. Qi, "Tool requirement and pre-scheduling optimization model of the tool flow system of a digital workshop[J]," *Proceedings of the Institution of Mechanical Engineers - Part B: Journal of Engineering Manufacture*, vol. 231, no. 5, pp. 838–849, 2017.

[11] D. Zhou, M. Sheng, J. Luo, R. Liu, J. Li, and Z. Han, "Collaborative data scheduling with joint forward and backward induction in small satellite networks," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3443–3456, 2019.

[12] M. Rikters, R. Ri, T. Li, and T. Nakazawa, "Japanese–English conversation parallel corpus for promoting context-aware machine translation research," *Journal of Natural Language Processing*, vol. 28, no. 2, pp. 380–403, 2021.

[13] M. Khatami and S. H. Zegordi, "Coordinative production and maintenance scheduling problem with flexible maintenance time intervals," *Journal of Intelligent Manufacturing*, vol. 28, no. 4, pp. 857–867, 2017.

[14] T. Lu, E. Yao, Y. Zhang, and Y. Yang, "Joint optimal scheduling for a mixed bus fleet under micro driving conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2464–2475, 2021.

[15] B. Lv, R. Wang, Y. Cui, Y. Gong, and H. Tan, "Joint optimization of file placement and Delivery in cache-assisted wireless networks with limited lifetime and cache space," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2339–2354, 2020.

[16] X. Xu, Y. Xue, M. Cui, Y. Yuan, and L. Qi, "Joint optimization of energy conservation and migration cost for complex systems in edge computing," *Complexity*, vol. 2019, no. 12, pp. 1–14, Article ID 6180135, 2019.

[17] J. Zhu, G. Zhang, Z. Zhu, and K. Yang, "Joint time switching and transmission scheduling for wireless-powered body area networks," *Mobile Information Systems*, vol. 2019, no. 4, pp. 1–11, Article ID 9620153, 2019.

[18] K. Abe, Y. Matsubayashi, N. Okazaki, and K. Inui, "Multi-dialect neural machine translation for 48 low-resource Japanese dialects," *Journal of Natural Language Processing*, vol. 27, no. 4, pp. 781–800, 2020.

[19] J. R. Lee, "Potential of machine translation in headline translation: f: focusing on Korean-Japanese translation of headlines[J]," *Interpretation and Translation*, vol. 21, no. 2, pp. 119–144, 2019.

[20] S. D. Budiwati and M. Aritsugi, "Word reordering on multiple pivots for the Japanese and Indonesian language pair," *Machine Translation*, vol. 35, no. 4, pp. 611–636, 2021.

[21] N. Wang, J. Xu, and F. Ming, "Integrating voice features into Japanese-English hierarchical phrase based model[J]," *Acta Scientiarum Nauralium Universitaties Pekinensis*, vol. 53, no. 2, pp. 305–313, 2017.

[22] W. Xu and M. Carpuat, "EDITOR: an edit-based transformer with repositioning for neural machine translation with soft lexical constraints :an edit-based transformer with repositioning for neural machine translation with soft lexical constraints[J]," *Transactions of the Association for Computational Linguistics*, vol. 9, no. 1, pp. 311–328, 2021.

[23] I. Goto, "Neural network machine translation technology[J]," *Broadcast Technology*, vol. 70, p. 21, 2017.

[24] T. Bezdan, C. Stoean, A. A. Naamany, N. Bacanin, and K. Venkatachalam, "Hybrid fruit-fly optimization algorithm with K-means for text document clustering," *Mathematics*, vol. 9, 2021.

[25] S. Li, X. Ao, F. Pan, and Q. He, "Learning policy scheduling for text augmentation," *Neural Networks*, vol. 145, pp. 121–127, 2022.

[26] X. Wu, L. Wang, Y. Xia et al., "Temporally correlated task scheduling for sequence learning," in *Proceedings of the*

*International Conference on Machine Learning*, Beijing China, July 2020.

[27] N. Bacanin, M. Zivkovic, T. Bezdan, K. Venkatachalam, and M. Abouhawwash, “Modified firefly algorithm for workflow scheduling in cloud-edge environment,” *Neural Computing & Applications*, vol. 34, no. 11, pp. 9043–9068, 2022.

[28] L. Jie, “Modeling and analysis of translation ability of applied English talents based on data mining algorithm,” *International Conference of Social Computing and Digital Economy*, pp. 211–214, 2021.

[29] Z. Wenjie, “English automatic translation platform based on BP neural network for phrase translation combination,” in *Proceedings of the International Conference on Frontier Computing*, Springer, Singapore, September 2022.