A scheduling algorithm for social workflow optimization based on hybrid cloud streaming mechanism

Min Chen*

Heilongjiang Bayi Agricultural University(HLAU), Daqing, Heilongjiang, China

Email:qweer1212138@163.com

Abstract. In order to solve the current workflow scheduling problem, a new optimal scheduling algorithm is proposed in this paper, which has the advantages of high cost, poor time convergence, poor cost accounting quality and so on. Then, a self-evolving workflow structure is constructed to integrate the decentralized workflow into work subprocesses that can be processed independently, and secondary costing is carried out according to the granularity of workflow. Using the numerical characteristics of work subflow, the recursive algorithm is used to calculate the overall cost of each work subflow, which reduces the computational complexity. The simulation results show that the proposed method has higher cost efficiency than the traditional collinear chaotic particle cost calculation and integrated evolutionary scale calculation.[1]

Keywords: social workflow; Chaos cloud shunt; Cost accounting; Working substream

1 Introduction

With the rapid development of cloud computing technology, a large number of network planning methods also come into being. However, due to the complexity of the social work process of big data, the cost accounting of one confirmation can only be confirmed once, resulting in the existing scheduling algorithm cannot meet the requirements of high time and high cost, so the promotion of this method is limited to a certain extent.[2]

Aiming at the "bottleneck" problem faced by this field at present, the applicant has designed several adaptive scheduling methods for social work oriented to big data according to the application requirements of big data, and made a preliminary exploration of their applications. Thelwall and others put forward an adaptive scheduling method of workflow in big data society on the basis of Larange's self-deletion. This method can dynamically configure the Lagrange check factor while configuring big data, and can automatically delete the Cache factor during the configuration process. However, the traditional sparse representation method based on sparse representation usually adopts one-to-one sparse representation method, and does not make full use of the convergence problem caused by the decrease of global service quality, which leads

A. Rauf et al. (eds.), *Proceedings of the 3rd International Conference on Management Science and Software Engineering (ICMSSE 2023)*, Atlantis Highlights in Engineering 20,
https://doi.org/10.2991/978-94-6463-262-0_11

to its high computational complexity. On the basis of comprehensive cost control, Bal gives an adaptive scheduling algorithm for social workflow of big data, and proves that it can perform real-time calculation in the area that can be received, and the calculation result has a high convergence speed. However, this algorithm is mainly based on fixed control, and its calculated traffic index vector is not very flexible, which causes some serious problems that the queuing cache can not be updated in time when the data is very random, thus affecting the practicability of the algorithm. Mohammad based on the congestion control of stream cache, he suggested that the availability of the overall load of the algorithm should be improved through random mapping, and then the dynamic convergence effect of the algorithm should be enhanced through the congestion control of stream cache. However, this method needs to use heterogeneous methods to rebuild the availability, and when the workflow parameters are randomly changed, the robustness will be reduced due to frequent Cache switching.

Aiming at the problems existing in current social work, this paper proposes an adaptive sorting optimization algorithm based on big data.[3] Using the Larrangian self-deletion theory, Thelwall et al. proposed an adaptive social workflow sorting algorithm based on the random characteristics of Lagrange check factors, which can carry out arbitrary configuration of a large number of social work processes, and can automatically delete the buffer factor when publishing. However, since this method needs to be compared one by one in different Spaces, it will not cause convergence problems due to the decline of the overall service level, which leads to the complexity of the algorithm. Bal [4] proposes a new adaptive scheduling algorithm for social workflow based on integrated cost control. The results show that this algorithm can calculate the real-time cost of enterprises efficiently and has a high convergence rate. Based on the blocking control characteristics of the flow buffer, Mohammad [5] improved the random mapping method to improve the load utilization of the whole system, and introduced the flow buffer to improve the dynamic convergence of the algorithm.

However, this method requires a variety of methods to realize the availability reconstruction of the system. When the parameters in the operation process are random, the robustness of the system will be reduced due to the update of the buffer. In order to overcome the shortcomings of existing algorithms, this paper proposes a new algorithm based on the principle of hybrid cloud streaming to adapt to the needs of a large number of workflows. Firstly, the time factor method is used to construct the self-evolving structure of chaotic cloud, and the segment caching technique is used to realize the adaptive cache diffraction. Secondly, the reverse evaluation method is used for buffer separation of the system, and the most efficient branch branch is used for detection, so that the convergence performance is improved. NS2 simulation results show that the proposed algorithm is effective.

2 Time factor construction of self-evolving structure of flow chaotic cloud

In the process of building the social work of big data, the upload and quality indicators of each substream are effectively divided into different stream dimensions. Therefore, the corresponding stream has good self-iteration ability and can be reconstructed according to the time factor. The current algorithms do not take this problem into account. In the implementation process, due to the unstable upload of substreams and the stability of quality indicators, it is difficult for the iterative algorithm to evaluate the whole process, resulting in a sharp decline in the convergence speed. Therefore, the time domain analysis method should be fully introduced to classify the data streams at each stage, then carry out cost accounting, and finally carry out secondary optimization, so as to improve the convergence of the algorithm, reduce the overall cost, and achieve the best effect.[6]

2.1 Complete calculation process

At any given moment, the optimization of costing can be divided into three levels: evolution structure reconstruction, flow dimension partitioning and cost control. In this paper, three performance indexes are proposed: structure evolution chaos degree, segmentation cost and convergence cost, and the processing of chaotic cloud flow is introduced in detail.[7]

(1) According to the possible distribution of the sub-business, the flow dimension and classification mode of the sub-stream at the next moment are estimated in advance, and the access is carried out according to the principle of the lowest cost;

(2) Reconstruct the data obtained in the first stage, and then conduct secondary access according to the principle of the lowest access cost until all the working sub-businesses are processed;

2.2 Transfer mechanism of chaotic cloud gas

According to the above steps, the whole system is divided into two parts.[8]

Firstly, the initial cost is estimated based on the current cost probability by using the discrete parameters of a single working subflow. Since the work subbusiness is discontinuous in all dimensions and has some randomness, its initial cost can be determined according to its size. In terms of the time factor t0, the cost vector of the flow is reconstructed from the accumulated particle size, and its cost is the lowest, reaching the size of a single particle. In this way, the process conforms to the T-shaped distribution, and its characteristic index is λ, and in the chaotic cloud, the distribution of each working substream takes on the characteristic of t, and its time factor X satisfies the same specific index. In addition, in the actual process, each work sub-business is assigned with its own characteristics, so the cost function MQ is:

$$M_Q = \min(\gamma E) \quad (1)$$

Where, EX is the first-level average cost index, and EX2 is the second-level average cost index. After obtaining the cost function in the bifurcating process of chaotic cloud, it is also necessary to verify the time cost of different working substreams by segment. If the distribution probability of the cost also satisfies Formula (1), it indicates that the working substream will also be in an independent evolving load. If the distribution probability of the cost does not meet MQ, it is necessary to construct the evolutionary structure of the work substream, form a new digital feature MQ, and continue the iteration. The judgment criterion is shown in Formula (2).

$$\gamma T(x) = \begin{cases} 0, & x < M_Q \\ 1, & x > M_Q \end{cases} \quad (2)$$

$$(M_Q)^{P'} = M_Q(x) \quad (3)$$

Where, the construction rate p' of the cost function satisfies the Poisson distribution with parameter λ, and its convergence index x satisfies the following characteristics, as shown in Formula (4).

$$\min M_Q(x) = \begin{cases} 1 - x, & x < 0.5 \\ 1, & x > 0.5 \end{cases} \quad (4)$$

Although Formula (4) completes the construction of the cost function, the process is in discrete state, so it is necessary to carry out interval quadratic mapping of the cost function in discrete state. If the cost function p(x) of a certain working substream is set as μ, then the accounting success rate p satisfies Formula (5).

$$P = \min M_Q(x) + \max M_Q(x) \quad (5)$$

Since there is a significant positive fluctuation relationship between the chaotic cloud space δand the accounting success rate of each working substream [9], the accounting cost can be divided according to the time factor, and the extraction period in the segmentation process is Δt.

3 Algorithm Flow

The content is constructed by the self-evolving structure of the flow chaos cloud based on time factor. The algorithm process in this paper is carried out as follows:

(1) The work substream is acquired first, and the data is processed according to the principle of first come first processing;

(2) For any working substream, search for all processable particle size in the chaotic cloud space. Before the processing is completed, cost accounting is carried out according to the formula; otherwise, the arrived particle size will continue to be received until the granularity of all working substreams is processed;

(3) Calculate the cost of the whole chaotic cloud space;

(4) End of algorithm. The flow chart of the whole algorithm is shown in Figure 1.

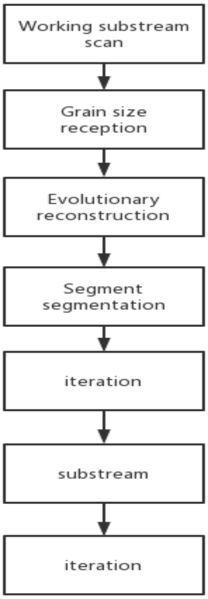

Fig. 1. Flow chart of social workflow adaptive scheduling optimization algorithm

4 Simulation experiment

4.1 Simulation environment Settings

For evaluating scheduling optimization algorithm is proposed in this paper, the simulation experiments using NS - 2 simulation platform for cost accounting mechanism, which affects the chaos particles (CollinearityChaosParticleCostAccounting, 3C_PA mechanism) and integrated evolution scale accounting (IEA mechanism) are simulated and compared. Experimental simulation parameters are shown in Table 1.

Table 1. Simulation parameters table

parameter	Numericalvalue
Flowsimulationtime/h	12
Flowdensitydensity(PCS/s)	10、20、30、40、50、60、70
Granularityperiod/s	＜512
Granularityarrivalprobability	＜0．6
Grainsizecoarseness	1024bit

In the simulation experiment, in order to verify the superiority of the proposed algorithm, the experiment focuses on the simulation and comparison of four indexes: particle size processing density, cost accounting cycle, particle size loss and convergence rate of flow coarseness.

5 Conclusion

In this paper, a new method is proposed to solve the problems of high cost, poor time convergence and low cost accounting quality existing in the current social workflow adaptive scheduling algorithm. Firstly, the time model of social work process is established by using the big data model, and the chaos cloud flow is divided, and the cost effectiveness evaluation is carried out. The simulation results show that the cost of workflow can be reduced effectively and the efficiency of workflow can be improved.

Project type and number: 2022 Innovation and Entrepreneurship Training Project for College Students in Heilongjiang Province S202210223064. Project name: Research on the integration of "Internet +" into community social work in the post-epidemic Era

References:

1. Shi Yufeng. Data Scheduling Optimization Method for P2P Streaming Media Based on Big Data [J]. Electronic Testing, 2018,11 (1):92-94.
2. JENKINSR, BURTONAM. Asurveyof100%energyaccu-racyinWSN [J]. Autonomousrobots, 2013,451 (31):435-441.
3. THELWALLM WILKINSOND, SUKHVINDERU. Dataminingemotioninsocialnetworkcommunication: GenderdifferencesinMySpace [J]. JournaloftheAmericanSocietyforInformationSocietyforInformation-ScienceandTechnol - ogy, 2010, 21 (1) : 190-199.
4. BALM. Roughsetstheoryassymbolicdataminingmethod: anapplicationoncompletedecisiontable [J]. InformationSciencesLetters, 2013, 2 (1) : 111-116.
5. MOHAMMADMR. Mediumaccesscontrolindistributedcognitivenet [J]. IEEEWirelessCommunication, 2013, 17 (4) : 41-51.
6. HAMDAOUIB, ENSANS ASurveyonclusteringalgorithmsforwirelesssensornetworks [J]. OpticsandLasersinEngi-neering, 2013,56 (19):1541-1549.
7. Li Mingchu, Xu Lei, Sun Weifeng, et al. Grid resource allocation Model Based on Incomplete Information Game [J]. Journal of Software, 2012, 23(2):428-438.
8. Wu Wenqian. Dynamic workflow Scheduling Method Based on elastic resources in hybrid cloud environment [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2019:6-15. (in Chinese)
9. HEMPELM, PENGDM Asurveyofenergyefficientcom pressionandcommunicationtechniquesformultimediain - networks [J]. IEEECommunicationsSurveys, 2013, 14 (99) : 2541-2549.