A Cloud Computing Task Scheduling Method Based on Genetic Algorithm

Huabin Zhang

1579401432@qq.com

Shandong University of Science and Technology, Can only equip the academy, Yantai, Shandong Province

ANSTRACT: In order to solve the traditional Internet service ability and expanding an acute shortage of, cloud computing has come into being, and has universal service in today's Internet services. Due to the huge scale of cloud computing and the diversity of applications and tasks, the performance requirements of cloud computing have been significantly improved. The characteristics of cloud computing brings the challenge of task scheduling, affect the performance and reliability of the system. To solve these problems, this paper puts forward cloud computing task scheduling algorithm based on genetic algorithm method. Firstly, the model of task and cloud computing resources and the objective function were established. Then, based on the task and cloud resource model, the gene genetic algorithm for task scheduling was proposed. Finally, through the experiment show that the algorithm has good results on the cloud resource scheduling, can reduce the total cost of users, to reduce the energy consumption.

Keywords—Cloud computing, genetic algorithm, task scheduling

1. INTRODUCTION

With the progress and development of society, network resources have become an indispensable part of social development. Cloud computing is a model used to realize the sharing of configurable computing resources and simple and fast network access. These resources can be quickly configured and produced with little management resources or interaction with service providers [1]. The most common ones are Web search engines and webmail, such as Google and Baidu.You only need to search any resources you want on the search engine through the mobile terminal. Data resources are shared through the cloud, which also generates a variety of application fields such as storage cloud, medical cloud, financial cloud and so on. Cloud computing refers to on-demand access to computing resources through the Internet. Compared with the traditional Internet, cloud computing has high reliability, scalability, low price and so on. These characteristics make the cloud platform can reduce costs, improve agility and shorten the time to value realization, expand more easily and effectively, and solve the problem of performance attenuation of large information systems caused by insufficient computing and management capabilities. For example, IaaS provided by cloud service platform can produce large-scale application deployment, and can also provide access to computing resources with high operability and scalability [2]. At the same time, cloud computing has the advantages of unlimited storage capacity and reducing computer operating costs. The significant advantages of cloud computing make the social demand for cloud computing more and more [3]

ICIDC 2023, June 02-04, Nanchang, People's Republic of China

DOI 10.4108/eai.2-6-2023.2334608

However, despite its many advantages, the adoption of cloud computing still suffers from great opposition [4]. Due to the huge scale of cloud computing, being prone to hacker attacks, resources being misused, and diverse types of applications and tasks, the performance requirements of cloud computing are significantly improved. Verify the performance and reliability of the system that affects the system. For example, Amazon's cloud computing service suffered a large-scale failure, resulting in the failure of services such as Disney's streaming media subscription service Disney +, Netflix, Slack, Robinhood, and Coinbase, the largest cryptocurrency exchange in the United States. Even directly implicated in the paralysis of Amazon's own delivery business. Therefore, task scheduling and performance optimization in cloud computing is very important.

To solve these problems, this paper puts forward cloud computing task scheduling algorithm based on genetic algorithm, the method of methods can be divided into three steps.In the first step, the model and objective function of task and cloud computing resources are established. In the second part, the gene genetic algorithm for task scheduling is established based on the task and cloud resource model. Finally, experiments show that the algorithm has good results for cloud resource scheduling, which can reduce the total cost of users and reduce energy consumption.

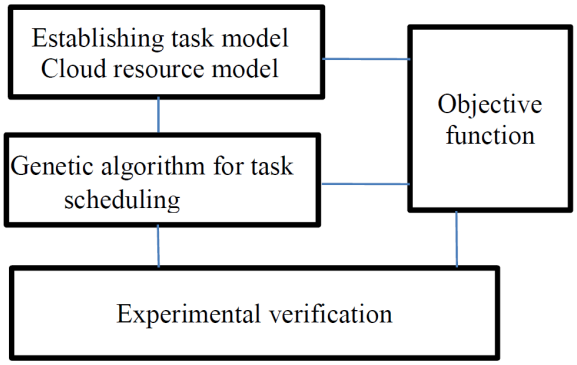

Fig. 1. Cloud computing task scheduling framework based on genetic algorithm

2.RESEARCH BACKGROUND

2.1Cloud Computing Technology

Cloud computing can be configured and produced quickly with few management resources or service provider interactions, and it has been well developed in various fields. At the same time, it has the ability to reveal the hierarchical structure existing in the real network [5], and it can be used as a guideline that individuals or organizations can implement. To artificially manage their social circle or change their communication mode or channel to protect their privacy from being disclosed [6-7]. Cloud computing also has advantages such as improved performance: computers in cloud computing systems run faster and boot up faster than traditional computers because they have fewer loaders and processes in memory.

Unlimited storage capacity: The amount of storage in cloud computing is very large, and can even reach the point of almost unlimited.

Data reliability: If a user's personal computer goes wrong or even crashes, all of the user's data is in the cloud computing platform and safe and still accessible.

Reducing software cost: Users can obtain most of the required software without purchasing expensive software applications.

Easier teamwork: Multiple users can more easily collaborate on documents and projects.

Device independence: Users are no longer linked to a single computer or network. Characteristics of.

Education cloud: A development of education informatization, education cloud can virtualize any educational hardware resources needed, and then it is introduced into the Internet to save, to provide a fast education cloud service platform to educational institutions and students and teachers.The universal use of the latest education field model is based on cloud technology and electronic learning [8], many university are use the MOOC school online courses.Cloud-based learning models will continue to be developed rapidly. At present, e-learning will enter a brand new era through cloud computing [9]

2.2Characteristics of Cloud Computing

Cloud computing has the characteristics of high reliability, super scale, scalability, virtualization and so on. High reliability cloud platform uses a variety of measures such as data multi-replica fault tolerance, computing nodes isomorphic and interchangeable to ensure the high reliability of services, through the very large scale data center, multi-data center technology to achieve data center city disaster recovery. The hypervior layer is used to isolate the damage of the underlying physical hardware, and mature and highly available services (such as AWS LBS, SNS, SQS) are used to decouple, reduce the complexity of application system use and improve system reliability. Real-time data synchronization and transaction rollback mechanism are implemented to ensure data integrity and unity. The large-scale cloud computing center has a considerable scale, and the number of servers of many companies supplying cloud computing has reached hundreds of thousands or millions. For example, the number of cloud computing servers of Google has reached more than 1 million, and the number of cloud computing servers of Amazon, IBM, Microsoft and Yahoo also has hundreds of thousands. Scalability In cloud computing, physical or virtual resources can be rapidly scaled horizontally with strong elasticity. Through automated processing, resources can be quickly pruned. Cloud service users can access nearly unlimited physical or virtual resources through the network at any time and place, and do not have to worry about resource size and capacity planning.

The deployment, monitoring and security functions used in the cloud computing system further ensure the reliability of services. The fault tolerance measures of extremely cheap cloud computing allow it to use cheap nodes to constitute the cloud.

Improved document format compatibility: Users do not need to worry about compatible files to create other users, applications or operating systems on their computers, such as desktop office cloud systems with strong adaptability based on the development capabilities of cloud service providers [10]

2.3The challenge of cloud computing

The use of cloud service platforms by organizations may lead to their emergence highly dependent on the availability and reliability provided by the cloud platform. And if the cloud service provider does not provide a timely and appropriate disaster recovery mechanism, if there is a problem, it will have unimaginable consequences for the organization.

The most obvious disadvantage of cloud computing is network dependency. Service interruption or power outage will make access to cloud computing service platforms and other facilities problematic. In addition, bandwidth is also one of the important issues affected by cloud computing. For example, during the peak period of the application, if the service of the cloud service provider is fully loaded during the peak period, then the availability and performance will be greatly reduced. Even though many cloud service providers have special room for handling the peak period, and other unpredictable risks, external attacks or server failures, etc. Any one of these poses a threat to customer access.

In the best case scenario, the cloud vendor won't go bankrupt and a larger company won't swallow it up. If this happens, the server may be shut down abruptly, and the organization should plan for this in advance.

Once the organization has submitted the data to the management system responsible for by the cloud service provider, the provider will no longer have full control over these submitted data, and it is difficult for the provider to take some technical and administrative measures to protect these data. It will face problems such as lack of availability, lack of integrity, lack of intervention, lack of confidentiality and lack of isolation.

3.TASK SCHEDULING MODELING FOR CLOUD COMPUTING

3.1Cloud Computing task and resource modeling

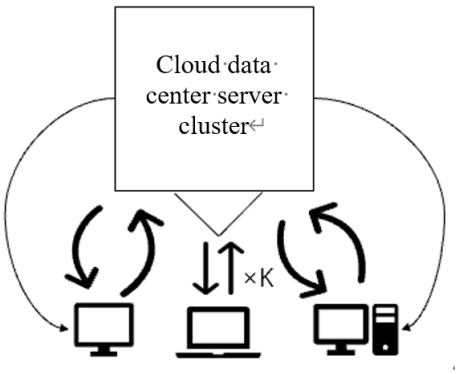

Fig. 2. An example of cloud computing task scheduling

Cloud computing is generated by the combination of Internet, multi-core chips, virtualization and other technologies. It is very important for the task scheduling strategy of cloud platform, because efficient management of resources in cloud platform is the essence of task scheduling [11]. In view of the above problems, as shown in Figure 2, In the computer field, ant colony

algorithm is used to optimize the information used by paths, as proposed in article [12], and pheromone matrix is initialized when it stalls to improve its performance. Article [13] proposes to reduce its completion time by improving the scheduling algorithm of particle swarm optimization, and article [14] improves the efficient memory constraint parallel algorithm. By using fuzzy mean clustering based on memory constraints to reduce operating costs, this paper proposes a research method of cloud computing task scheduling based on genetic algorithm [15]. In this method, the basic environment of cloud computing is first modeled, and the cloud computing task model, cloud computing resource model and objective function composed of time, cost and other factors are established. Genetic algorithm is used to realize reasonable scheduling of cloud computing tasks and resources to meet the demands of time, cost and other factors. Finally, experimental analysis shows the effectiveness of this method, which plays a positive role in improving resource use and task scheduling. An important problem in cloud computing research is task scheduling.

Suppose that there are d_n tasks waiting to be allocated, which can be expressed as follows.

$$\mathrm{Duty}_{\mathrm{series}} = \{\mathrm{d1}, \mathrm{d2}, \mathrm{d3}, \ldots\ldots, \mathrm{dn}\} \quad (1)$$

Where $\mathrm{Duty}_{\mathrm{series}}$ represents the total task.

To facilitate the division of different tasks in the total task, assume that each task model is as follows.

$$\mathrm{D_i} = \{\mathrm{c_i}, \mathrm{d_i}, \mathrm{me_i}, \mathrm{time_i}\} \quad (2)$$

Where $\mathrm{c_i}$ is the task amount requirement, $\mathrm{d_i}$ is the disk space requirement, $\mathrm{me_i}$ is the memory space requirement, and $\mathrm{time_i}$ is the time requirement.

Network resource model is a formal description of the managed objects (i.e., network resources). It is a high generalization and abstraction of communication network resources. It describes all kinds of resource classes, resource class attributes and the relationships between resource classes in mobile networks, and provides a uniform abstract description of network resources, such as wireless network resource model, transportation network resource model, and spatial resource model.

Here, the resources on the server side are defined as virtual machines or sub-servers, and are denoted as C

$$C = \{C1, C2, C3, \ldots., Cn\} \quad (3)$$

And each server-side resources are defined in detail.

$$Ci = \{Memory、Number\ of\ cores、Network\ bandwidth\} \quad (4)$$

3.2Modeling task scheduling in Cloud Computing

After the definition of the task model is completed, we are faced with the calculation of the server performance, namely the time allocated to the task. Meanwhile, the time demand timei is composed of various data quantities, such as upload delay

$$T_{iup} = \frac{c_i}{R_i} \tag{5}$$

Here, R_i is the transmission rate of network bandwidth allocation. The task processing time is represented as follows

$$T_j = \frac{c_i}{F_j} \tag{6}$$

Where F_j represents the processing capacity of the server.

The download time is expressed as follows

$$T_{idown} = \frac{c_i}{R_i} \tag{7}$$

The total time is denoted by

$$T_{itotal} = sum\left(T_{iup} + T_j + T_{idown}\right) \tag{8}$$

After obtaining the amount of data for the allocated time of the server, the corresponding energy consumption should also be calculated as to achieve the optimal effect.

$$E_i = \left(\frac{c_i}{F_j}\right) * P_i \tag{9}$$

Here, $P_i$represents the computational power of the server, denoted by

$$E_{up} = \left(\frac{c_i}{F_j}\right) * P_i \tag{10}$$

$$E_{down} = \left(\frac{c_i}{F_j}\right) * P_j \tag{11}$$

In summary, the total time model can be obtained:

$$T_{(0-n)}\, sum = \sum_1^n \left(T_{iup} + T_j + T_{idown}\right) \tag{12}$$

The cost model is denoted by

$$cost = T_{total} \tag{13}$$

Where T_{total} denotes the sum of multiple tasks

The total energy consumption is denoted by

$$Engergy = E_i = \left(\frac{c_i}{F_j}\right) * P_i + \left(\frac{c_i}{F_j}\right) * P_i + \left(\frac{c_i}{F_j}\right) * P_j \tag{14}$$

Overall objective:

$$Target = fun(Wt * Time, Wc * Cost) \tag{15}$$

Here, fun is defined as the selection function, Wt is expressed as the weight of time, and Wc is expressed as the weight of consumption.

4.OPTIMIZATION OF MULTI-OBJECTIVE TASK SCHEDULING BASED ON GENETIC ALGORITHM

4.1Gene genetic algorithm

Compared with the traditional cloud computing task scheduling algorithms such as Min-Min method [16] and Sufferage method [17], the cloud computing task target scheduling method based on genetic algorithm (GA) firstly modeled the cloud computing IaaS resources, and established the computing power, cost calculation model and consumption time model of multiple servers in cloud computing.IaaS cloud computing task model is established. The objective function composed of time, cost and other factors was constructed. The matching degree and matching pros and cons of tasks and resources could be obtained by calculating each component of the function.Genetic algorithm is adopted to achieve the target optimization matching between tasks and resources. Genetic algorithm is designed and proposed according to the evolution law of organisms in nature. It is a computational model simulating the biological evolution process of natural selection and genetic mechanism of Darwinian biological evolution, and a method to search for the optimal solution by simulating the natural evolution process.The algorithm by means of mathematics, the use of computer simulation calculation, converting the problem solving process similar to the evolution of chromosomes gene of crossover and mutation process.

4.2Genetic algorithm for task cloud service resource scheduling

Given a set of independent tasks $Duty$

$$Duty = \{d1, d2, d3, \ldots\ldots, dn\} \tag{16}$$

The constraint of each independent subtask in is

$$D_c = \{C_c, M_r, N_w\} \tag{17}$$

Where D_c represents the constraints of the task, C_c represents the computing power constraint of the task, M_r represents the memory requirement constraint of the task, and N_w represents the network bandwidth constraint of the task.

At the same time, each subtask also has its own required requirements, i.e

$$D_r = \{time, cost\} \tag{18}$$

Where $time$ represents the time requirement of the task, and $cost$ represents the consumption requirement of the task.

Similarly, the gene genetic algorithm for task cloud service resource scheduling also has resource scheduling requirements

$$C = (C_1, \ldots\ldots C_n)$$

$$Ci = \{ Number\ of\ cores\ memory\ bandwidth\} \tag{19}$$

The overall objective

$$Target = Fun\{\ \ (Wt * Time, Wc * Cost\} \tag{20}$$

On the server side, the same is true

$$C = (C_1, \ldots . C_n) \tag{21}$$

$$Ci = \{ Number\ of\ cores\ memory\ bandwidth \}$$

4.3Algorithm Description.

In this paper, genetic algorithm is proposed to solve the problems of slow running speed and high cost consumption of traditional algorithm. The building blocks of the basic genetic algorithm:

- Chromosome encoding method: using a fixed length of binary string to represent a group of individuals.
- Individual fitness evaluation: Is a measure of an individual's dominance in a population, used to distinguish between "good and bad" individuals. Fitness is calculated using the fitness function.
- Genetic operators: Selection operator, crossover operator, mutation operator. In the process of optimization, if it is difficult to know the solution steps of the problem in advance, searching in the solution space becomes one of the more extensive strategies.
- The operating parameters of the basic genetic algorithm Q: M: population size, that is, the number of individuals in the population, generally taken as 20100. T: The termination evolution generation of the genetic operation, generally taken from 100 to 500. Crossover probability (0.4 to 0.99) Pm: Mutation probability, generally from 0.0001 to 0.1.

After the discovery of the optimization problem, and presents a cloud computing task scheduling algorithm based on genetic algorithm to solve the above problems. Firstly, the initial iteration number is input, and then its fitness value is calculated, and the one with high fitness is selected as the next generation elite and new individuals are continuously generated. Finally, the individual that achieves the effect of the optimization algorithm is output. In the algorithm, C represents the encoding mode of the individual. E is the individual fitness evaluation function; P is the initial population; M is the population size; D is the selection operator. F is the crossover operator; A is the mutation operator; T is denoted as the genetic operation termination condition algorithm.

Parameter description:

The process of the whole algorithm can be divided into the following steps:

Initialize population P and set the fitness value of all individuals to 0.

Set the initial iteration number t=0 to start the iteration process.

Calculate the fitness value of each individual in the population and update the fitness value of each individual in the population P(t).

For each individual, crossover operation or mutation operation is selected according to a certain probability to generate a new individual.

Select the individuals with the highest fitness value as the next generation elite to ensure that they enter the next generation population.

The new individuals and elites generated will form the next generation population P(t+1).

Iteration number t+1, repeat the above process.

When the final result has reached the termination condition, the iterative process is finished and the result is displayed

Genetic algorithm is an optimization algorithm that simulates natural selection. By simulating natural processes such as heredity, crossover and mutation, it iteratively optimizes individuals in the population to find the global optimal solution or approximate optimal solution.

```
SGA= (C，E，P，M，D，F，A，T)，
Procedure SGA
begin
   initialize P (0);
   t=0;
while (t ≤ T) do
   for i= 1 to M do
            Evaluate fitness of P (t );
   end for
   for i= 1 to M do
                Select operation to p (t );
        end for i =1 to M/2 do
              Crossover operation to P (t);
    end for
        for i = 1 to M do
              Mutation operation to P (t);
end for
for i= 1 to M do
Mutation operation to P (t );
      end for
      for i= 1 to M do
      P (t+1) =P (t)
      end for
      t=t+1;
   end while
   end;
```

5.SIMULATION EXPERIMENTS

5.1Experimental parameter setting

Through the establishment of simulation experimental environment, we can verify the effectiveness of the algorithm.

Experimental environment is processor Intel Core, memory, operating system.

Simulation experiment environment is that the number of cloud servers is 1, the number of tasks is 10, the maximum number of iterations is 50, like table 1 and 2.

TABLE 1. SERVER SELF COEFFICIENT

Server self coefficient		
parameters	***Meaning***	***data***
Fi	Processing power of cloud services	10GHz
Ci	Resources owned by the server	2G
Pi	Server computing power	10KW
Ri	The transmission rate of the network bandwidth allocation	30M

TABLE 2. QUANTITY SOUGHT

Quantity sought	
parameters	***Subhead***
Ei	Energy consumption generated by task processing
Ti	Time required for task processing

5.2 Analysis of experimental results

Through the calculation and comparison of multiple groups of data, it can be found that the genetic algorithm has relatively less resource and energy consumption demand for the server, and at the same time, the processing time of the task also shows a downward trend as shown in the following figure 3 and figure 4

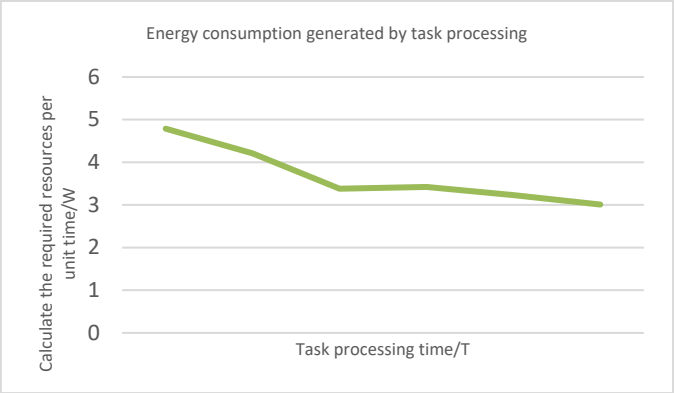

Fig. 3. Energy consumption generated by task processing

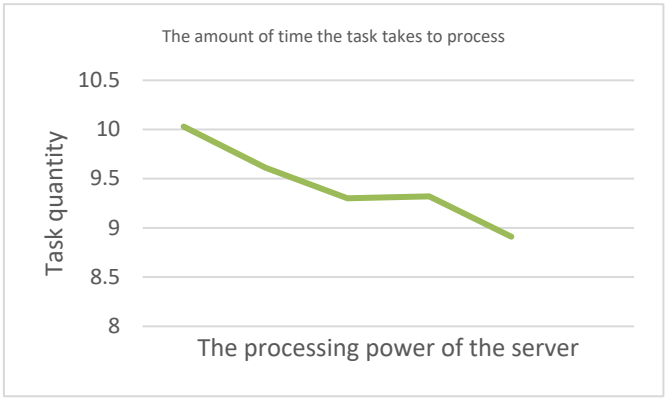

Fig. 4. The amount of time the task takes to process

6.CONCLUSION

In the application process of cloud computing, we found that cloud computing is huge in scale, prone to hacker attacks, resources are misused, and applications and task types are diversified, which leads to significant improvement in the performance requirements of cloud computing. The characteristics of cloud computing bring challenges to task scheduling. To verify the performance and reliability of the system, this paper shows a method of cloud computing task scheduling algorithm based on genetic algorithm, which is divided into three steps. The first step is to establish the model of task and cloud computing resources and the objective function. The second part is to establish the gene genetic algorithm for task scheduling based on the task and cloud resource model. The above experiments can clearly show that the algorithm can reduce the burden of cloud resource scheduling, reduce the demand for energy, and reduce the time and money spent by users.

REFERENCES

[1] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

[2] Sreenu, K., & Malempati, S. (2019). MFGMTS: Epsilon constraint-based modified fractional grey wolf optimizer for multi-objective task scheduling in cloud computing. IETE Journal of Research, 65(2), 201-215.

[3] Durao, F., Carvalho, J. F. S., Fonseka, A., & Garcia, V. C. (2014). A systematic review on cloud computing. The Journal of Supercomputing, 68, 1321-1346.

[4] Al-Hujran, O., Al-Lozi, E. M., Al-Debei, M. M., & Maqableh, M. (2018). Challenges of cloud computing adoption from the TOE framework perspective. International Journal of E-Business Research (IJEBR), 14(3), 77-94.

[5] Moscato, V., Picariello, A., & Sperli, G. (2019). Community detection based on game theory. Engineering Applications of Artificial Intelligence, 85, 773-782.

[6] Mittal, S., Sengupta, D., & Chakraborty, T. (2021). Hide and seek: outwitting community detection algorithms. IEEE Transactions on Computational Social Systems, 8(4), 799-808.

[7] Waniek, M., Michalak, T. P., Wooldridge, M. J., & Rahwan, T. (2018). Hiding individuals and communities in a social network. Nature Human Behaviour, 2(2), 139-147.
[8] Samyan, N., & St Flour, P. O. (2021). The impact of cloud computing on e-Learning during COVID-19 pandemic. International Journal of Studies in Education and Science, 2(2), 146-172.
[9] Abdalrahman, A. E. M., & Altaieb, M. O. (2018, April). E-learning based on Cloud Computing technology. In 2018 1st International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-6). IEEE.
[10] Pan, H., Yuan, Y., Song, W., & An, Z. (2019). The Design and Implementation of Secure Cloud Desktop System. The International Conference on Signal and Information Processing, Networking and Computers, 212-218.
[11] Zhu, C., Leung, V., Hu, X., Lei, S., & Yang, L.T. (2013). A Review of Key Issues That Concern the Feasibility of Mobile Cloud Computing. IEEE International Conference on Green Computing & Communications & IEEE Internet of Things & IEEE Cyber. IEEE Computer Society, 769-776.
[12] Ning, J., Zhang, Q., Zhang, C., & Zhang, B.. (2018). A best-path-updating information-guided ant colony optimization algorithm. Information ences, s 433-434, 142-162.
[13] Ebadifard, F., & Babamir, S. M. (2018). A PSO-based task scheduling algorithm improved using a load-balancing technique for the cloud computing environment. Concurrency and Computation: Practice and Experience, 30(12), e4368.
[14] Tukkoji, C., & Seetharam, K. (2020). Memory constraint parallelised resource allocation and optimal scheduling using oppositional GWO for handling big data in cloud environment. International Journal of Cloud Computing, 9(4), 432-452.
[15] Gong, D.W., Jing, S., & Miao, Z. (2018). A set-based genetic algorithm for interval many-objective optimization problems. IEEE Transactions on Evolutionary Computation, 22(99), 47-60.
[16] Braun, T. D., Siegel, H. J., Beck, N., Bölöni, L. L., Maheswaran, M., Reuther, A. I., ... & Freund, R. F. (2001). A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems. Journal of Parallel and Distributed computing, 61(6), 810-837.
[17] Panda, S. K., & Khilar, P. M. (2013). MSSA: A M-Level Sufferage-Based Scheduling Algorithm in Grid Environment. In Distributed Computing and Internet Technology: 9th International Conference, ICDCIT 2013, Bhubaneswar, India, February 5-8, 2013. Proceedings 9 (pp. 410-419). Springer Berlin Heidelberg.