Research on Multi-Objective Optimization Method of Edge Cloud Computing Virtual Machine Placement

Lingpeng Shi¹,a, Shida Lu²,b, Tianbo Feng³,c, Xiumin Zhao⁴,d, Xiaolu Chen⁵,e, Haoyang Cui⁶,f*
¹,²,³,⁴,⁵State Grid Shanghai Electric Power Co. Ltd. Information & Telecommunication Company, Shanghai, China
⁶Shanghai University of Electric Power, School of electronic and information engineering, Shanghai, China
aLingpeng_Shi@163.com, bShida_Lu@126.com, cTianbo_Feng@163.com
dXiumin_Zhao@163.com, eXiaolu_Chen@163.com, fcuihy@shiep.edu.cn

Abstract—With the emergence of demand for massive computing tasks in the edge cloud of the city, the disordered computing in the edge cloud leads to the high energy consumption of CPU computing and the problem of too long time delay caused by the blockage of computing tasks. This has become the first technical difficulty to solve in the construction of edge cloud. The virtual machine placement method is optimized by computing energy consumption and computing delay. First, the LRR physical host screening model is constructed to detect the status of the physical hosts and form a list of migrated physical hosts; secondly, the MMT time scale model is constructed to generate the list of migrated virtual machines; finally, the GA algorithm is used to place the virtual machines. The simulation results show that this algorithm can reduce the CPU energy consumption of the edge cloud center by 20.21% and the time delay by 16.11%. This optimization method has a good theoretical guidance effect on the construction of cloud computing.

1. INTRODUCTION

With the advent of the era of 5G, the Internet of Things, and the gradual increase in cloud computing applications, the pressure on data processing for cloud computing centers is becoming more and more intense [1]. If the resource allocation is unreasonable, in addition to the waste of resources, it will also bring greater computing pressure to the cloud computing center [2]. Therefore, it is of great significance to realize the multi-objective optimization of Virtual Machine Placement (VMP).

Virtual machine placement is an important step in virtual machine migration. Effective optimization of VMP can improve the utilization of physical resources, reduce energy consumption and achieve load balancing. From the perspective of the purpose of the virtual machine placement problem, there are currently three main types: energy consumption optimization technology for VMP [3]; VMP technology in the special architecture of the cloud data center; and the design of the placement method.

In recent years, the literature on the placement of virtual machines has focused on the following two aspects: changes in energy consumption models and improvements in intelligent algorithms. In terms of energy consumption model, it is divided into two types: multi-objective and single-objective [4]. In the early articles, when considering the energy consumption model of the data center, it was relatively single, and the typical goals were: energy consumption optimization, performance optimization, and resource
allocation optimization [5]. In order to optimize the data center from the aspects of data center load balancing, reducing energy consumption, and improving resource utilization, literature proposed a series of intelligent algorithms based on the combination of multi-objective optimization problems and intelligent algorithms. Multi-objective solving algorithm. Literature [6] uses pheromone positive feedback mechanism and heuristic search to add ant colony algorithm to solve, realizes the integration strategy of virtual machine placement in cloud environment, and obtains results in cloud data center load balancing, reducing energy consumption, and improving resource utilization. Literature [7] first takes minimization of energy consumption and maximization of resource utilization as the optimization goals. By imitating the pheromone mechanism, the classical bat algorithm is discretized and improved, and the improved discrete bat algorithm is used to solve the model. Literature [8] sets the goal as the waste of physical resources in the data center and the time delay of the data center, and optimizes the two goals at the same time through an improved dual fitness genetic algorithm. Literature [9] established an optimization model to minimize energy consumption and resource waste, using Pareto dominance concept combined with ant colony algorithm to solve the problem. Literature [10] uses CPU, MEM and hard disk occupancy rates as evaluation indicators to describe the load of the system, and uses linearized evaluation to calculate the resource utilization of the host to maximize resource utilization. Literature [11] designed a multi-objective discrete differential evolution of network-aware virtual machine placement algorithm. Under the framework of differential evolution, based on the network-wide communication cost, maximum link utilization, hardware resource constraint violations, and link capacity constraints Violation degree comprehensively evaluates individuals and selects the best.

However, genetic algorithm has the inherent deficiency of premature convergence, and the current improvement schemes based on genetic algorithm pursue high intelligence, precision and reckless to avoid premature convergence, which is very easy to cause high operation energy consumption. In this paper, MMT and LRR algorithms are used to optimize the genetic strategy and improve the virtual machine placement process, which can not only avoid the problems of low efficiency and premature caused by the influence of decision coding, but also significantly reduce energy consumption and improve the sense of user experience.

2. Model Building

2.1. The workflow of GA-VMP
The workflow specifically includes the following 4 steps: Step1 Complete physical host status detection based on the robust partial regression LRR method to form a list of candidate physical hosts for migration; Step2 Select the MMT algorithm based on the minimum migration time to complete the virtual machine selection and form a list of candidate virtual machines for migration; Step3 is based on the GA-VMP method to complete the virtual machine relocation; Step4 Repeat Step1 ~ Step3, set a cycle, and it will end when the time period is reached.

Figure 1. workflow of the GA-VMP
Figure 1 shows the workflow of the GA-VMP virtual machine placement strategy during the entire virtual machine migration process.

2.2 GA-VMP related terms

1) Cloud data center energy consumption model

The cloud data center is mainly composed of a large number of physical hosts, so its energy consumption is mainly composed of the energy consumption of all components of the physical host. Literature [12-15] believes that the energy consumed by a physical server (dual-core CPU, 4 memory, a disk, 2 PCI slots, a motherboard, etc.) is approximately: CPU accounts for 41%, memory accounts for 18%, and disk accounts for about 41%. 7%, PCI slots accounted for 22%, and motherboards accounted for 12%. According to this idea, this paper designs the energy consumption of a single physical server as:

\begin{equation}
E(U_{cpu})=E_{idle}+(E_{max}-E_{idle})U_{cpu}
\end{equation}

\begin{equation}
U_{cpu}(t)=\sum_{i=1}^{M}\sum_{c=1}^{PE} r_{j}(t)\frac{mips_{i,c}}{MIPS_{i,c}}
\end{equation}

\begin{equation}
E(U_{mem})=E_{idle}+(E_{max}-E_{idle})U_{mem}
\end{equation}

\begin{equation}
E(U_{disk})=E_{idle}+(E_{max}-E_{idle})U_{disk}
\end{equation}

\begin{equation}
E(U_{bw})=E_{idle}+(E_{max}-E_{idle})U_{bw}
\end{equation}

\begin{equation}
E_{total}=E(U_{cpu})+E(U_{mem})+E(U_{disk})+E(U_{bw})
\end{equation}

Among them, \( r_{j}(t) \) represents the index collection of virtual machines \( M \) allocated to the physical host. \( U_{cpu}(t) \), \( U_{mem}(t) \), \( U_{disk}(t) \), \( U_{bw}(t) \) indicates the CPU usage, memory usage, disk usage, and network bandwidth usage of the physical host at time. 0% \( \leq U_{cpu}(t) \), \( U_{mem}(t) \), \( U_{disk}(t) \), \( U_{bw}(t) \) \( \leq 100\% \). Indicates the energy consumption of the physical host CPU, memory, disk space, and network bandwidth when it is idle, that is, when \( U_{cpu}(t)=0\% \), \( U_{cpu}(t)=0\% \), \( U_{mem}(t)=0\% \), \( U_{disk}(t)=0\% \), \( U_{bw}(t)=0\% \). Indicates the energy consumption of the physical host at full load, that is, when \( U_{cpu}(t)=100\% \), \( U_{mem}(t)=100\% \), \( U_{disk}(t)=100\% \), \( U_{bw}(t)=100\% \).

\( \text{Mips}_{i,c} \) is the mips request situation of the \( c \) processing element of the \( i \) virtual machine. Is the overall MIPS computing power of the \( c \)-th processing element of the \( j \)-th physical host.

The number of MIPS requested by a virtual machine changes as the application changes, so the resource usage rate of the physical host should also change as the application changes. Therefore, the energy consumption of the physical server must be counted within a certain period of time. So according to formula (1) can be evolved into formula (7):

\begin{equation}
E_{total}(t)=E(U_{cpu}(t))+E(U_{mem}(t))+E(U_{disk}(t))+E(U_{bw}(t))
\end{equation}

In this way, the total energy consumption of the \( j \)-th physical host in the time period can be calculated according to formula (8):

\begin{equation}
E_{j}=\int_{t_0}^{t_1} E_{total}(t) \, dt
\end{equation}

The energy consumption of the entire cloud data center is formula (9), where \( M \) is the number of physical hosts.

\begin{equation}
E=\sum_{j=1}^{M} E_{j}
\end{equation}

2) SLA violation Time per Active Host (SLATAH).

When a cloud client submits a job to the cloud computing platform, the lack of resources will cause SLA violations. In the process of virtual machine allocation, an important performance indicator is the SLA time per host of each physical host. It reflects the online time of the physical host with high quality of service.

\begin{equation}
\text{SLATAH}=\frac{1}{M} \sum_{i=1}^{M} \frac{T_{i,j}}{T_{i,j}}
\end{equation}

The number of hosts and virtual machines in the cloud data center are denoted by \( M \) and \( N \).
respectively, where $i$ is the time when the CPU utilization of the physical host reaches 100%. $T_{aj}$ is the time that the physical host is online and active.

3) Performance Degradation due to Migrations (PDM) after virtual machine migration.

\[
PDM = \frac{1}{N} \sum_{j=1}^{N} \frac{C_{d_j}}{C_{r_j}}
\]

Where $C_{d_j}$ is the estimated value of performance degradation caused by virtual machine migration, and $C_{r_j}$ is the total CPU MIPS computing power during the entire time period during which the virtual machine is requested.

4) The combined index (ESV) of energy and SLA violations.

The SLA violation rate is calculated as formula (12):

\[
\text{SLA violation} = \text{SLATAH} \times \text{PDM}
\]

The indicators of overall energy consumption balance are:

\[
\text{ESV} = E \times \text{SLA violation}
\]

3. GA-VMP’s Virtual Machine Placement Algorithm

3.1. The specific process of GA-VMP virtual machine placement

Suppose there is a collection of $N$ virtual machines

\[
\{\text{VM}_i(p_{ei}, m_{ipsi}, t_{si}, d_{i}) | i=1, \ldots, N\}
\]

VM physical hosts to be reassigned to the cloud data center during the virtual machine placement phase \(\{\text{M}_j(PE_j, MIPS_j) | j=1, \ldots, M\}\) run on. Each virtual machine requests $p_{ei}$ processing elements, the request includes $m_{ipsi}$ computing requirements, the start execution time of VM is $t_{si}$, and the completion time is $t_{si}+d_{i}$. The migration operation was not interrupted or performed within $d_{i}$ a period of time. This article does not put the migration factor only on the single dimension of the CPU usage of the physical host, but expands the dimensions of physical resources, including memory size and disk space size, Network bandwidth, etc.

Assuming that each physical host can accommodate any type of virtual machine, and the energy consumption model of the physical host has a linear relationship with the utilization rate of various physical resources, then the goal of the virtual machine placement is to minimize the energy consumption of the cloud data center and make the most of it. It is possible to complete the execution of more virtual machines, so the objective function is:

\[
G = \min \sum_{j=1}^{M} E_j
\]

\[
E_j = \int_{t_{0}}^{t_{1}} E_{\text{total}}(t) \, dt, \quad E = \sum_{j=1}^{M} E_j
\]

As mentioned earlier, virtual machine placement is to complete the mapping of virtual machines to physical hosts, which is a multi-objective optimization problem. Therefore, GA-VMP introduces a multi-objective genetic algorithm to solve the virtual machine placement problem.

The specific meaning of the genetic algorithm will not be described in detail due to the space relationship. The specific steps of the GA-VMP solution are as follows:

- The chromosome encoding creates an initial population for $s$ chromosomes, and $s$ is the size of the population, which realizes the mapping from the virtual machine number to the chromosome.
- Calculation of fitness fitness function Fitness is converted from the objective function. Under the conditions of a given population, the evolution value of each chromosome is calculated.
- The calculation of the new family population completes the creation of a new family population by performing the following steps and completes the update.
- The selection stage is based on the evolution value of the chromosome, and an individual with two parents is selected from the current population.
- The crossover stage uses the crossover probability to create a new offspring by modifying the chromosomes of both parents.
- In the mutation stage, the mutation will be completed in a certain position of the chromosome through the mutation probability.
The acceptance phase in the current situation, a new child will be produced from the next generation. The replacement phase completes the replacement operation by assigning the current generation of individuals to the next generation. If the termination condition is met, the algorithm will stop execution and return an individual with the highest evolution value. Otherwise, the algorithm returns to the fitness calculation phase to continue the cycle. Figure 2 shows the flow of GA-VMP.

![Figure 2. the flow of GA-VMP](image)

The surface is the pseudo code of the virtual machine placement process of GA-VMP genetic algorithm. The code can be implemented in Java language with a small amount of modification.

3.2. Implementation of GA-VMP

GA-VMP is a genetic algorithm, which is mainly used to solve the multi-objective optimization problem of virtual machine placement. The objective function of GA-VMP is to minimize the energy consumption of the cloud data center. This article uses a hierarchical structure to reflect the chromosome relationship between individuals. This hierarchical structure has 3 levels:

- Level 1 consists of initialized individuals.
- Level 2 consists of a series of nodes, representing a group of virtual machines in the cloud data center.
- Level 3 consists of a series of nodes, representing a group of physical hosts running on physical hosts.

Through this hierarchical structure, the placement of virtual machines to physical hosts is reflected, and the fitness function completes the calculation of the evolution value of each chromosome. This process is performed in parallel for virtual machine placement, which can reduce the calculation time for virtual machine migration. For the pseudo code of the fitness function, see Algorithm 1.

**Algorithm 1** Calculation of the fitness value of each chromosome

**Input:** Chromosome  
**Output:** Fitness of the chromosome

1. Power Of Datacenter = 0  
2. For each host ∈ collection of hosts do  
3. utilization Mips=host. get Utilization Of CPU  
4. utilization Mem = host. get Utilization Of Mem  
5. utilization Disk = host. Get Utilization Of Disk  
6. Utilization B w = host. Get Utilization Of B w  
7. Power Of Mips = get Power (host, utilization Mips)  
8. Power Of Mem = get Power (host, utilization Mem)  
9. Power Of Disk = get Power (host, utilization Disk)
10 Power Of Bw = get Power (host, utilization Bw)
11 Power Of Host = power Of Mips + power Of Mem + power Of Disk + power Of Bw
12 Power Of Datacenter = power Of Datacenter + power Of Host
13 End For
14 Evaluation value(chromosome) = 1.0/power Of Datacenter

The GA-VMP virtual machine placement method in this article combines the existing robust local regression LRR physical host state detection method and the minimum migration time MMT virtual machine selection strategy in Cloudsim, forming a new type of virtual machine migration model called LRR-MMT-GA.

4. EXPERIMENTAL RESULTS AND ANALYSIS
This article analyzes the performance of the GA-VMP algorithm proposed in this article from the aspects of virtual machine placement energy consumption and delay, and compares it with the following three algorithms:
1. Optimization Sequence Algorithm (Optimization Sequence Algorithm, OSA): A virtual machine sequence placement algorithm optimized by the idea of interval bubbling.
2. Simulated Annealing (Simulated Annealing, SA): A Monte-Carlo iterative stochastic optimization algorithm based on the similarity between the annealing process of solid substances in physics and general combinatorial optimization problems.
3. Particle Swarm Optimization (PSO): A self-adaptive multi-strategy heuristic algorithm that uses multiple particles to form a population for iterative calculation to find the optimal solution.

The simulation in this paper is implemented in Cloudsim and MATLAB R2019b simulation platform, and genetic algorithm is introduced as the benchmark algorithm. The initial size of the population is 20 and the number of iterations is 200. This algorithm is unique in terms of optimal configuration. The experiment used 100 physical nodes and 200 virtual machines. The CPU resources requested by the virtual machines were uniformly set to (0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4) to simulate the different requests of the virtual machines. At the same time, the memory is also set to (0.25,0.5,1,1.5,2,2.5,3,4) the power consumed by the physical node when the CPU utilization is 0% and 100% is 175W and 250W respectively.

![Figure 3. CPU energy consumption](image-url)
As shown in Figure 3, under the same virtual machine load, the comparison of CPU energy consumption occurs under four different algorithms. It can be seen that with the increase in the number of virtual machines placed, the CPU energy consumption of the four placement algorithms is on the rise. The OSA algorithm has the highest CPU energy consumption, SAA and PSO are relatively low, and GA-VMP has the lowest CPU energy consumption. Compared with the PSO algorithm, the CPU energy consumption of the algorithm in this paper is reduced by 20.21% on average. The proportion of CPU memory consumption fluctuates with the increase in the number of virtual machines. As shown in Figure 4, there is no significant increase in memory consumption, indicating that the server has a strong virtual machine placement capacity. In addition, the memory consumption rate of the GA-VMP virtual machine placement algorithm is slightly lower than the other three algorithms, and it has stronger placement capabilities.

Taking the placement of 20 virtual machines as an example, the particle swarm optimization algorithm with the best delay performance is selected and compared with the algorithm in this paper. The delay effect is shown in Figure 5, and the average delay is reduced by 16.11%. Since the particle swarm optimization algorithm is prone to local optimal solutions, it needs to perform secondary placement and migration according to the allocation rate of virtual machines, which consumes a certain
amount of time. The GA-VMP virtual machine placement algorithm mentioned in this article avoids localization to a certain extent through genetic iteration. The frequent appearance of the optimal solution can effectively improve the delay performance and enhance the user experience. In addition, the consumption rate of CPU and memory has been reduced to varying degrees.

5. CONCLUSION
This paper proposes a GA-VMP virtual machine placement method, which uses MMT and LRR algorithms to optimize genetic strategies. While improving the virtual machine placement process, it can reduce the energy consumption of virtual machine placement and enhance the user experience. Based on the future research of this article, the following two aspects can be considered: (1) Consider combining genetic algorithm with particle swarm algorithm, ant colony algorithm and other methods to design a method that can better jump out of the local optimal solution. (2) The granularity of the target virtual machine can be refined and analyzed to obtain the impact of different granularity on the placement performance of the virtual machine.

ACKNOWLEDGMENT
This study was supported by the Science and Technology Project of State Grid Shanghai Electric Power Co., Ltd (52090F20007L). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

REFERENCES
[1] SHANG P J, WANG R J, WANG Q D. A novel power management for CMP systems in data-intensive environment [J]. IEEE Transactions on Computers, 2016, 65(6): 1816-1830.
[2] MAO Y Y, ZHANG J, LETAIIF K B. A Lyapunov Optimization Approach for Green Cellular Networks with Hybrid Energy Supplies [J]. IEEE Journal on Selected Areas in Communications, 2015, 33(12): 2463-2477.
[3] BELOGLAZOV A, BUYYA R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers [J]. Concurrency and Computation: Practice and Experience, 2012, 24(13): 1397-1420.
[4] GU Y H, YI P. Method of Service Function Chain Adjustment Based on Multi-VM Live Migration. Journal of Chinese Computer Systems, 2017, 38(5): 1022-1027.
[5] Wang JV, Cheng CT, Tse CK. A power and thermal-aware virtual machine allocation mechanism for cloud data centers [C]//IEEE International Conference on Communication Workshop. IEEE, 2015:2850-2855
[6] Zhang Congyue, Fu Xiong, Qiao Lei. Research on virtual machine placement based on multi-objective optimization in cloud computing environment [J]. Computer Applications and Software, 2021, 38(03): 32-38.
[7] Lin Kaiqing, Li Zhihua, Guo Shujie, Li Shuangli. Multi-objective optimization virtual machine placement algorithm in cloud environment[J]. Computer Applications, 2019, 39(12): 3597-3603.
[8] Huang Zhaonian, Li Haishan, Kui Qiang. Research on virtual machine placement based on improved algorithm in cloud computing [A]. Electronic Technology Academic Committee of China Shipbuilding Engineering Society. Paper of Equipment Technology Development Forum 2017 of Electronic Technology Academic Committee of China Shipbuilding Engineering Society[C]. Electronic Technology Academic Committee of China Shipbuilding Engineering Society: China Shipbuilding Engineering Society, 2017:5.
[9] F. Farahnakian et al., "Using Ant Colony System to Consolidate VMs for Green Cloud Computing," in IEEE Transactions on Services Computing, vol. 8, no. 2, pp. 187-198, 1 March-April 2015, doi: 10.1109/TSC.2014.2382555.
[10] Yan Jianen, Zhang Hongli, Xu Haiyan. Virtual machine placement method based on load optimization [J]. Intelligent Computers and Applications, 2019, 9(06): 178-180+183.

[11] Zang Weifei, Lan Julong, Hu Yuxiang, Chen Wentao. Network-aware virtual machine placement algorithm based on multi-objective discrete differential evolution[J]. Computer Engineering, 2019, 45(06): 96-102.

[12] Li Shuangli, Li Zhihua, Yu Xinrong. Virtual machine placement method based on multi-objective optimization [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2020, 32(03): 356-367.

[13] Zhou Zhen, Wang Hui, Li Junfeng. Virtual machine placement strategy based on family genetic algorithm[J]. Computer Engineering and Design, 2021, 42(02): 482-488.

[14] Xu Shengchao. Using genetic algorithm to complete the optimization of virtual machine placement strategy[J]. Computer and Modernization, 2020(12): 25-31+42.

[15] Liu Dan, Sui Xin, Li Li. Virtual machine placement energy-saving optimization algorithm based on cloud computing [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2015, 38(06): 150-153.