

Research on Resource Allocation Algorithm Based on Energy Optimization and Immune Algorithm in Cloud Computing Environment

Lu Dai¹, Ji OuYang¹ and Caichang Ding^{2,*}

¹Computer College, Dongguan University of Technology, Dongguan, 523808, China

²School of Computer Science, Yangtze University, Jingzhou, 434023, China

Abstract: This paper will explore the key factors affecting the resource allocation performance; reveal the essential relationship among CPU dynamic power consumption, resources computing capacity, resources supply voltage strategy, the amount of resources, supply voltage, and the frequency of CPU; and solve the huge energy consumption problem and the poor performance of load balancing in the process of resource allocation. The issues include: design and analysis of the key technology of resource allocation in cloud computing environment; combined with time span, load balancing and energy optimization, design of multi-objective optimization model for resource allocation and heuristic resource allocation algorithm. This research has important significance for further study on resource allocation and energy optimization in cloud computing environment.

Keywords: Cloud computing, energy optimization, immune algorithm, resource allocation.

1. INTRODUCTION

Computing technology is developing towards generalization, large scale data centers have to be built for satisfying the society's increasing requirement to cloud computing however its operation has to be maintained with a great deal of energy [1]. It becomes an urgent issue that how to increase resource utilization rate and reduce energy consumption under the premise of ensuring high performance of cloud computing system and providing generalized service.

Resource allocation is a key technology of cloud computing. At present, most of the studies on resource allocation mainly focus on improving the overall performance by balancing the load of data center, but they didn't give full consideration to how to reduce energy consumption. The better solution for energy saving in the process of seeking resource allocation can not only save electric energy, but also reduce heat production of server so as to cut heat dissipation cost, improve system stability and extend service of hardware [2, 3]. By estimation, China Data Center consumed 7 million KWh power in 2011 which takes up 5% of China's total power consumption. When 1 KWh power is consumed, 0.555kg CO₂ will be produced, which will result in unprecedented huge pressure on energy supply and environment [1].

It can be seen from above analysis that this paper not only introduces the latest results of IA into cloud computing, but also organically integrates resource allocation and optimization of energy consumption and provides new tools and

methods for studies in the field of cloud computing. Comparing with existing analysis models and measures, the solutions proposed by this paper have distinct features and development prospects.

2. THE OBJECTIVE OF STUDY

2.1. Status Quo & Analysis of Chinese and Overseas Studies

The study of resource allocation under the environment of cloud computing mainly includes three aspects: the first is realizing the balance of resource requirement and allocation with modes of market mechanism; the second is how to realize the purpose of saving energy by energy-efficient operation of data center and increase of resource utilization and; the third is how to realize load balance of system resources and minimization of task execution time.

By studying the relationship between server's energy consumption and resource utilization rate, literature [4] proposed a kind of energy saving algorithm on the basis of initial allocation of cloud computing virtual machine. Dhiman G. *et al.* scheduled virtual host according to competition to memory, but they didn't take competition to other resources into account [5]. Li Bo *et al.* carried out allocation of virtual resources according to the thought of maximizing server utilization rate, but it resulted in excessive migration operations in the process [6]. Literature [7] realized efficient virtual host scheduling, but the solving of NP problem and excessive use of sorting algorithm resulted in over high complexity. Li Qiang *et al.* designed a multi-objective optimizing genetic algorithm for virtual machine placement that is based

*Address correspondence to this author at the School of Computer Science, Yangtze University, Jingzhou, 434023, China; E-mail: dail@dgut.edu.cn

on goal constraint at application service level [8]. Khan S. *et al.* designed method of game theory for optimizing energy consumption and response time of system [9]. Though Chinese and overseas scholars carried out effective studies in above fields, the problem of resource allocation under cloud computing environment are still facing the challenges as below:

(1) As a key technology of cloud computing, the problem of resource allocation is still in the phase of study. With respect to key technologies that have an influence on resource allocation, mutual coordination and interdependence between resource allocation and task scheduling, there is still no unified industrial standard formed yet.

(2) The energy consumption in the resource allocation process under cloud computing environment is influenced by many factors, so it is crucial to establish a mathematical model and carry out intensive study on its performance for essentially uncovering the relationship among the dynamic energy consumption of CPU and resource computing capacity, the supply voltage strategy of resources, the amount of resources, supply voltage and CPU frequencies as well as key factors that have influence on system performance.

2.2. The Content of Study

This paper aims at systematically and intensively studying the problem of resource allocation which is based on energy consumption optimization and immune algorithm under cloud computing environment. Main objectives of study include:

(1) Analyzing key technologies of resource allocation under cloud computing environment, coordination and interdependence between resource allocation and task scheduling, multi-objective constraint conditions of time span and load balancing for laying theoretical basis for establishing study of integral frame of cloud computing resource allocation.

(2) Uncovering and quantitatively expressing relation among the dynamic energy consumption of CPU and resource computing capacity, the supply voltage strategy of resources, the amount of resources, supply voltage and CPU frequencies for providing new clues for establishing theoretical research and application of energy consumption optimization model of cloud computing resource allocation.

(3) Verifying the effectiveness and the practicability of key technologies proposed by the mode of integrating theory, imitation and test for providing scientific basis for resource allocation algorithm which is based on cloud computing platform and integrates intelligent optimization and energy consumption optimization.

3. ANALYSIS AND MODELING OF KEY TECHNOLOGIES

3.1. Method of Study and Technical Route

This subject plans to carry out the study with the method of integrating theoretical analysis, simulation verification

and practical arrangement. Three study groups of theoretical analysis, simulation test and test data analysis are set as a whole. The theoretical analysis group is responsible for establishing performance analysis model for cloud computing resource allocation, designing relevant algorithm and optimizing heuristic allocation algorithm. The simulation test group is responsible for realizing resource allocation algorithm, designing test methods, testing platform and realizing software, implementing simulation test, analyzing and maintaining simulation data. The test data analysis group is mainly responsible for designing and arranging test scheme, collecting, analyzing, accounting, and sorting test data as well as discussing existing problems and possible method of improvement together with simulation test group. See Fig. (1) for the technical route that this study plans to take and the detailed content is as follows.

3.2. Stablising the Optimization Analysis Model

Integrating the analysis and description of section 2.2, the following technical route and study methods will be taken for intensively developing the optimization model for energy consumption of cloud computing resource allocation:

(1) Establishing the energy consumption optimization model

The dynamic energy consumption is caused by electrical charging and discharging of capacitances of each node and its basic expression is: $P = A \times C \times v^2 \times f$, of which A represents the average circuit hopping times within a cycle and it is called flipping frequency and related to charging and discharging frequency;

$$E_{ijl} = r \times [(v_k(i))_j]^2 \times (f_k(i))_j \times f \times ETC'[i, j, k] \quad (1)$$

Where, $r = A \times C$ represents the intrinsic property of given resource, $v_k(i)_j$ represents the supply voltage of the resource r_i that executes task t_j when the supply voltage strategy is s^i and DVS level is k ; $f_k(i)_j$ represents the operating frequency at this time; $ETC'[i, j, k]$ represents the k th component of $ETC'[i, j]$. Substituting the k th component of formula (2), then formula (3) is translated into:

$$E_{ijl} = r \times f \times [(v_k(i))_j]^2 \times ETC(i, j) \quad (2)$$

The energy consumption that the resource execute allocated tasks can be preliminarily described as:

$$E_i = r \times f \times \left\{ \sum_{j \in T(i), k \in L(j)} [(v_k(i))_j]^2 \times ETC(i, j) + [v_{\min}(i)]^2 \times f_{\min}(i) \times Idle_i \right\} \quad (3)$$

Where, $T(i)$ represents the set of tasks allocated to the resource r_i , $L(j)$ represents the set of DVS levels when the resource r_i executes tasks, $v_{\min}(i)$ and $f_{\min}(i)$ respectively represent the voltage and the frequency when the resource r_i is switched to hibernation in idle time and $Idle_i$ represents

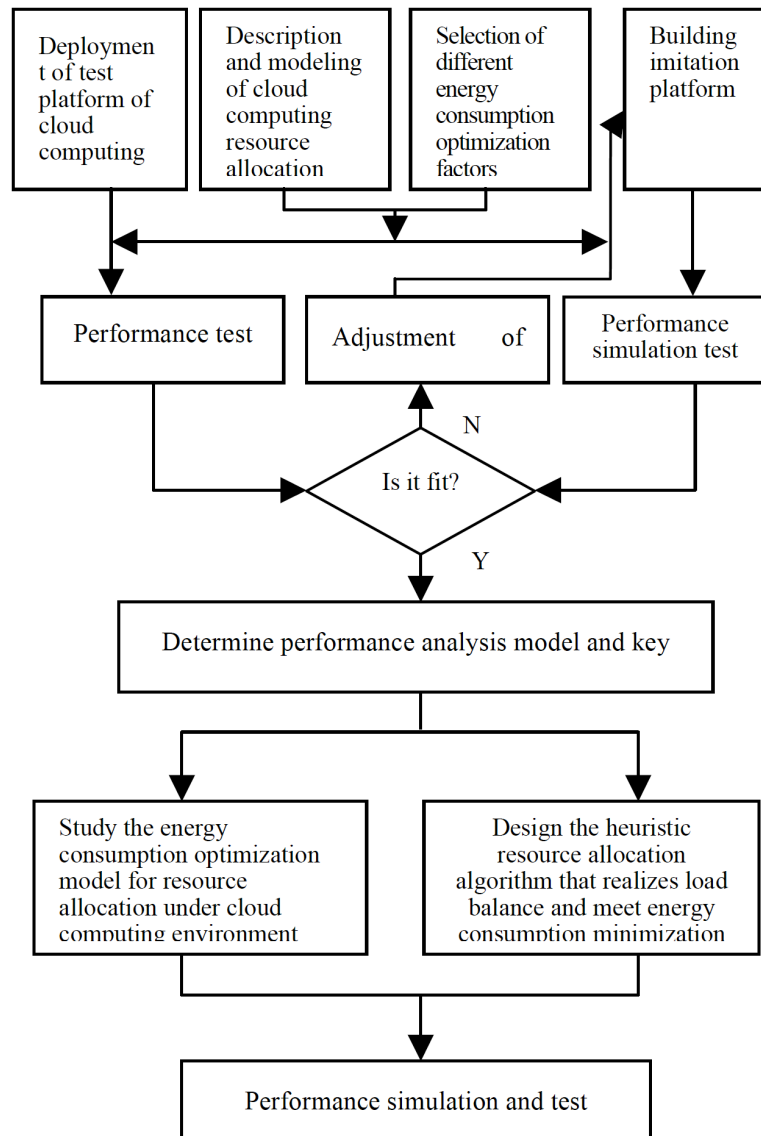


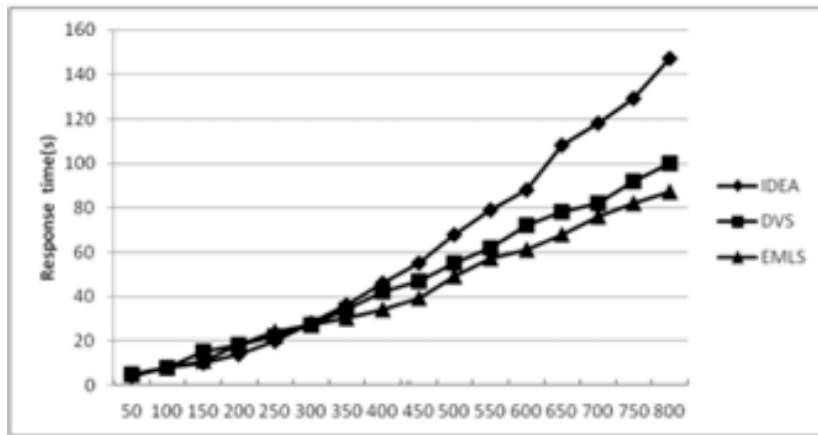
Fig. (1). The technical route to be taken by this paper.

the idle time of resource. The detailed expressions of $v_{\min}(i)$, $f_{\min}(i)$ and $Idle_i$ will be monitored and computed by designing practical and feasible methods in the process of paper implementation.

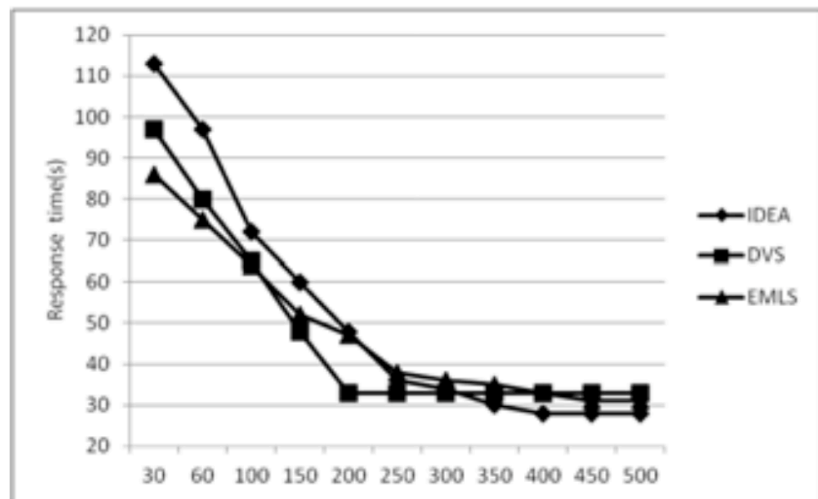
4. SIMULATION EXPERIMENT

In this section, we analyze the performance of our model based on the experimental results. For testing the effectiveness and superiority of the resource allocation model in cloud computing, the existing resource allocation approaches such as the IDEA [10], EMLS [11], and DVS [12] were used. Firstly, we compare the response time of the three resource allocation algorithms in Fig. (2). The Y-axis represents response time, and the X-axis denotes generations and number of tasks, respectively. Secondly, we compare the makespan of the three resource allocation algorithms in Fig. (2). The Y-axis represents makespan, and the X-axis denotes generations and number of tasks, respectively. Finally, we compare the energy consumption of the three resource allo-

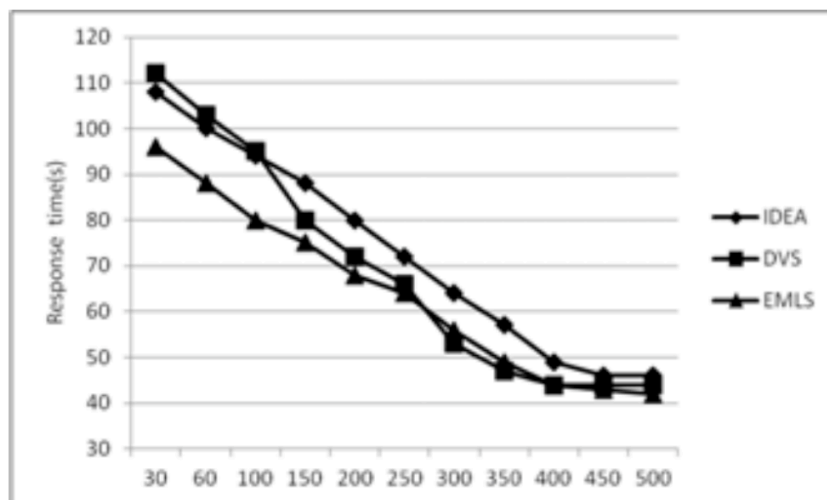
cation algorithms in Fig. (3). In Fig. (3), the Y-axis shows energy consumption, and the X-axis denotes lower utilization threshold. Response time is the amount of time taken between submission of a request and the first response that is produced. As shown in Fig. (2), DVS has the best response time performance compared to the other three algorithms. EMLS has relatively close response time, and the response time of IDEA increases significantly when the number of tasks increases in cloud computing environments. When the number of tasks is 400, the response time of ICGA is relatively minimal and its minimum value is close to 26.1 s. When the number of tasks is 600, the response time of ICGA is the smallest in most cases and its minimum value is close to 35.2 s. When the number of tasks is 800, the response time of ICGA is relatively minimal and its minimum value is close to 40.8 s. It is evident that ICSA is more efficient compared with the other three algorithms.



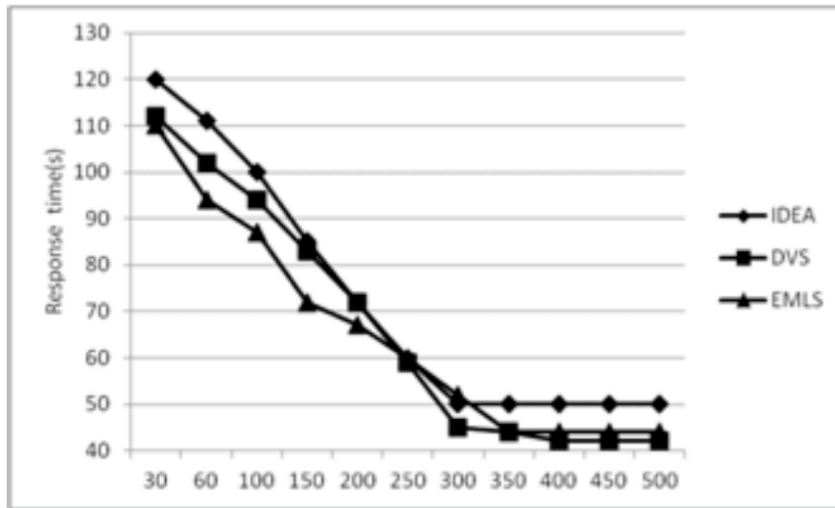
(a) Number of tasks



(b) Generations

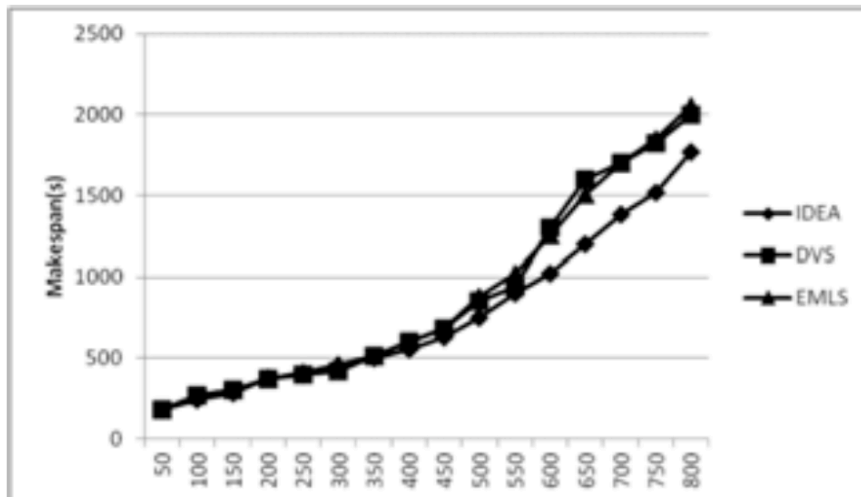


(c) Generations

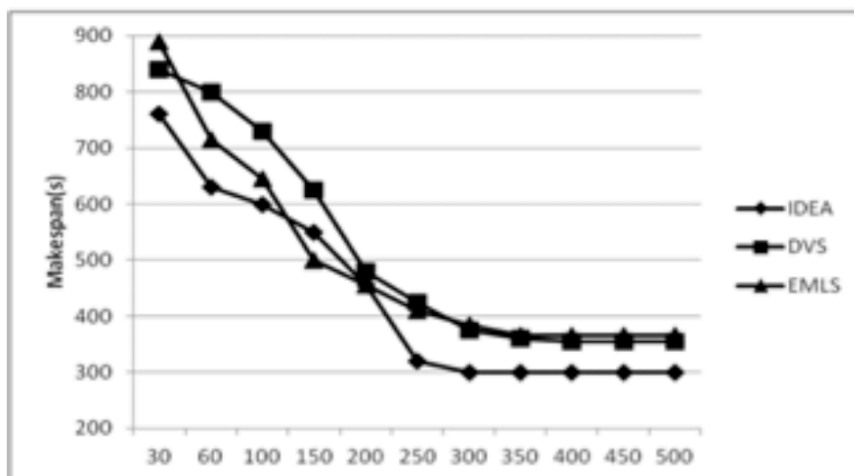


(d) Generations

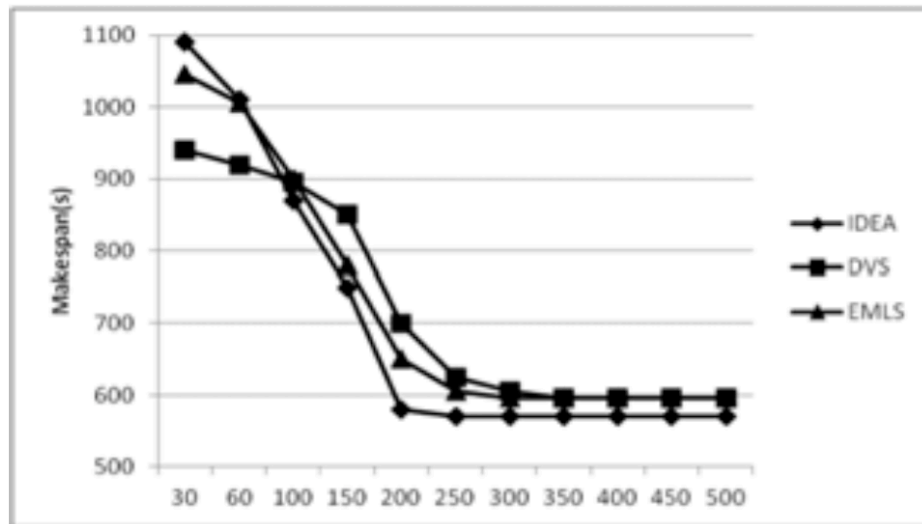
Fig. (2). Comparison of the response time of the four resource allocation algorithms. (a) Different tasks. (b) n= 400.(c) n= 600.(d) n= 800.



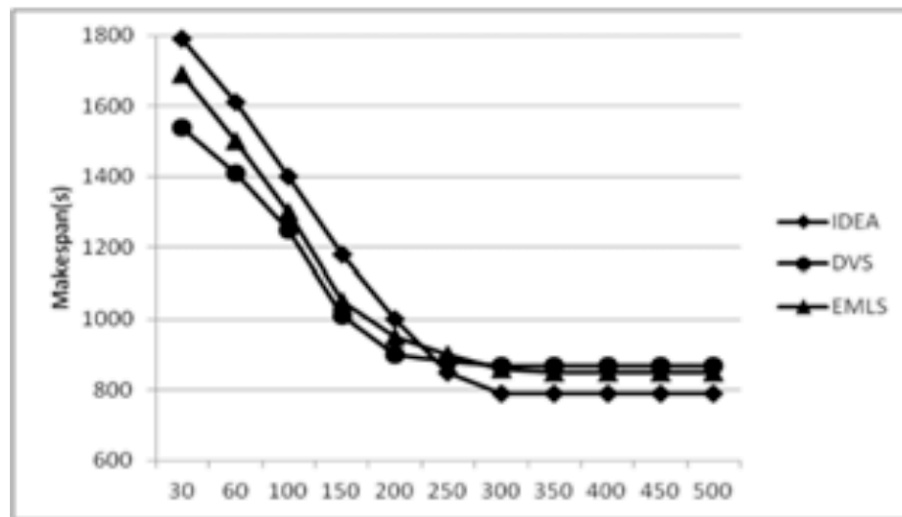
(a) Number of tasks



(b) Generations



(c) Generations



(d) Generations

Fig. (3). Comparison of the makespan of the four resource allocation algorithms. (a) Different tasks. (b) n= 400.(c) n= 600.(d) n= 800.

It is clearly evident from Fig. (3) that when the number of tasks in cloud computing environments is higher, the difference in makespan becomes more apparent. When the number of tasks is 400, the makespans of EMLS, IDEA, DVS are 245.8, 258.2, 291.6s, respectively. When the number of tasks is 600, the makespans of EMLS, IDEA, DVS are 493.2, 523.5, 576.8s, respectively. When the number of tasks is 800, the makespans of EMLS, IDEA, DVS are 762.5, 796.4, 835.2s, respectively. Thus, we can decrease the computation complexity of traditional resource allocation algorithms to increase the overall performance. Energy efficiency is one of the key technologies of resource allocation in cloud computing. When the lower utilization threshold increases, the energy consumption of the system is also rapidly reduced.

It can be observed that there is a significant difference among the three resource allocation algorithms. The DVS is well enhanced and balanced on exploration and exploitation and has better stability and scalability (Fig. 4).

CONCLUSION

On the basis of tracking the theories of cloud computing, the key technologies of resource allocation, the designs and optimization of energy consumption and the designs of immune algorithm for a long term, this paper thoroughly investigated the status quo of the research field of cloud computing and the requirements of practical engineering application and will carry out intensive study on core problems involved.

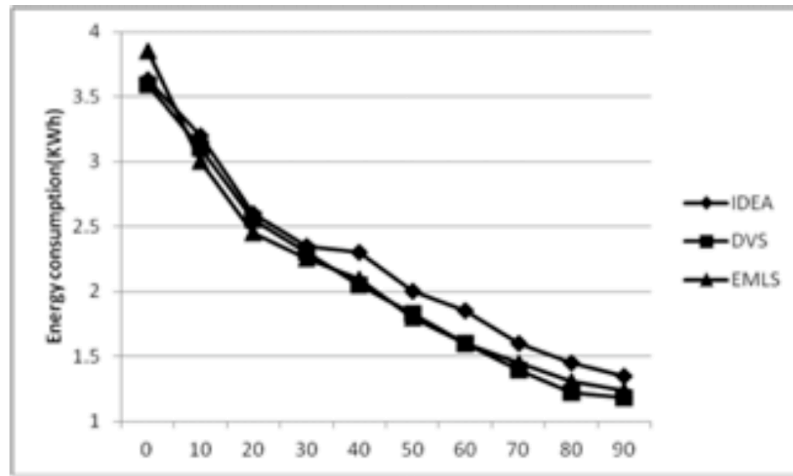


Fig. (4). Different values of the utilization thresholds.

As a challenging novel subject, it carries out study on the resource allocation under cloud computing environment by integrating immune algorithm and energy consumption design and optimization technology. The study is mainly based on the knowledge and technologies in the fields of cloud computing, electrical engineering and intelligent computing. There have been many basic theories and research achievements in such fields, which lays sound theoretical basis and provides referential technologies for the study to be carried out under this subject.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] C. Huang, C. Guan and H. Chen, "An adaptive resource management scheme in cloud computing", *Engineering Applications of Artificial Intelligence*, vol. 26, no.1, pp. 382-389, 2013.
 [2] A. Beloglazov, J. Abawajyb and R. Buyyaa, "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing", *Future Generation Computer Systems*, vol. 28, no. 5, pp. 755-768, 2012.
 [3] J. Ryan, B. R. Paul, "Energy efficient virtual machine allocation in the cloud: An analysis of cloud allocation policies", In: *Proceed-*

ings of the 2011 International Green Computing Conference and Workshops, 2011, pp.137-151.
 [4] C. Hsu , S. Chen and C. Lee, "Energy-aware task consolidation technique for cloud computing", In *Proceedings of the 3rd IEEE International Conference on Cloud Computing Technology and Science*, 2011, pp. 115-121.
 [5] S. Srikantaiah, A. Kansal, and F. Zhao, "Energy Aware Consolidation for Cloud Computing", In *Proceedings of the 2008 Conference on Power Aware Computing and Systems*, 2008, pp. 201-208.
 [6] A. Beloglazov, A.V. Vasilakos and V. Lesser, "Evolutionary stable resource pricing strategies", In *Proceeding of the ACM SIGCOMM*, 2009, pp. 734-739.
 [7] M. Cardoso, M. Korupolu and A. Singh, "Shares and utilities based power consolidation in virtualized server environments", In: *Proceedings of the IFIP/IEEE Integrated Network Management*, 2009, pp. 327-334.
 [8] S. Khan and I. Ahmad, "A cooperative game theoretical technique for joint optimization of energy consumption and response time in computational grids", *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 3, pp. 346-360, 2009.
 [9] T.V.T Duy, Y. Sato and Y. Inoguchi, "Performance evaluation of a green scheduling algorithm for energy savings in cloud computing", In: *Proceedings of the 2010 IEEE International Symposium on Parallel and Distributed Processing, Workshops and Phd Forum*, 2010, pp.1-8.
 [10] J-T. Tsai, J-C. Fang, and J-H. Chou, "Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm," *Computers & Operations Research*, vol. 40, no. 2, pp. 3045-3055, 2013.
 [11] Y. Kessaci, N. Melab, and E-G. Talbi, " A multi-start local search heuristic for an energy efficient VMs assignment on top of the OpenNebula cloud manager," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 1-20, 2013.
 [12] J. Kołodziej, S.U. Khan, L. Wang, M. Kisiel-Dorohinicki, and S.A. Madani, E Niewiadomska-Szynkiewicz, A.Y. Zomaya, "Security, energy, and performance-aware resource allocation mechanisms for computational grids," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 944-959, 2013.

Received: November 26, 2014

Revised: January 08, 2015

Accepted: January 21, 2015

© Dai et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.