2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)
GLA University, Mathura, Oct 26-28, 2017

An improved Harmony Search algorithm with Group technology model for scheduling workflows in cloud environment

Nidhi Chaudhary
P G Scholar
Department of Computer Science and Engg.
NITTTR, Chandigarh,
frmnidhichoudhary@gmail.com

Mrs. Mala Kalra
Assistant Professor
Department of Computer Science and Engg.
NITTTR, Chandigarh,
malakalra2004@yahoo.co.in

***Abstract*—Cloud computing is a latest innovation in technology in which provide services like infrastructure and applications to end-users as per pay-as-you-go model. It can provide virtualized services according to the requirements of end-users varying with time. The resources of the cloud are available everywhere and user can access these resources anytime. Cloud computing provides virtualization of physical machine into several virtual machines, which increases the resource availability. It provides high performance to the users, and profit to the service providers of cloud. Cloud resources are allocated dynamically on demand of users and shared by multiple numbers of users. The challenge is to optimally map the resources to users in order to satisfy users QoS constraints, known as task scheduling. Due to NP-complete nature of scheduling problem, exact solutions cannot be found in finite time. Therefore, many researchers apply heuristic or random search techniques as Particle swarm optimization (PSO), Ant colony optimization (ACO), Genetic algorithm (GA), Cuckoo Search (CS), Min-Min algorithm etc. to get solutions of such problems which are near optimal. We continuing our research in that direction too and we are proposing a workflow scheduling algorithm for cloud environment based upon Harmony Search algorithm merged with Group technology aiming to reduce makespan. We compared our harmony search based algorithm with Particle swarm optimization (PSO) and Genetic algorithm (GA) and Min-Min scheduling algorithm. The experimental results show that our proposed algorithm provides better results in terms of execution time.**

Keywords—Cloud Computing, Harmony search, Group technology, Workflow Scheduling, Min-Min

I. INTRODUCTION

Demand of computing resources has enlarged in past few years as per the advancement in computing technologies. To provide computing resources to everyone at anytime and anywhere is the main challenge for organizations. Hence, new solutions are offered by IT persons. Cloud computing is the latest among all the solutions provided recently. Main advantage of cloud computing is that it can be accessed as a utility service like water, electricity etc. This means individuals pay as per their usage of resources. Hence additional costs such as purchasing and maintenance of infrastructure are reduced. Efficient resource management has a large impact in minimizing the costs of operations and increasing the effectiveness of technology. Hardware virtualization technology is used for efficient management of resources [5]. Scheduling is important for proper management of resources. Through cloud scheduling methods, resources provided by cloud managers can be used effectively and user request can be served properly. Cloud scheduling can be divided into two categories. When all tasks are independent, tasks can assigns to the processors independently without any predefined order of execution. In case of dependent scheduling (workflows), scheduling is very complicated. A workflow defines the dependency between tasks that have to be executed on accessible belongings. Scheduling plays an important role in information processing environment. In past decade, many metaheuristic algorithms have been developed for cloud environment which were used to resolve workflow scheduling problem. Geem [8] proposed a novel metaheuristic technique, Harmony search method, which can be applied to solve various optimization problems. Harmony search is a latest population based metaheuristic technique based upon, improvisation of music players. Idea is to find out the better harmony memory based upon all the existing music harmonies and replacement of the worst harmony by New Harmony memory. Advantage of harmony search was, by considering all the existing solution vectors, new vectors are formed. Harmony search algorithm helps in finding better solutions. Hence Harmony search can be applied to solve workflow cloud scheduling problem to enhance the various QoS parameters of cloud.

II. RELATED WORK

Scheduling workflows in cloud while satisfying QoS constraints is one of the main challenging problems. From the literature survey, it has been found that various metaheuristic techniques have been used to solve scheduling problems for workflows in cloud environment. Previous research work in this domain used ACO, Genetic algorithm, PSO, Cuckoo search and modified versions of these algorithms. Fard et al.[1] proposed a framework and

978-1-5386-3004-4/17/$31.00 ©2017 IEEE

heuristic method for static scheduling with multiple goals at a time like makespan, cost and energy consumption. According to the optimization process objectives has been analyzed. Wang et al.[2] introduced a multi constrained scheduling method of multi-workflows (MQMW) in which the problem of multiple QoS requirements from different users at the same time was addressed. The method can fulfill the different QoS requirement of the different users and scheduling of multiple workflows can be done at a time. Pandey et al.[3] proposed a PSO based metaheuristic for scheduling workflows considering execution cost and data transmission cost as optimization criteria. PSO based proposed algorithm gives three times better makespan and less data transmission cost than best resource selection algorithm. Malnik et al. [7] proposed a polyrhythmic harmony search algorithm (PHS) to solve workflow scheduling problem for cloud which provided better convergence speed. A Hybrid Harmony Search algorithm was presented in [9] with local search method to minimize the execution time of job shop scheduling problem. Harmony search has been used to solve flow shop scheduling problem in [10] and for permutation flow shop problem in [17] and for load balancing problem in cloud environment [21]. Pashaki et al. [5] used Group Technology (GT) model with Genetic algorithm for consolidation problem in cloud computing and achieved better results as compared to standard genetic algorithm. Pashaki et al. [6] used Group Technology model with Cuckoo search algorithm for resource allocation problem in cloud computing and achieved better results. This motivated us to apply Harmony search algorithm merged with Group technology model to solve workflow scheduling problem in cloud environment which may reduce the makespan of workflow.

III. PROBLEM DEFINITION

Most of the e-scientific applications are nowadays articulated as workflows. We can represent a workflow by using a Directed Acyclic graph (DAG) G (T, E), where T gives the number of jobs in a workflow and dependencies between tasks are defined by E. To fulfill the user's constraints, we have to map every task T to certain virtual machine. From the literature survey it has been observed that the main target of scheduling algorithms is to lessen the makespan. As per the NP- Complete nature of workflow scheduling problem, researchers applied various metaheuristic techniques to find out a satisfactory solution. So the aim is to develop a better workflow scheduling algorithm which may reduce the makespan of workflow.

IV. PROPOSED WORK

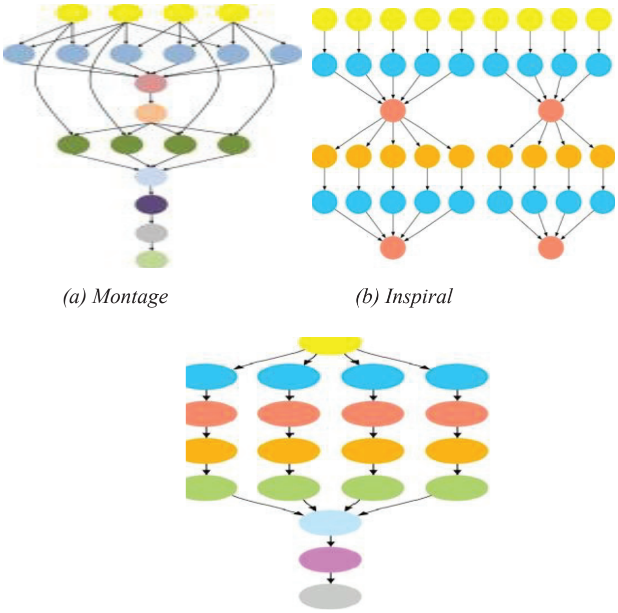
In this research work, we propose an algorithm based on Harmony search merged with group technology model (HSGT) for solving scheduling problem for workflows in cloud environment. Harmony search method merged with group technology model is an extended form of the harmony search method which gives better performance. We are considering execution time (makespan) as an objective function and trying to minimize the execution time of workflow. Makespan is the completion time of last task. For every virtual machine, we summed up the running time of every task which is allocated on that machine and the largest number is said to be fitness function of that harmony. Makespan is calculated by max $\{FT_i\}$ where FT_i is the finish time of i^{th} task. Makespan is considered as fitness function. Pseudo code for proposed Harmony search merged with group technology (HSGT) algorithm.

1. Initialize the algorithm parameters as follows: Harmony Memory Considering Rate (HMCR)=0.7, Pitch Adjustment Rate (PAR)=0.3, Maximum number of iteration (MaxItr) =100 as stopping criteria,
2. Size of Population=50.
3. Generate Initial population vector of harmony memory by using Min-Min scheduling algorithm.
4. Update the harmony memory vector by using Harmony memory considering rate (HMCR).
   i. If (HMCR<0.7)
      a. Copy the harmony vector from the original harmony memory.
      b. If (PAR<0.4) New Harmony vector is generated by considering the minimum vector from the existing harmony member.
5. Calculate penalty of the calculated solution.
6. If (Stopping criteria met)
   a. Generate the best harmony vector as solution vector.
   b. Else, Update solution vector by using HMCR and PAR.

V. EXPERIMENTAL SETUP AND RESULT ANALYSIS

A. Experimental setup

In this section, we present the experiments done to evaluate proposed Harmony Search merged with Group Technology (HSGT) algorithm. We compared HSGT with three algorithms, Discrete PSO[3], GA[5] and Min-Min[35] scheduling algorithm. We considered three realistic workflows: (a) Montage workflow is used to generate custom mosaic of sky using the set of input images, it is I/O bound workflow and does not take much CPU processing time. (b) Inspiral is used to detect gravitational waves generated by various universal events, it is CPU bound workflow and takes large processing time. (c) Epigenomics is used to map the epigenetic state of human cells on genome. The DNA sequence is divided into multiple chunks that can be operated simultaneously. It is also a CPU bound workflow. These workflows are shown in figure 1.



(a) Montage (b) Inspiral

(c) Epigenomics

Figure 1. Organization of workflows used in experiments

We present the results achieved by simulating our proposed HSGT algorithm in WorkflowSim Toolkit[36]. Table 1 shows the parameters setting for HSGT and Table 2 showing parameters setting for WorkflowSim toolkit.

Table 1: Setting of parameters for HSGT

Parameter	Value
Size of population	50
Maximum iteration	100
HMCR	0.7
PAR	0.3

To estimate the performance of the given algorithmic procedure, execution time (makespan) is used as a parameter.

Table 2: Setting of Parameters for WorkflowSim

Entity	Parameter	Value
Datacenter	Total Datacenter	1
	Total Host	20
	Manager Type	Time- Shared
VM	Total VMs	5
	MIPS	400
	Memory (RAM) of VMs	512
	Bandwidth	1000
	Policy	Time-Shared

B. Result Analysis

In this section simulation result of the algorithms are shown with three considered realistic workflows (a) Montage (b) Inspiral and (c) Epigenomics. We have assumed five different types of VMs in a data center with the processing speed 1000 MIPS. Results of these three workflows with different number of tasks are given below. Since HSGT is a metaheuristic algorithm, each experiment is executed 10 times to obtain the average execution time. The graph based on the values of experiments done with Montage, Inspiral and Epigenomics workflows with varying number of tasks is given in figure 2, 3 and 4 respectively.

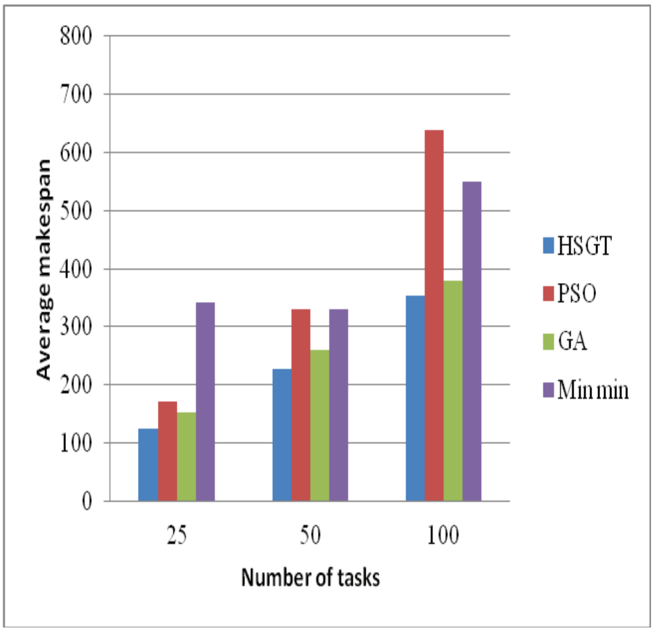


Figure 2. Makespan of HSGT, DPSO, GA and Min-Min algorithm on Montage Workflow.

From figure 2 we can clearly observe that HSGT gives lesser makespan as compared to DPSO, GA and Min-Min scheduling algorithms when the number of virtual machines is fixed. Makespan produced by HSGT in case of montage workflow is 30% lesser than DPSO, 12% lesser than GA and around 40% lesser than Min-Min scheduling algorithm.

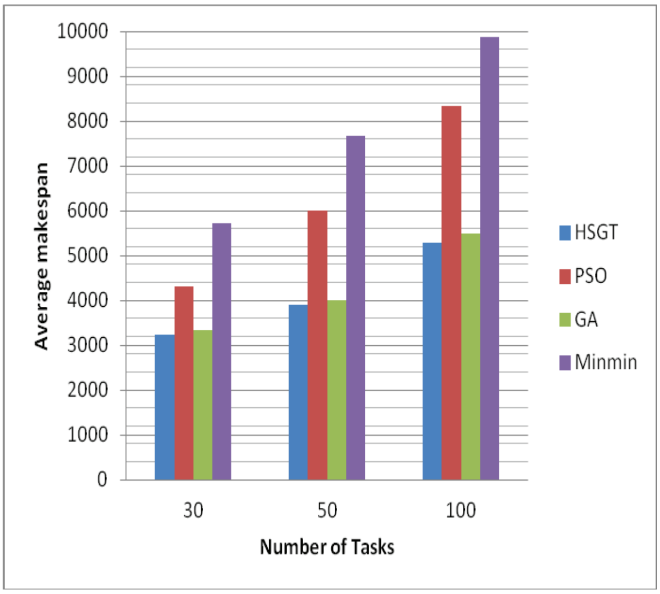


Figure 3. Makespan of HSGT, DPSO, GA and Min-Min algorithm for Inspiral Workflow.

From figure 3 we can clearly observe that makespan of HSGT algorithm is around 50% lesser than DPSO, around 10% lesser than GA and around 50% lesser than Min-Min scheduling algorithm.

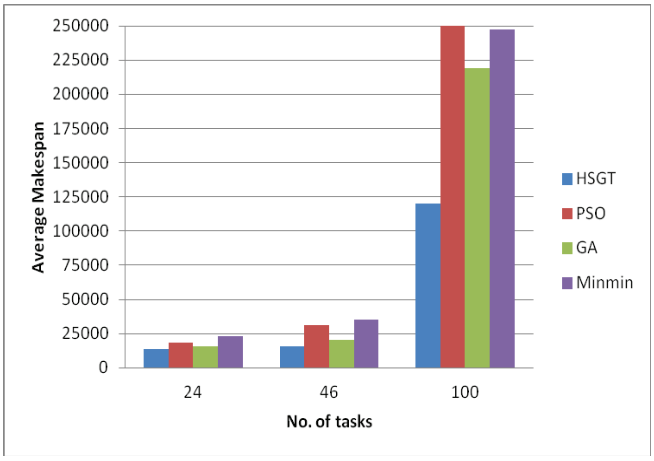


Figure 4. Makespan of HSGT, DPSO, GA and Min-Min algorithms for Epigenomics Workflow.

Figure 4 evidently shows that makespan of HSGT algorithm is around 31% lesser than DPSO, 13% lesser than GA and around 60% lesser than the Min-Min scheduling algorithm on Epigenomics workflow.

In Figure 5 comparison of makespan is given by considering different population size for Montage, Inspiral and Epeginomics workflows. Number of task remained fixed in this case.

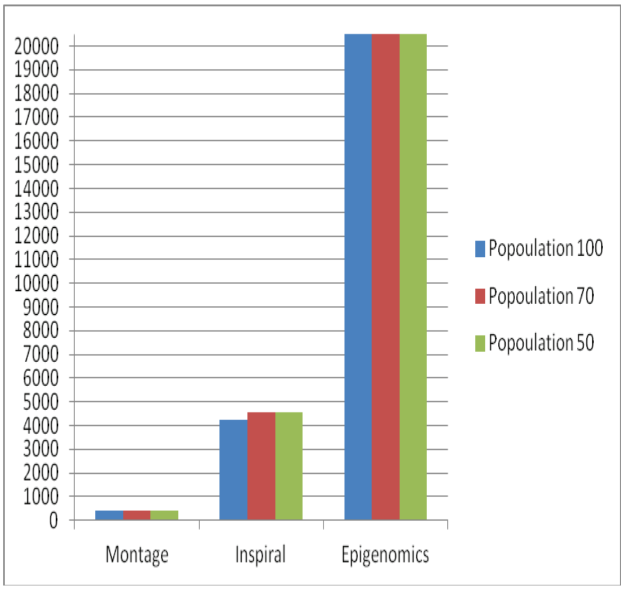


Figure 5. Makespan of HSGT algorithm on different workflows upon varying population size.

Figure 5 evidently shows that when number of population increases makespan of the HSGT algorithm decreases. This implies that HSGT algorithm would perform better in case of greater population size.

IV. CONCLUSION AND FUTURE WORK

In this paper, we presented a framework based on Harmony Search merged with Group Technology (HSGT) technique to optimize workflow scheduling for cloud focusing on reduction in makespan. From the experimental simulation we can sees that HSGT gives smaller makespan as compared to Discrete PSO, GA and Min-min scheduling algorithm. We compared our proposed algorithm with other existing scheduling algorithms for different workflows like Montage, Inspiral and Epigenomics on different number of tasks.

In future, we will extend the algorithm by considering other constraints like cost and deadline of tasks. Future research can be carried out on hybrid algorithms like merging of harmony search with genetic algorithm.

REFERENCES

[1] H. M. Fard, J. J. D. Barrionuevo, R. Prodan, T. Fahringer, "A Multi-Objective Approach for Workflow Scheduling in Heterogeneous Environment", 12th IEEE International Conference on Cluster, Cloud and Grid Computing, pp. 300-309, 2012.

[2] H. Wang, L. Cui, M. Xu, Y. Bi, "A multiple QoS constrained scheduling strategy of multiple workflows for cloud computing," IEEE international symposium on parallel and distributed processing with applications, pp. 629-634, 2009.

[3] S. Pandey, L. Wu, S. M. Guru, R. Buyya. "A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments," 24th IEEE international

conference in Advanced information networking and applications (AINA), pp. 400-407, 2010.

[4] R. Hassan, B. Cohanim, O. D. Weck, and G. Venter. "A comparison of particle swarm optimization and the genetic algorithm." In 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, pp. 1897. 2005.

[5] S.S. Pashaki, , E. Teymourian, R. T. Moghaddam. " New approach based on group technology for the consolidation problem in cloud computing- mathematical model and genetic algorithm." Comp. Applied Math. Vol. 48, pp. July 07, 2016.

[6] S.Pashaki, E.Teymourian, V. Kayvanfar, GH.M.Komaki, A. Sajadi. "Group technology-based model and cuckoo optimization algorithm for resource allocation in cloud computing." IFAC-PapersOnLine , vol. 48.3, pp.1140-1145 , December 31, 2015.

[7] M. Melnik , T. Trofimenko "Polyrhythmic Harmony Search for Workflow Scheduling." Procedia Computer Science. vol. 66, pp. 468-76, Dec 31, 2015.

[8] Z.W. Geem , J.H. Kim, G.V. Loganathan, "A new heuristic optimization algorithm: harmony search", Simulation, vol. 76(2),pp. 60-68, Feb 1, 2001.

[9] Y. Yuan, H. Xu, J.Yang, " A hybrid harmony search algorithm for the flexible job shop scheduling problem", Applied Soft Computing. vol. 13(7), pp.3259-3272, July 31, 2013.

[10] Q.K .Pan, P.N. Suganthan, J.J. Liang, and M. F. Tasgetiren, "A local-best harmony search algorithm with dynamic sub-harmony memories for lot-streaming flow shop scheduling problem." Expert Systems with Applications, vol. 38 issue 4, pp.3252-3259 , April 30, 2011.

[11] R. Shiyas, and V.M. Pillai, "A non-linear programming model for the machine grouping and a genetic algorithm based solution methodology", International Journal of Computer Applications, vol. 38(8), pp.37-41, January 2012.

[12] X.S.Yang and S. Deb, "Cuckoo search via Lévy flights", IEEE World Congress on Nature & Biologically Inspired Computing, NaBIC. pp. 210-214, 2009.

[13] M. Zineddine, "Vulnerabilities and mitigation techniques toning in the cloud: A cost and vulnerabilities coverage optimization approach using Cuckoo search algorithm with Lévy flights", Computers & Security, vol. 48, pp.1-18, Feb 28, 2015.

[14] F.Tao, Y. Feng, , L. Zhangand, T.W. Liao, "CLPS-GA: A case library and Pareto solution-based hybrid genetic algorithm for energy-aware cloud service scheduling", Applied Soft Computing, vol. 19, pp.264-279, June 30, 2014.

[15] A.A. Visheratin, M.A. Melnik, and D.Nasonov, "Workflow Scheduling Algorithms for Hard-deadline Constrained Cloud Environments", Procedia Computer Science, vol. 80, pp.2098-2106, 2016.

[16] J. Zhang, H. Huang, and X .Wang, "Resource provision algorithms in cloud computing: A survey", Journal of Network and Computer Applications, vol. 64, pp.23-42, April 30, 2016.

[17] M. Komaki, S. Sheikh, and E. Teymourian, "A Hybrid Harmony Search algorithm to minimize total weighted tardiness in the permutation flow shop", IEEE Symposium on Computational Intelligence in Production and Logistics Systems (CIPLS), pp. 1-8, December 2014.

[18] R . Raju, R.G. Babukarthik, D.Chandramohan, P. Dhavachelvanand, T.Vengattaraman. "Minimizing the makespan using Hybrid algorithm for cloud computing", IEEE 3rd International Conference in Advance Computing (IACC), pp. 957-962, February, 2013.

[19] M. Yakhchi, G.M. Ghafari, S Yakhchi , M, Fazeli , A, Patooghi , "Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in cloud computing infrastructures." 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), pp. 1-5, 2015 May 27, 2015.

[20] H. Sajedi, and M. Rabiee, "A Metaheuristic Algorithm for Job Scheduling in Grid Computing", International Journal of Modern Education and Computer Science, vol. 6(5), pp.52-59, May 2014.

[21] O. Norouzpour, N. Jafarzadeh, "Using Harmony Search Algorithm for Load Balancing in Cloud Computing", Indian Journal of Science and Technology, vol. 8(23), Sep 21, 2015.

[22] X.Wang, X.Z. Gao and K. Zenger, "The Overview of Harmony Search, An Introduction to Harmony Search Optimization Method" Springer International Publishing, pp. 5-11, 2015.

[23] S. Rao, S.V.L. Narasimham, M.R. Raju and A.S. Rao, "Optimal network reconfiguration of large-scale distribution system using harmony search algorithm." IEEE Transactions on Power Systems, vol. 26(3), pp.1080-1088, August 26, 2011.

[24] M. Mahdavi , M. Fesanghary , E. Damangir, "An improved harmony search algorithm for solving optimization problems", Applied mathematics and computation, vol. 188(2), pp.1567-79, May 15, 2007

[25] G. Soni and M. Kalra, "A novel approach for load balancing in cloud data center", IEEE International in Advance Computing Conference (IACC), pp. 807-812, 21 February 2014.

[26] Z. Zhu, G. Zhang, M. Li and X. Liu, "Evolutionary multi-objective workflow scheduling in cloud", IEEE Transactions on Parallel and Distributed Systems, vol. 27(5), pp.1344-1357, May 1 , 2016.

[27] S. Mistry, A. Bouguettaya, H. Dong and A.K.Qin, " Metaheuristic Optimization for Long-term IaaS Service Composition", IEEE Transactions On Service Computing, vol. x, no. x, March 2016.

[28] S. Joshi and S. Kaur, "Cuckoo search approach for virtual machine consolidation in cloud data centre", International Conference in Computing, Communication & Automation (ICCCA), pp. 683-686, May 2015.

[29] M. Yakhchi, S.M. Ghafari, S. Yakhchi, M. Fazeli and A. Patooghi, " Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in cloud computing infrastructures", 6th International Conference in Modeling, Simulation, and Applied Optimization (ICMSAO), pp. 1-5, May , 2015.

[30] M.Kalra and S.Singh, "A review of metaheuristic scheduling techniques in cloud computing", Egyptian Informatics Journal, vol. 16, issue 3, pp.275-295, November 2015.

[31] H.H.Sharif, El-Kilany, K.S. and M.A. Helaly. "A genetic algorithm approach to the group technology problem", In International Multi Conference of Engineers & Computer Scientists, vol. 2008, pp. 1755-60 , 2008

[32] http://www.ibm.com/developerworks/websphere/techjournal/1206_dejesus/1206_dejesus.html last accessed on 29/07/2016

[33] http://www.appcore.com/wp-content/uploads/types-of-cloud-computing.png last accessed on 29/07/2016

[34] http://www.nist.gov/itl/csd/cloud-102511.cfm last accessed on 02/08/2016

[35] G. Liu, J. Li, J. Xu. "An improved min-min algorithm in cloud computing". In Proceedings of the 2012 International Conference of Modern Computer Science and Applications 2013 (pp. 47-52). Springer Berlin Heidelberg.