IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-6, June 2019 E-ISSN: 2347-2693

Application of Genetic Algorithms: Task Scheduling in Cloud Computing

Srishti Garg[1*], P. K. Chaurasia[2]

[1,2]Dept. of Information technology, Babasaheb Bhimrao Ambedkar University, Lucknow, India

Corresponding Author: srishti.garg18@gmail.com1

DOI: https://doi.org/10.26438/ijcse/v7i6.782787 | Available online at: www.ijcseonline.org

Accepted: 12/Jun/2019, Published: 30/Jun/2019

***Abstract*:** Cloud computing implies to the delivery of information technology (IT) services which functions by retrieving the resources from the Internet; implementing various web-based tools and applications, in opposition to directly connecting to a server. In a nutshell, cloud computing works on the purpose of taking all the efforts involved in processing large quantities of data from the device carried around and moving that work to huge computer clusters far away in a virtual space. The internet becomes the virtual information space "THE CLOUD", and all the data, work and applications are available from any device which when connected to the internet, anywhere in the world accesses it. Cloud computing is the provisioning of business computing model and providing multifarious facilities over the internet. Data that are looked over by third parties or other person at various remote locations can be assessed by individuals and various other business organizations through Cloud Computing applications. A cloud environment is categorized into computing clouds and data clouds.Task scheduling is considered to be the core feature and plays an important role in maintaining the quality of service in the cloud computing environment. The application of genetic algorithm in cloud computing task scheduler environment is a topic gaining popularity in the recent years. But, achieving an efficient task scheduling methodology is a major attribute for harnessing the potential of cloud computing applications in an effective manner. The objective of this paper is to discuss the application of heuristic algorithms; the use of GAs to minimize the total scheduling time and execution cost of tasks improves task completion time and maximize resource utilization in cloud computing framework by a task scheduler genetic algorithm.

***Keywords*:** *Cloud Computing; Genetic Algorithm; Selection Operation; Crossover Operation; Task Scheduling; Mutation Operation; Fitness Function.*

I. INTRODUCTION

With the worldwide growth of access of Internet and huge amounts of data in terms of their volume, velocity, and variety through the Internet, cloud computing becomes rapidly increasing in the industry, academia and society. Cloud computing is the product integration of traditional computer and network technologies such as Distributed Computing, Parallel Computing and Virtualization. Cloud computing involves provision of on-demand computing and storage services with high and dynamic performance with scalability of virtualized resources over the Internet. Cloud computing is a reliable computing paradigm. However, with the rise in energy consumption of the cloud data centres has become a prominent problem. Task scheduling is a significant feature to measure the overall performance of the cloud computing which substitutes to be the major motivation in the formulation of the paper. The amount of task that cloud computing needs to deal with and how to make use of the available resources to allocate the cloud task reasonably is enormous [1].

II. DESCRIPTION OF TASK SCHEDULING IN CLOUD COMPUTING

The basic objective of cloud computing is to provide an optimal scheduling of the tasks, so that it can be accessed by the users and the entire cloud system with optimal operation time, improved Quality Of Service (QoS) at the same time and load balancing. Task scheduling is for the optimal matching of tasks and resources. It refers to the designing and providing of a scheduled management framework for engineering implementation in clouds. It is an important aspect in cloud computing resource management as it helps to minimize energy and timespan.

The cloud is mainly to provide users with QoS**.** The basic framework is composed of four main components: user portal, information service, task scheduler and cloud data centre with physical machines (PM). The user portal provides an

interface for users to submit task unit. The task unit further divided into small tasks to be executed in PMs as shown in Fig.1 [2].

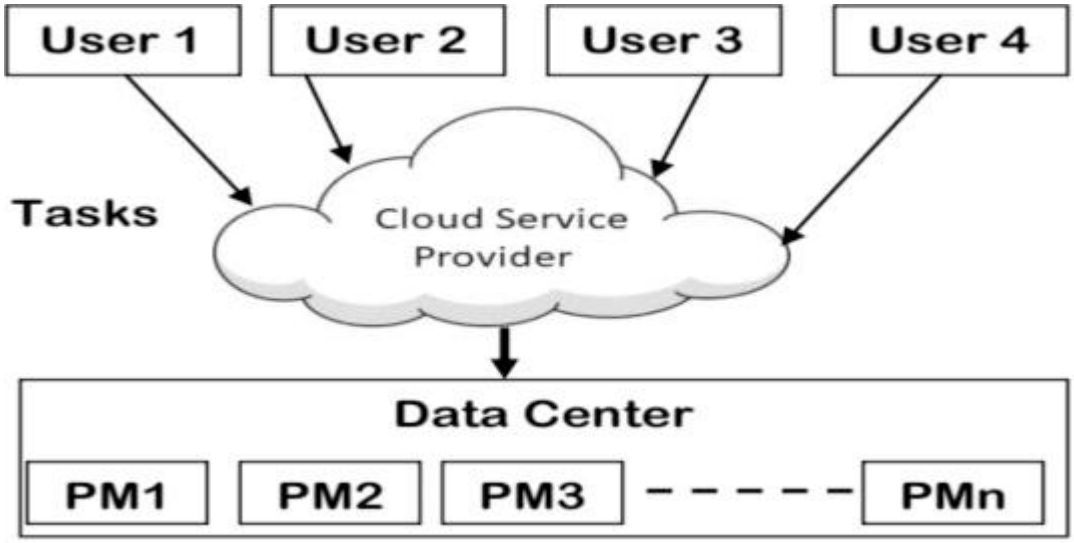

Figure 1: Components of Cloud Computing

The Information Service Provider (ISP) keeps the details of resource utilization and other log information to help scheduler to schedule a task to a PM in a data centre. The scheduler accepts the task unit from the user portal and uses ISP to choose the appropriate PM in a cloud data centre. After the task unit completes its execution, the result and the new status of the resource will be sent back to the ISP's for another scheduling.

The scheduling problem covers the following assumptions:

- The input of the task is the minimum child task, which means that large tasks have been split into small tasks and the length of the minimal task is random.
- All cloud computing resources are mapped to virtual machines and the performance of all virtual machines is random.
- The number of tasks is much larger than the number of virtual machines [3].

III. GENETIC BASED TASK SCHEDULING ALGORITHM

The essential characteristics a Cloud provider should guarantee are scheduling of user's tasks optimally and appropriate utilization of the Cloud resources with the best throughput.

THE PROBLEM STATEMENT

The complexity of scheduling user tasks in the cloud computing environment is proportional to the number of user's tasks which implies that the cloud service provider needs a "good-enough" algorithm to efficient scheduling of the tasks on the cloud. The proposed task scheduling algorithm in the cloud environment works by making some modifications in the basic GA. The selection method namely the Tournament selection method is deployed to overcome the limitation of the population size by selecting the best chromosomes. The proposed algorithm is, therefore, named to be the Tournament Selection Genetic Algorithm (TS-GA). The modifications work is made by taking into account the parents in each of the population size besides the produced child after the crossover process [4].

PRINCIPLES OF GENETIC ALGORITHM

Genetic Algorithm is a proliferating growing area of Artificial Intelligence. Genetic algorithm is a search and optimization "adaptive heuristic" technique which is in line with Charles Darwin's Theory of Natural Evolution which coins the principle of "Survival of the Fittest". This method works by scheduling of the tasks assigned to resources .The genetic determination is based on the mathematical value of the fitness function for each parameter in the task scheduling procedure.

GENERATION OF INITIAL POPULATION: The population to be initialized is the set of all individuals. These are used to find out the solution in a genetic algorithm optimally. The individual is defined to be all the possible solutions and the

representation is as of a chromosome. This makes the genetic operations more effective. The next generation is formed by the selected individuals with consideration of the initial population and performing operations on them. A specific criterion is used for selecting the mating chromosomes.

FITNESS FUNCTION: The function which takes a solution after population initialization as the input and produces the suitability of the solution as the output. The fitness value depicts the precedence and its performance of an individual in the population. The fitness function is defined to be a motivating factor as the survival rate of individuals is highly dependent on the fitness value.

SELECTION: The selection operator in genetic algorithm is defined by selecting two parent chromosomes from the size of the population according to their fitness value. For this selection strategies such as Roulette Wheel, Tournament Selection, and Rank Selection are used to select the best chromosomes.

CROSSOVER: Crossover operation can be achieved by selection of two parent individuals and then creation of a new individual tree by performing alternation and reformation in the parts of selected parents [5].

MUTATION: Mutation is the next step after crossover operation which is to enhance the genetic diversity in the generated population. By following the process of mutation probability, the new offspring after crossover are mutated at each locus point in the initial chromosome. This is done to avoid homogeneity and enable the genetic algorithm to provide better solutions after reproduction and crossover operations.
The flowchart of genetic algorithm is shown in Fig.2:

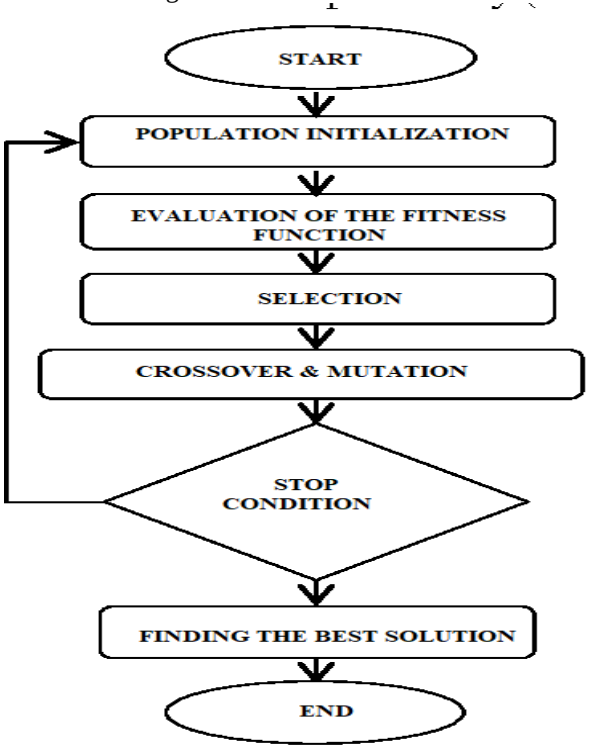

Figure 2: Flowchart for Genetic Algorithm

IV. IMPLEMENTATION OF TASK SCHEDULING GENETIC ALGORITHM (TS-GA)

A modified GA is propounded to enhance completion time for scheduling and execution of the tasks on the virtual machines in the same time, minimization of the total cost of usage of the resources and maximization of utilization of the resources in cloud computing framework. The basic idea of the algorithm is to implement that after each selection in the population, a solution is obtained which might satisfy good fitness function. Further, this solution is not removed from the population, but it is chosen

and added to the population when next iteration is started. This step is considered as a good step as some of the iterations can generate the best solution [6].

INITIALIZATION OF THE POPULATION: The population is randomly generated using BINARY ENCODING (0, 1).The binary bit is obtained by encoding the virtual machines and the executed tasks into the bit.

THE FITNESS FUNCTION: Task scheduling in the Cloud computing works with the objective of reducing the completion time for task execution within the available resources. The fitness function determines the measure of fitness value for a given solution in solving the problem quantitatively. The results generated by the fitness function should be mostly instinctive. The best/worst candidates should have best/worst score values. Therefore, for the scheduling purpose of cloud computing, we calculate the execution time of each task for each virtual machine which, in turn, reduces the completion time. Figure 3 shows representation of tasks and virtual machines.

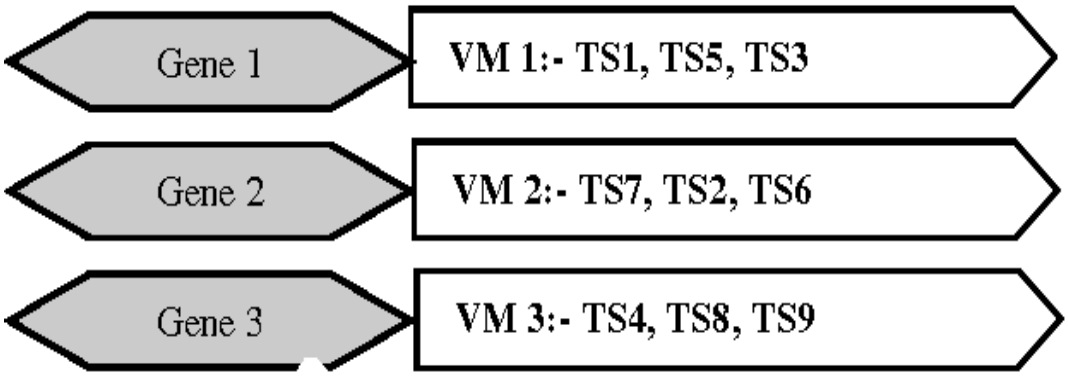

Figure 3: Representation of Tasks (TS's) and Virtual Machines (VM's)

SELECTION PROCESS: Cloud computing deploys parallel implementation. For this purpose Tournament selection method is taken into consideration; as it is found to be more efficient computationally and is more responsive to the parallel implementation. The Tournament Selection works by selecting two individuals randomly from the population. This method helps in overcoming the limitation of the population size. For this beginning is done by choosing a random number r between 0 and 1. Considering if $r < k$ (where k is a parameter, for example, 0.95), the fitter of the two individuals is selected to be a parent; otherwise the less fit individual is selected. The individuals which are not selected are then returned to the original population and could be selected again.

CROSSOVER: The new crossover in TS-GA works in a different strategy from the crossover operation in the original genetic algorithm. The new crossover operation works by considering the two chromosomes also as offspring which are selected to the crossover process to generate two off springs. This crossover produces four children [7]. The two best children are picked from them after the crossover process. Fig. 4 shows the process of crossover in genetic algorithm.

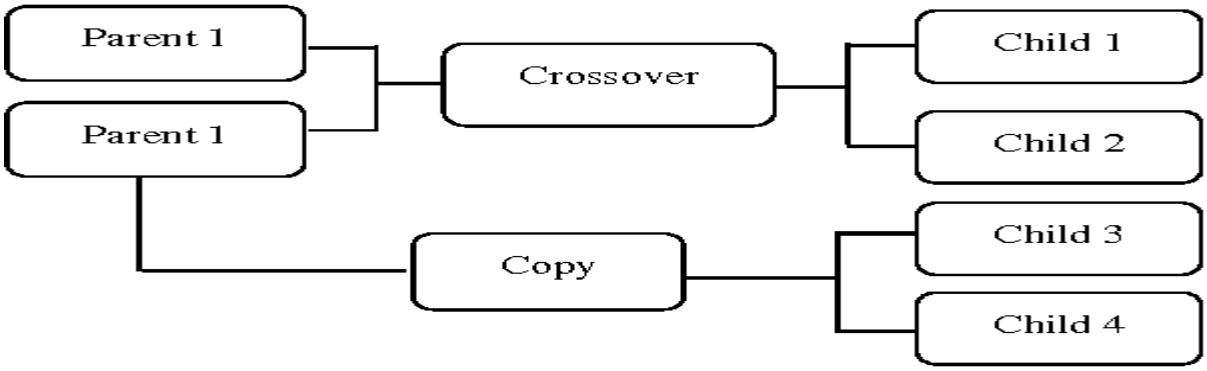

Figure 4: Crossover Process

INITIALIZATION OF THE SUBPOPULATION: The new populations after crossover (i.e., the subpopulation) are added into old populations (i.e., the parents); after performing each iteration [8]. The introduction of this step helps in the enhancement of the diversity of population [9].

KEEPING THE BEST SOLUTION: There is also a probability of a solution being present in the population that might hold a sound fitness value and satisfies the criterion of a good fitness function. But the solution is not selected neither removed from the population during the crossover process. This solution with good fitness value is chosen and added to the population size when next iteration is started to satisfy the optimization criteria in the TS-GA.A best solution can be obtained after some of the iterations. This step is considered as good step [10] [11].

V. CONCLUSION

Therefore, Cloud computing can be detailed to be an emerging technology which enables the organization to utilize hardware, software and applications without any upfront cost over the internet. The challenge before the cloud service provider is that in a dynamic environment computing resources need to be managed efficiently and effectively; This facilitates the computing device not to be in an under-utilized or over-utilized state . A good task scheduling technique is always required for the dynamic allocation of the task to avoid such a situation.[12]
The following interpretations can be drawn from the above implementation:

- The Tournament Selection method is implemented as the selection process as it helps in obtaining the possible best solution.
- The next generations with best solutions are facilitated by considering the solutions not chosen during the selection process and adding to the new population size [13].
- The new generated population after crossover operation is introduced by considering parents individuals as new child [14].
- After each iteration is performed, new populations after crossover (i.e., subpopulations) are added into old populations (i.e., parents).

Therefore, the task scheduler genetic algorithm in this paper targets to implement the effective minimization of the completion time and cost, and maximization in the resource utilization of the cloud computing framework [15].

REFERENCES

[1] S. H. Jang, T. Y. Kim, J. K. Kim, and J. S. Lee, "The study of genetic algorithm-based task scheduling for cloud computing," International Journal of Control and Automation, Vol. 5,pp. 157-162,2012.
[2] T. Goyal and A. Agrawal, "Host Scheduling Algorithm Using Genetic Algorithm In Cloud Computing Environment," International Journal of Research in Engineering & Technology (IJRET), Vol-1, 2013.
[3] B. Furht, "Armando Escalante Handbook of Cloud Computing," ISBN 978-1-4419-6523-3, Springer 2010.
[4] F. Etro, "Introducing Cloud Computing," in London Conference on Cloud Computing For the Public Sector, pp. 01-20,2010.
[5] R. Kaur and S. Kinger, "Enhanced Genetic Algorithm based Task Scheduling in Cloud Computing," International Journal of Computer Applications, Vol. 101, 2014.
[6] J. W. Ge and Y. S. Yuan, "Research of cloud computing task scheduling algorithm based on improved genetic algorithm," in Applied Mechanics and Materials, pp. 2426-2429, 2013.
[7] V. Vignesh, K. Sendhil Kumar, and N. Jaisankar, "Resource management and scheduling in cloud environment," International Journal of Scientific and Research Publications, Vol. 3, p. 1, 2013.
[8] Z. Zheng, R. Wang, H. Zhong, and X. Zhang, "An approach for cloud resource scheduling based on Parallel Genetic Algorithm," in Computer Research and Development (ICCRD), 2011, 3rd International Conference on, pp. 444-447, 2011.
[9] S. Singh and M. Kalra, "Scheduling of Independent Tasks in Cloud Computing Using Modified Genetic Algorithm," Computational Intelligence and Communication Networks (CICN), 2014 International Conference on, pp. 565-569, 2014.
[10] R. Buyya, R. Ranjan, and R. N. Calheiros, "Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities," in High Performance Computing & Simulation, 2009. HPCS'09. International Conference on, pp. 1-11, 2009.
[11] B. Kruekaew and W. Kimpan, "Virtual Machine Scheduling Management on Cloud Computing Using Artificial Bee Colony," in Proceedings of the International Multi-Conference of Engineers and Computer Scientists, 2014.
[12] M. Mitchell, An introduction to genetic algorithms: MIT press, 1998.
[13] R. N. Calheiros, R. Ranjan, C. A. De Rose, and R. Buyya, "Cloudsim : A novel framework for modelling and simulation of cloud computing infrastructures and services," 2009, arXiv preprint arXiv: 0903.2525, 2009.
[14] R. Sahal and F. A. Omara, "Effective virtual machine configuration for cloud environment," in Informatics and Systems (INFOS), 2014 9th International Conference on, pp.-15-20, 2014.
[15] D. M.Abdelkader, F.Omara," Dynamic task scheduling algorithm with load balancing for heterogeneous computing," Journal of system Egyptian Informatics, Vol.13, pp.135–145, 2012.
[16] International Journal of Scientific Research in Computer Sciences and Engineering (ISSN: 2320-7639)
[17] International Journal of Scientific Research in Network Security and Communication (ISSN: 2321-3256)

AUTHORS PROFILE

Srishti Garg, B.Tech. (Computer Science and Engineering) is currently pursuing M.Tech. (Software Engineering) in the Department of Information Technology from Babasaheb Bhimrao Ambedkar University, Lucknow, Uttar Pradesh, India. She received her Bachelor Degree in Computer Science Engineering from Uttar Pradesh Technical University, Lucknow (India) in 2014. Her research area includes Cloud Computing, Virtualization.

Pawan Kumar Chaurasia is working as an Assistant Professor in the Department of Information Technology, Babasaheb Bhimrao Ambedkar University, (A Central University), Lucknow, India. He has published more than 30 International Journal papers. His research area includes Software Reliability, Software Quality and Software Testing.

