Resource Management in Cloud Computing using SVM with GA and PSO

G. Anushuya, K. Gopikaa, S. Gokul Prasath Dr. P. Keerthika
Department of Computer Science and Engineering
Kongu Engineering College,
Perundurai,Erode.

***Abstract*— In the modern world,Resource allocation has become a major issue.In particular,To accomplish a tasks with lower cost has become a significant issue when the resources are getting less.The main aim of this research is to fully utilize the hardware resources using Machine Learning algorithms.A prediction mechanism Support Vector Regression(SVR) is used to estimate whether it is possible to redistribute the resources based on the current status of all the virtual machine in the physical machine.A resource dispatch mechanism Genetic algorithm(GA) is designed for reallocation of resources.The experimental results of GA and Particle Swarm Optimization(PSO) are compared to determine the better one based on the generation in which the fitness function is attained.In addition,our proposed mechanism maintains a desirable performance in the cloud environment.**

Keywords— Cloud computing,SVM,PSO,GA,

I. INTRODUCTION

The Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. Within that cloud system, massive distributed data center infrastructure, virtualized physical resources, virtualized middleware platform as well as applications are all being provided and consumed as services.Cloud computing being a Business to Business (B2B) model,users can access the software and other services through the Internet.Since cloud computing has fully utilized the virtualization technology that can provide users with high flexibility and extension,many Industries used cloud computing in the business model. Hence, it is an important issue of resource management for the cloud system to allocate the virtual machines (VMs) based on the user's request.Internet Service Providers(ISPs) will formally sign a contract, called Service Level Agreement (SLA) with the consumers to determine the price of each level of service, the contracted content which sometimes refers to some performances metrics of resources, such as the performance of CPU, the capacity of memory, and the response time. ISPs can evaluate their applications to determine different prices according to the performance metrics of resources.

As mentioned Chenn-Jung Huang[1], Resource management being an important issue is mostly applied in some arrangement of working tasks. Different strategies of selections will bring different cost and efficiency. Therefore, how to find a best resource allocation strategy, especially for the limited resource, aiming at each kind of goal is a challenging work. This research work focused on the application of Evolutionary Algorithms (EA) in the area of cloud computing. After considering the individual representation of some algorithms, the gene coding representation of GA is more suitable in this work. Besides, numerous researchers have proposed GAs to deal with the optimization problems, and the Schema theorem for GA proposed by Holland [2] illustrated that GA is a robust searching approach.Hence,We chose GA as the resource allocation algorithm in this work, and we will compare with the other EAs as proposed by Frederic Nzanywayingoma[3], such as the PSO algorithm to determine the better one based on the generation in which the fitness function is attained.The main advantage of GA compared with other EA methods is that it only needs a fitness function to evaluate the quality of different solutions and there is no necessary to offer a particular algorithm to solve a given problem.

SVR is a kind of supervised machine learning method that recognizes patterns and analyze data, mostly used for classification and regression analysis. The major difference between the SVR and traditional regression techniques is that the SVR employs the structural risk minimization (SRM) approach, rather than the empirical risk minimization (ERM) approach typically adopted in statistical learning.The SRM attempts to minimize an upper threshold on the generalization rather than minimize the training error, and is expected to perform better than the traditional ERM approach. Added to that, the SVR is a convex optimization, which guarantees that the local minimization is the unique minimization. Many researchers have adopted SVR to deal with the classification and regression problems.Users can handle the overall schedule more efficiently with this method. In addition on comparing three regression approaches, including SVR, Back-propagation Neural Network, and Partial Least Squares, to predict the Cold Modulus of Silicon Ceramic, and the results showed that SVR obtained better performance in root mean square error than the other two methods.

An application service prediction module built with SVR is used to estimate whether there is a necessary for reallocation of resources. When new VMs are demanded, the system will consider the individual loading to proceed the prediction with SVR. If the result needs to change the allocation strategies, the system will utilize GA and PSO to perform reallocation of resource and compare the efficiency between them based on the criteria stating that at which genertion they achieve the fitness function.

II. RESOURCE MANAGEMENT ALGORITHM

Fig. 1 illustrates the architecture of the resource management scheme for cloud computing proposed in this work. An application service resource pool is used to collect all applications and an application monitor is used to record the overall utilization of system resources. Additionally, a physical machine resource pool is used to provide resource, CPU or Memory, for the hosts and a look-up table,such as the remaining resource table is used in determining the strategies of increasing or decreasing the number of virtual machines (VMs) requested by each application service.

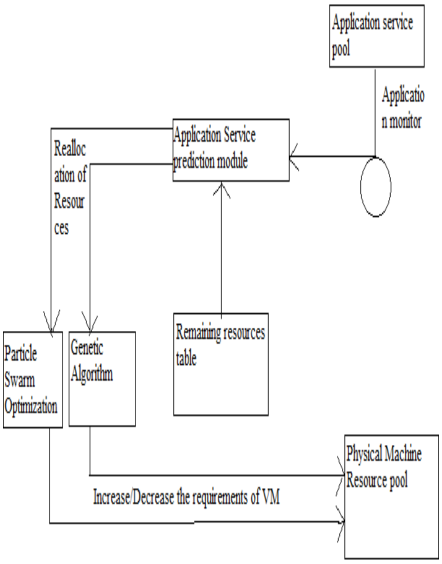
Application service pool
Application monitor
Application Service prediction module
Reallocation of Resources
Particle Swarm Optimization
Genetic Algorithm
Remaining resources table
Physical Machine Resource pool
Increase/Decrease the requirements of VM

Fig. 1. The architecture of the resource management system.

Application Service Prediction Module

This module mainly estimates the requirements of the resource in the applications to remind the system of creating or collecting the VMs. In addition, a well-known time series predictor, namely SVR, is embedded in this module to predict whether reallocation of resources is needed with the assistance of a lookup table, that includes the remaining usable resource table, which records all related usable resources in the applications. With the above-mentioned methods, this module will determine whether the VMs should be increased or decreased for the request from the application services.

Support Vector Regression (SVR)

To solve a nonlinear regression or functional approximation problem, the SVR nonlinearly maps the input space into a high-dimensional feature space using an appropriate kernel representation, such as polynomials and radial basis functions with Gaussian kernels. This approach is utilized to build a linear regression hyperplane in the feature space, which is nonlinear in the original input space.When a user request for a VM,the role of SVM is to determine whether reallocation of resources are needed based on the current status of all the VMs present in the Physical machines(PMs).

Resource Reallocation Module

As shown in Fig. 1, this module collects the information from the Application Service Prediction module and a lookup table, along with SLAs requested by the customer, to determine the number of VMs requested by each application service. GA is applied in this work to deal with the optimization problem, and the fitness function is designed.

Genetic algorithm (GA)

As mentioned by Safwat A. Hamad[4] VMs which are established in the PMs are transferred into a binary code as the initial population, called chromosomes. Each element in the chromosome is either 0 or 1, and the higher fitness value will be kept to generate the next generation during the procedure of recombination and mutation. The new generation will run the same steps as of their parents did until the stop criteria is satisfied. In addition, we do not need to set fixed time interval to activate GA because the system will proceed the adjustment strategies according to the real-time demand of VMs. Once a new request of VM arrives, the system will run GA to adjust the overall allocation of resources.

Before finding the best solution with GA, a set of initial population is generated in random, and the chromosomes in matrix format, with n VMs, and m resources. A horizontal row stands for the location of each VM and a vertical column represents the distribution of each resources in the VM.Notably if the resource is not established in VM then the entry of the matrix aij is set to 0.Meanwhile, before GA is utilized in the initialization, the system will run SVR to evaluate the number of resource utilization.The system has proceeded the allocation of VMs for each individual demand in the initialization step, and the operation of GA is to finely adjust the overall allocation to achieve better performance.

Fitness function

We first evaluate the resources of the physical machines that initially can provide the resource distribution for all virtual machines.Here we determine the fitness function by considering each VMs which has the maximum resource utilization.The new generation proceeds as of the parents till the fitness function is attained.

Selection, reproduction, crossover, and mutation

Better individuals get higher chance to survive and create offspring. In this work, we adopt roulette wheel selection and one-point recombination as the recombination method, respectively, and the recombination rate is 100%. To make sure that the sum of each column corresponds to the total VMs that can be allocated by a physical machine, we restrict each recombination occurring only in the same column. Notably, we chose one-point recombination, which has lower overhead, as our recombination type because of the real-time demand of VMs in this work.The mutation rate is set as the reciprocal of the size of the initial population in this work. A bit from the element in matrix is chosen in random to perform the NOT operation on it.The purpose of setting the termination conditions is to reduce the computing time and to avoid unnecessary evolution. In general, the process is repeated until one of the given termination conditions has been reached.

Particle Swarm Optimization(PSO)

PSO is used to solve the optimization problems. In PSO, each single solution in the search space called as a particle. All of particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the flying of the particles. The particles fly through the problem space by following the current optimum particles. PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In every iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called pbest. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called gbest. When a particle takes part of the population as its topological neighbors, the best value is a local best and is called lbest.After finding the two best values, the particle updates its velocity and positions with following equations.

v[] = v[] + c1 * rand() * (pbest[] - present[]) + c2 * rand() * (gbest[]-present[])
present[]=present[]+v[]

v[] is the particle velocity, present[] is the current particle (solution). pbest[] and gbest[] are defined as stated before. rand () is a random number between (0,1). c1, c2 are learning factors. usually
c1=c2=2.
The pseudo code of the procedure is as follows For each particle

```
Initialize particle END
Do
For each particle Calculate Fitness value
If the fitness value is better than the pBest Set the current value as the new pBest.
End
Choose the particle with the best fitness value of all the particles as the gBest
For each particle
Calculate particle velocity Update particle position End
```

Comparisons between GA and PSO

Many Optimization algorithm follows same procedure as follows:
1.Random generation of an initial population 2.Reckoning of a fitness value for each subject. It will directly depend on the distance to the optimum. 3.Reproduction of the population based on fitness values. 4.If requirements are met, then stop. Otherwise go back to 2.

As suggested by Frederic Nzanywayingoma[3]Both GA and PSO starts with the initial population.Next generation is produced as of the parents and the iteration continues till the fitness value is achieved nearer to the optimum value.The iteration continues till the stop criteria is achieved.However, PSO does not have genetic operators like crossover and mutation. Particles update themselves with the internal velocity. They also have memory, which is important to the algorithm. Compared with GAs, the information sharing mechanism in PSO is significantly different.

In GAs, chromosomes share information with each other. So the whole population moves like a one group towards an optimal area. In PSO, only gBest (or lBest) gives out the information to others. It is a one -way information sharing mechanism. The evolution only looks for the best solution.On comparison PSO attains to be the best since it reaches the fitness value in a generation before than the generation in which the GA reaches its fitness value.

III. CONCLUSION

In this work, a resource allocation system is performed effectively with the help of two optimization algorithms.Here the reallocation of resources is done by GA and PSO based on the prediction of SVM.The efficiency of both the algorithms is compared based on the generation in which the fitness value is attained.In the future work,reallocation of resources is done by taking the best solution of GA and this solution is given as a initial population for PSO,this inturn refers to the reallocation of resources by hybrid approach,This attains a better desirable performance.

REFERENCES

[1] Chenn-Jung Huang, Yu-Wu Wang, Chih-Tai Guan, Heng- Ming Chen, and Jui-Jiun Jian, Applications of Machine Learning to Resource Management in Cloud Computing, International Journal of Modeling and Optimization, Vol. 3, No. 2, April 2013.

[2] J. H. Holland, Adaptation in natural and artificial systems, Ann Arbor, MI: Univ. Michigan Press, pp. 89-97, 1975.

[3] Frederic Nzanywayingoma, Prof. Yang Yang, Analysis of

Particle Swarm Optimization and Genetic Algorithm based on Task Scheduling in Cloud Computing Environment, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 1, 2017.

[4] Safwat A. Hamad, Fatma A. Omara, Genetic-Based Task Scheduling Algorithm in Cloud Computing Environment, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016.

[5] Rajib Kumar Bhattacharjya, Introduction to Particle Swarm Optimization.