CHAPTER 8

CONCLUSION AND FUTURE WORK

8.1 CONCLUSION

Cloud resource provisioning is a highly sensitive task which needs to be done with greater concern for increasing the user satisfaction level as well as the profit level of the cloud service providers. Admission control is one of most popular technologies which can perform optimized resource provisioning in the presence of heterogeneous environment with varying user requests. Better and effective admission control task can lead to improved utilization of cloud resources with increased profit for cloud service providers. This is achieved by introducing various research methodologies that attempt to allocate the resources with the concern for increased profit and reduced makespan.

In the first research called machine learning based admission control, SVM and ANN algorithms are used for performing effective scheduling. Both the proposed approaches attempt to control the admission control process by allocating the tasks to the resources that can provide more Return on Investment (ROI). This is done by learning the profit and the non-profit parameters of the cloud resources and the tasks in an enhanced manner.

In the second research methodology namely Position Balanced Parallel Particle Swarm Optimization Approach, optimized resource allocation and scheduling is done. PB-PPSO is used as the classifier to extract
rules from the historical data. The initialized particles randomly generate rule from historically executed data. The optimal rules are found after much iteration based on the multi objective fitness values such as makespan and price.

In the third research work namely, NMF clustering based PB-PPSO, computation overheads occurring in the previous work is reduced considerably. This is done by grouping the tasks and resources together based on similarity using NMF based clustering approach. After clustering of tasks and resources, the rules are generated for cluster of tasks and resources using the methodology called the PB-PPSO. Thus the number of generated rules can be reduced considerably in case of arrival of more number of tasks also.

In the fourth research work, Data Locality Aware Prioritized User Job Scheduling (DLAPUJS) increases the profit of cloud service providers by reducing the data transfer cost considerably and giving priority for the users from who high profit can be earned. KNN classification algorithm is used in this work to categorize the cloud users based on their prioritization level. The users are categorized based on the demand submitted by them and the high prioritized user’s task would be scheduled first. After categorization, the data required by the user task would be located in the cloud environment.

The overall findings of this research work conclude that the proposed research methodologies can accomplish the admission control process in the cloud computing environment to the satisfaction of user specified constraints. The experimental tests conducted prove that the proposed research namely DLAPUJS provides better results than the existing and previous research works in terms of increased profit and decreased number of VMs initialized as well as execution time with reduced data transfer cost.
8.2 FUTURE WORK

The current research has focused on increasing the profit with improved user satisfaction level using four different approaches like machine learning based admission control, PB-PPSO, NMF-PB-PPSO, DLAPUJS. The future enhancements are given below.

- In future, Different SLA parameters that focus on improving the cloud resource provider’s reputation level can be considered to improve resource allocation strategies.
- Some optimized scheduling algorithms can be considered for load balancing problems that focus on avoiding the fault tolerance problem.
- Migration of virtual machines across the resources which are located at different locations can be considered for enhancing the task execution.
- QoS negotiation techniques can also be introduced by improving the reputation level of cloud service providers considerably in addition to satisfaction levels of users.