IMPROVING THE QUALITY OF CLOUD SERVICE BY ENSEMBLE PREDICTION

Lydia jeba[1] , Rathna.M[2]

[1,2]*Faculty of Computing, Sathyabama University.*

[1]lydiajeba@gmail.com , [2]rathnamohandoss@gmail.com

ABSTRACT

Cloud Computing provides pay-as-you-go style service that attracts the user to use cloud rather than opting for traditional computing. Cloud provides Saas ,Paas and Iaas services. In recent days Iaas plays a vital role. In this paper, to improve the quality of cloud service, especially provision of virtual machine and planning of cloud capacity is increased by performing prediction. Prediction is done by ensemble prediction model. Thus by predicting the demands for virtual machine, the quality of cloud service will be improved. To quantify the prediction error Cloud Prediction Cost (CPC) method is suggested.

Keywords: cloud computing, cloud analytics, data mining.

INTRODUCTION

Due to the popularity of cloud services, there will be an increasing demand for virtual machine. The demands for resources will change dynamically, the changes in resource demand is not achieved by the traditional computing because it consumes more time and is not cost-effective. By using cloud services, the resources can be used anywhere and anytime, cloud provides pay-as-you-go style service; as the demands are on-the-fly. Thus cloud service is far better than the traditional computing.

The virtual machine available in the cloud contains different levels of RAM memory and hard drive. The virtual machine can be user specified of different types that are provided on request by the user. The time taken by the cloud to provide this virtual machine should be on-the-fly. If a cloud system, takes time to provide a requested virtual machine, then the cloud vendor has to pay the penalty for violating the Service Level Agreement (SLA) of cloud. Inorder to reduce the time delay for providing the virtual machine provision has to be made to provide virtual machine before it is requested. Planning the capacity of cloud faces many challenges. Since cloud is a dynamic environment, we should not know at what time and how many resources are requested. Two types of problem arises i) Over-estimation ii) Under-estimation.

If the cloud capacity is over estimated and more number of virtual machines are made available and if the user request or demand is less than the available resource, the unused physical resource will be idle. If the cloud capacity is under estimated and if the demand is higher than the available resource, then the requested resource have to be provided that will consume time. This will lead to violate the service level agreement of cloud. To avoid this problem, we have to predict the demands of virtual machine.

LITERATURE SURVEY

Ensemble prediction model is used for prediction. The virtual machine requested by the user contains different features such as VM type, customer id, priority, response time etc. The attribute VM type is used for predicting the future demands. To quantify the prediction error, cloud prediction cost is evaluated prediction cost and cost before prediction of the cloud is also examined. This cpc method will help to quantify the prediction error. The demand history of resources are used for predicting the future demands.

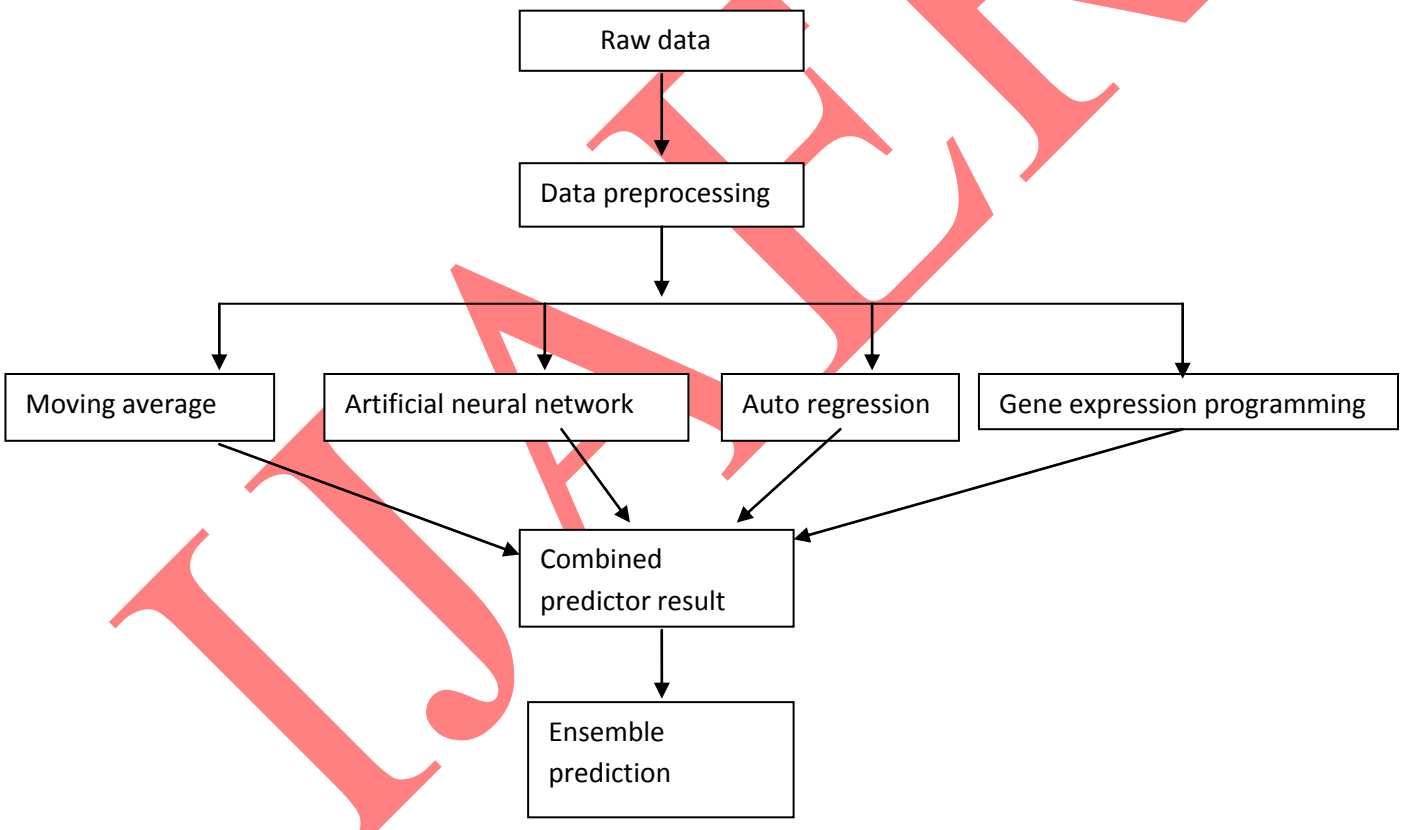
The attributes used for requesting the virtual machine is categorized. A text mining technique has been used for keeping the attribute in a common format. The system logs contains same type of messages. These messages are grouped into a common category. After categorization the rules for each category are formed. Based upon the rules prediction is done. The disadvantage of using Text mining is privacy problem.

Another technique used for mining console logs is Principal Component Analysis. The PCA will identify the normal pattern and make it as a group. By grouping the normal method abnormal method is detected. Thus the result of PCA is given to decision tree. It will helps to understand the detection result. Temporal data mining is another method used for mining the different events. For event summarization a hidden markov model is used. It is based on the learned hidden markov model and achieves short description length and high interpretability. These are the different techniques used for mining of log files.

In this paper, ensemble prediction is used. It uses requested history of virtual machine and prediction is done by mining the log files. Prediction is done by different predictors. The predictors used are moving average, gene expression programming, auto regression, artificial neural network. The result of these predictors will be sent to the ensemble predictor. By combining these results, the Ensemble predictor will give the best result of prediction. The predictors uses the requested history of log files and identify the best parameter combination with the training dataset. Inorder to eliminate the randomness of predictors, the results are computed by averaging 10 runs of each predictors. For each individual predictor, the grid search strategy is to find the best parameter.

Thus taking the average of all the values of best parameter, ensemble method achieves the best performance. In this for each individual customer prediction is done. The model predicts the future provisioning/deprovisioning of virtual machine demands. Also the cloud capacity is estimated with the active virtual machines. By calculating the estimated provisioning demands and deprovisioning demands, the capacity change in one time can be estimated. For calculating the provisioning demand for an individual customer time series technique model is used and global behavior of all the customers are inferred. To calculate the deprovisioning demand temporal characteristics of the virtual machine is used.

System Architecture



The above architecture diagram shows the proposed model. To improve the service quality of cloud more features are proposed to be incorporated into the prediction model. Prediction is done for individual customers, thus improving the prediction accuracy further by providing customer-oriented personalized prediction. For individual customer, the above steps are performed and the result will helps in improving the quality of cloud service. Different attributes for prediction are used, thus reducing the overload for the virtual machine provision.

CONCLUSION

The prediction is done for each individual customer using the history of log files. The customer-oriented personalized prediction helps to know the demands of virtual machine and cloud capacity planning. It will reduce the time delay for preparing virtual machine, thus improving the quality of cloud service.

REFERENCES

[1] W. Peng, T. Li, and S. Ma, "Mining log files for data-driven system management," SIGKDD Explorations, Volume 7, vol. 1, pp. 44–51, 2005.

[2] T. Li, F. Liang, S. Ma, and W. Peng, "An integrated framework on mining logs files for computing system management," in Proc. 2005 KDD, pp. 776–781.

[3] W. Peng, H. Wang, M. Liu, and W. Wang, "An algorithmic approach to event summarization," in Proc. 2010 SIGMOD.

[4] Y. Jiang, C.-S. Perng, T. Li, and R. Chang, "Intelligent cloud capacity management," in 2012 NOMS.

[5] Z. Gong, X. Gu, and J. Wilks, "Press: predictive elastic resource scaling for cloud systems," in 2009 CNSM.

[6] W. Xu, L. Huang, A. Fox, D. Patterson, and M. I. Jordan, "Detecting large-scale system problems by mining console logs," in 2009 SOSP.

[7] J. Kiernan and E. Terzi, "Constructing comprehensive summaries of large event sequences," in Proc. 2008 KDD.