Workload Prediction on Google Cluster Trace

Md. Rasheduzzaman, Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Md. Amirul Islam, Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Rashedur M. Rahman, Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

ABSTRACT

Workload prediction in cloud systems is an important task to ensure maximum resource utilization. So, a cloud system requires efficient resource allocation to minimize the resource cost while maximizing the profit. One optimal strategy for efficient resource utilization is to timely allocate resources according to the need of applications. The important precondition of this strategy is obtaining future workload information in advance. The main focus of this analysis is to design and compare different forecasting models to predict future workload. This paper develops model through Adaptive Neuro Fuzzy Inference System (ANFIS), Non-linear Autoregressive Network with Exogenous inputs (NARX), Autoregressive Integrated Moving Average (ARIMA), and Support Vector Regression (SVR). Public trace data (workload trace version II) which is made available by Google were used to verify the accuracy, stability and adaptability of different models. Finally, this paper compares these prediction models to find out the model which ensures better prediction. Performance of forecasting techniques is measured by some popular statistical metric, i.e., Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Sum of Squared Error (SSE), Normalized Mean Squared Error (NMSE). The experimental result indicates that NARX model outperforms other models, e.g., ANFIS, ARIMA, and SVR.

Keywords: ANFIS, ARIMA, Autocorrelation Function (ACF), Cloud Computing, MAE, MAPE, NARX, NMSE, Partial Autocorrelation Function (PACF), Predictive Model, RMSE, SSE, SVR

1. INTRODUCTION

Big data analytics and cloud computing are currently top of minds for most of the large organizations across the world. To get the valuable insights which can create competitive advantage and drive increased revenues, big data analytics is used now-a-days. On the other hand cloud computing has the potential to enhance business agility and productivity by reducing costs. Cloud computing is a way to increase capacity or add capabilities without investing in new infrastructure. Recently, cloud computing has become more popular to different indus-

DOI: 10.4018/ijghpc.2014070103
tries. System log data, produced by different industries are increasing rapidly. As a result, a growing number of enterprises and data centers are building efficient cloud environments. It make sense that IT organizations should look to cloud computing in order to support their big data projects as cloud computing offers a cost-effective way to support big data technologies. A recent survey conducted by GigaSpaces found that 86% of those IT organizations who think big data processing is important are considering moving their big data analytics to one or more cloud delivery models. The findings from the previous studies in the past decade ensure that, quantity of computing resources and the power efficiency of hardware is not only the main reason for extremely high energy consumption. Inefficient usage of resources is the main reason for extra consumption. Therefore, efficient predictive model is necessary to get prior information about future workload in data centers. Time series analysis can be a good tool to estimate future workload or resource usage.

The main focus of our study is to build a predictive model that can be used for predicting future workload in a cluster system. To find a better predictive model we use different forecasting models like Adaptive Neuro Fuzzy Inference System (ANFIS), Nonlinear Autoregressive network with exogenous inputs (NARX), Autoregressive Integrated Moving Average (ARIMA), and Support Vector Regression (SVR) and finally come up with a solution which performs well for real world cluster data. SVR is suitable for the complex and dynamic characteristics of the cloud computing environment. Experiments with Google trace data have shown that, in comparison with ANFIS, NARX, ARIMA, and SVR, NARX always has the minimum prediction error.

The rest of this paper is organized as follows. Section 2 describes the related work done on predicting future workload using different forecasting models. The system description of the ANFIS, NARX, ARIMA, and SVR model is given in section 3 respectively. In Section 4 we conduct the experiment with these four models. Section 5 also presents a comparative study of the results from those models and finally we draw the conclusion in section 6.

2. RELATED WORKS

A lot of researches have been conducted in forecasting time series. These studies have provided a better understanding of what might happen in future. Wang, Chau, Cheng and Qiu (2009) compared different artificial intelligence (AI) (such as SVM, ARIMA, ANFIS, ANN and GP) methods to build effective hydropower resource management and scheduling. They used four statistical performance measurement techniques to validate different AI models. The authors showed that AI methods are more powerful than traditional time series forecasting models. They authors reported that SVM had better result in both training and validation stage where ANFIS showed different performance in training and validation stage. Genetic Programming (GP) had better result in validation phase.

(Contreras) et. al. (2003) carefully studied the performance impact of the ARIMA model on next day electricity price forecasting. The authors proposed two different ARIMA models for predicting electricity price in Spain and California. In their research they showed that ARIMA model required only 5 hours to forecast future price for Spain, whereas it took 2 hours for prediction the price for California. Daz-Robles et al. (2008) designed a hybrid model by combining ARIMA and ANN model for forecasting air quality in urban area. They showed that traditional Box-Jenkins time series (ARIMA) model and multi-linear regression (MLR) models have limited accuracy. Khashei and Bijari (2010) built a novel predictive model by combing ARIMA and ANN model which gives better accuracy. They tested their model with three real-data set and argued that in case of higher performance this model can be used as appropriate alternative.
Related Content

A Novel Multi-Dimension Resource Recycling Mechanism for Cloud Data Centers
[www.igi-global.com/article/a-novel-multi-dimension-resource-recycling-mechanism-for-cloud-data-centers/136816?camid=4v1a](www.igi-global.com/article/a-novel-multi-dimension-resource-recycling-mechanism-for-cloud-data-centers/136816?camid=4v1a)

E-Portfolio to Promote the Virtual Learning Group Communities on the Grid
[www.igi-global.com/chapter/portfolio-promote-virtual-learning-group/64513?camid=4v1a](www.igi-global.com/chapter/portfolio-promote-virtual-learning-group/64513?camid=4v1a)

Principles of Soft Verification
[www.igi-global.com/article/principles-soft-verification/76920?camid=4v1a](www.igi-global.com/article/principles-soft-verification/76920?camid=4v1a)
Efficient Resource Allocation Mechanism for Federated Clouds
*International Journal of Grid and High Performance Computing* (pp. 74-87).
[www.igi-global.com/article/efficient-resource-allocation-mechanism-for-federated-clouds/141358?camid=4v1a](www.igi-global.com/article/efficient-resource-allocation-mechanism-for-federated-clouds/141358?camid=4v1a)