

Service Performance Pattern Analysis and Prediction of Commercially Available Cloud Providers

Shyam S. Wagle* Mateusz Guzek†, Pascal Bouvry†

University of Luxembourg

6, rue R. Coudenhove-Kalergi, Luxembourg City, Luxembourg

* Email: shyam.wagle.001@student.uni.lu

† Email: [firstname.lastname]@uni.lu

Abstract—The knowledge of service performance of cloud providers is essential for cloud service users to choose the cloud services that meet their requirements. Instantaneous performance readings are accessible, but prolonged observations provide more reliable information. However, due to technical complexities and costs of monitoring services, it may not be possible to access the service performance of cloud provider for longer time durations. The extended observation periods are also a necessity for prediction of future behavior of services. These predictions have very high value for decision making both for private and corporate cloud users, as the uncertainty about the future performance of purchased cloud services is an important risk factor. Predictions can be used by specialized entities, such as cloud service brokers (CSBs) to optimally recommend cloud services to the cloud users. In this paper, we address the challenge of prediction. To achieve this, the current service performance patterns of cloud providers are analyzed and future performance of cloud providers are predicted using to the observed service performance data. It is done using two automatic predicting approaches: ARIMA and ETS. Error measures of entire service performance prediction of cloud providers are evaluated against the actual performance of the cloud providers computed over a period of one month. Results obtained in the performance prediction show that the methodology is applicable for both short-term and long-term performance prediction.

Index Terms—Cloud providers; Pattern Analysis; Performance Prediction; Data Analytics

I. INTRODUCTION

The wide range of similar cloud services offered by the growing number of cloud service providers (CSPs) creates the challenging decision making problem of selection of cloud services. Cloud service users (CSUs) have access to the documented service level agreement (SLA) commitments in their SLA agreements, but they often are more concerned about the actual performance delivery of cloud service providers (CSPs). Cloud providers may fail to provide actual services, as stated in SLA commitments [1]. Service performance patterns of CSPs help cloud users to select appropriate cloud services from multi-cloud architecture according to their quality of service (QoS) requirements. Detailed information (historical, current and future performance) of cloud providers adds confidence to cloud users in decision-making. Due to the dynamic nature of service performance in cloud computing,

there are considerable fluctuations in the QoS, which results in problems of predicting performance of cloud providers. In cloud service selection, historical performance and future performance predictions may be more important than the current, instantaneous performance of cloud providers. Thus, capturing all the performance variability in different performance metrics is important to select the right one among the multiple alternatives of cloud services [2],[3].

Usually, cloud users expect to receive a certain level of service performance as specified in an SLA document. In order to provide the long-range performance information about cloud providers, service performance of cloud providers should be continuously monitored. Due to technical limitations and high cost of monitoring, many cloud users may not be able to continuously monitor service performance of their cloud providers. The comprehensive solutions are to provide users the observed data together with forecasts of future performance of cloud providers.

In this paper, performance of commercially available cloud providers is measured for a month on a daily basis using cloud-monitoring tool. Performance metrics *Uptime*, *Downtime*, *Outage Frequency*, *Latency*, *Response Time* and *Throughput* are considered to observe the performance of the cloud providers. These time series data do not follow the specific trend pattern and is unique for each cloud provider. In such case automatic forecasting method can create appropriate time series models [4]. We applied ETS and ARIMA prediction methods as automatic prediction methods, which gave very convincing cloud performance predictions according to the performance data collected from cloud providers. Comparison of error measures in both methods reveals that both techniques are appropriate for future performance prediction of cloud providers. In overall performance prediction, ARIMA method produced better result than ETS for our dataset collected from multiple cloud providers during one-month period.

The remainder of the paper is organized as follows: Section II provides the related works of our contribution. Prediction methods and accuracy calculation approaches are briefly described in section III. Few deviation patterns and actual patterns of service performance measured from multiple cloud

providers in selected service metrics are shown in section IV. Service performance predictions based on measured performance from cloud providers with possible error calculations are presented in section V and section VI respectively. Section VII concludes the paper with potential future works.

II. RELATED WORKS

Increasing number of cloud providers with similar service offer, SLA commitments, and similar service price make complicated to the cloud user to choose the appropriate cloud providers according to their requirements. To differentiate cloud providers according to their service performance, Wagle et al. [1], [5] have proposed service evaluation techniques to evaluate the performance of cloud providers. Beside the current performance behavior of the cloud provider, it is also very important to know the historical and future service behavior of cloud provider due to dynamic nature of cloud computing environment.

Most of the current research works have addressed QoS forecasting of web-based services. Li et al. [6] propose a web service selection algorithm based on QoS prediction mechanism. Their algorithm uses time series modeling based on structural equations to fit QoS values of web services, and dynamically predicts their future changes to support adaptive services selection. Godse et al. [7] propose a method that combines monitoring and extrapolation methodologies based on ARIMA models to predict service performance. The proposed method is used to support automating dynamic service selection methodology, which is robust in the face of varying QoS.

Other research works in web-based services predictions are focused on QoS management. Nobile et al. [8] propose an architecture that uses ARIMA models in order to predict future traffic characteristics. Zhu et al. [9] present a Grid-based framework that uses a time series prediction algorithm to forecast the future performance of parallel/distributed discrete event simulation (PDES). Zeng et al. [10] have investigated that performance metrics which can be predicted based on their historical data. Amin et al. [11] propose a forecasting approach considering the high volatility of QoS measures and have claimed that it improves the forecasting accuracy of QoS attributes and violations.

Likewise, some research works have addressed workload predictions in cloud computing. Zhang et al. [12], [13] have proposed prediction approach for user's cloud component QoS usage experiences. It forecast the QoS experience of users based on past user's experience. Panneerselvam et. al [14] have analyzed the workload demand of users to reduce the excess resource consumptions of cloud providers. A. Biswas et al. [15], [16] proposes an auto-scaling framework to control enterprises resources coming in to cloud but it does not consider itself the prediction of performances of cloud providers. Syu et al. [17] have applied Genetic Programming for time-aware dynamic QoS prediction. Calheiros et al. [18] have proposed workload prediction method using ARIMA method

and have analyzed impact of it in QoS prediction in cloud computing.

To the best of our knowledge, service performance pattern analysis including future performance prediction of cloud providers using real monitoring data is missing in the current research. Recent and future behavior of cloud provider helps in decision making to select appropriate cloud services to cloud users. These research works mentioned in this section show that it is crucial research issue to aware the cloud users with the current and future service performance pattern of cloud providers to select appropriate cloud services by cloud users. In this work, ETS and ARIMA forecasting methods have been implemented to predict the future behaviour in the service performance of the cloud providers.

III. PREDICTION METHODS AND PREDICTION ACCURACY

Prediction methods are broadly divided into qualitative and quantitative. Qualitative forecasting techniques are subjective, based on the opinion and judgment of consumers or experts. They are appropriate when past data is not available. Quantitative predictions are appropriate when past numerical data is available and are in some reasonable patterns. Delphi method, market research, and historical life-cycle analogy are some example of qualitative forecasting method whereas time series methods, and causal methods are some examples of quantitative forecasting method. Due to many drawbacks in simple and weighted moving average of quantitative prediction, exponentially smoothing methods are widely used in forecasting. The choice of prediction method is often constrained by data availability and data pattern. The pattern in the data will affect the type of forecasting method selected. The pattern in the data will also determine whether a time-series method will suffice or whether casuals model are needed. If the data pattern is unstable over time, a qualitative method may be selected. Thus the data pattern is one of the most important factors affecting the selection of a forecasting method [19]. Data may not follow the specific pattern in all the cases. In these circumstances, an automatic forecasting method is essential, which determines an appropriate time series model, estimate the parameters and compute the forecasts [4]. The most popular automatic forecasting algorithms are based on either exponential smoothing or ARIMA methods.

A. Exponential Smoothing

The exponential method involves the automatic weighting of past data with weights that decrease exponentially with time, i.e. the most current values receive a decreasing weighting. For example in each increment in the past is decreased by $(1-\alpha)$, where $\alpha \in (0,1)$ is the smoothing parameter. Generally, there are three exponential smoothing are in practice: Simple, Double and Triple exponential smoothing. The triplet(E,T,S) refers to the three components: error, trend and seasonality. So the model ETS(A,A,N) has additive errors, additive trend and no seasonality and so on. ETS can also be considered an abbreviation of ExponenTial Smoothing [4].

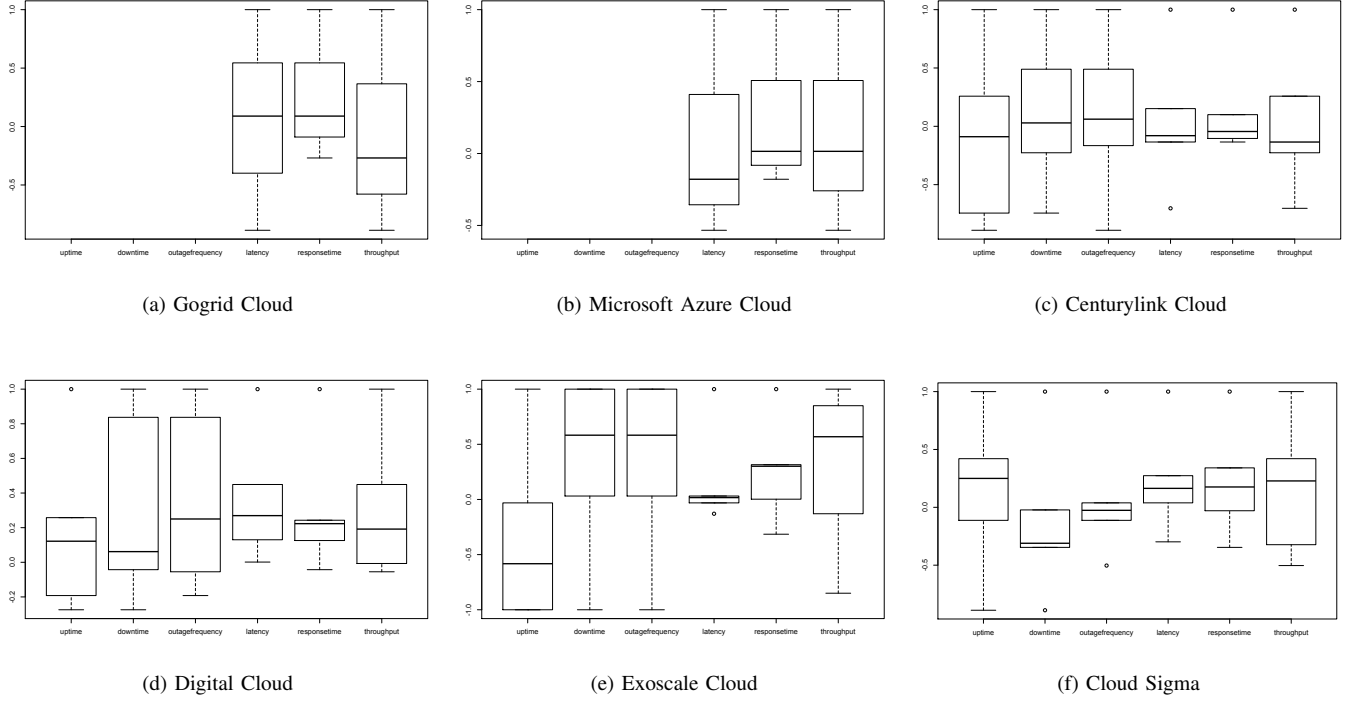


Fig. 1. Service Performance Deviation Pattern of CSPs

B. Autoregressive integrated moving average (ARIMA)

It is a generalization of an autoregressive moving average (ARMA) model [20] where data show evidence of non-stationary. ARIMA models are generally denoted ARIMA (p, d, q)(P,D,Q)_m where parameters p, d, and q are non-negative integers, p is the order of the autoregressive model, d is the degree of differencing, and q is the order of the Moving-average model. Furthermore, m refers to the number of periods in each season, and the uppercase P, D, Q refer to the autoregressive, differencing, and moving average terms.

C. Prediction Accuracy

Let y_t denotes the observation at time t and f_t denote the forecast of y_t . Then, define the forecast error $e_t = y_t - f_t$. The forecasts may be computed from a common base time, and be of varying forecast horizons [21]. Thus, we may compute out-of-sample forecasts f_{n+1}, \dots, f_{n+m} based on data from times $t = 1, \dots, n$. There are commonly used accuracy measures in the forecasting:

1) *Scale-dependent measures*: These accuracy measures are dependent with the scale of the data. These are useful when comparing different methods applied to the same set of data, but should not be used, for example, when comparing across data sets that have different scales. The most commonly used scale-dependent measures are based on the absolute error or squared errors:

$$\text{MeanSquaredError}(MSE) = \text{mean}(e_t^2) \quad (1)$$

$$\text{MeanAbsoluteError}(MAE) = \text{mean}(|e_t|) \quad (2)$$

$$\text{RootMeanSquareError}(RMSE) = \sqrt{e_t^2} \quad (3)$$

2) *Measures based on percentage errors*: The percentage error is given by $p_t = 100e_t/y_t$. Percentage errors have the advantage of being scale independent, and so are frequently used to compare forecast performance across different data sets. The most commonly used measures are:

$$\text{MeanPercentageError}(MPE) = \text{mean}(p_t) \quad (4)$$

$$\text{MeanAbsolutePercentageError}(MAPE) = \text{mean}(|p_t|) \quad (5)$$

Measures based on percentage errors have the disadvantage of being infinite or undefined if $y_t = 0$ for any t in the period of interest, and having an extremely skewed distribution when any y_t is close to zero.

3) *Scaled errors*: To make independent with the scale of the data Hyndman et. al [21] proposed scale independent error measure called Mean Absolute Scaled Error (MASE). It is used to determine the success of a model selection procedure.

$$MASE = \text{mean}(|q_t|) \quad (6)$$

where $q_t = e_t / ((1/(n-1)) \sum_{i=2}^n |y_i - y_{i-1}|)$. When $MASE \leq 1$, it gives smaller errors than the one-step errors from other scale dependent measures in the proposed method.

4) *Information Criteria(IC)*: Information criteria are chosen to choose the best predictive model selection. It is useful in comparison to IC value for another model fitted to same data set.

Akaike's Information Criteria (AIC):

$$AIC = -2\log(\text{Likelihood}) + 2p \quad (7)$$

The AIC [22] provides a method for selecting between the additive and multiplicative error models. Point forecasts from the two models are identical, so that standard forecast accuracy measures such as the MSE or MAPE are unable to select between the error types. The AIC is able to select between the error types because it is based on likelihood rather than one-step forecasts. Minimizing the AIC gives the best model for prediction.

Schwartz's Bayesian IC(BIC):

$$BIC = AIC + p(\log(n) - 2) \quad (8)$$

The BIC[23] is used to overcome the inconsistency and over fitted problem in AIC, BIC is used. It is also used in a similar manner like AIC.

$$AICc = AIC + 2(p + 1)(p + 2)/(n - p) \quad (9)$$

where p is the number of estimated parameters in the model. The AICc[24] is an asymptotically efficient information criterion that does an approximate correction for this negative bias. It is also used in a similar manner like AIC.

We have considered all the errors measures in this paper to evaluate the prediction accuracy. ACF1(Autocorrelation of errors at lag 1) is also considered for the error measurements.

IV. PERFORMANCE MEASUREMENT

To observe the performance of cloud providers, six major service parameters/metrics are considered in the measurement of the performance of the cloud providers: *Uptime*, *Downtime*, *Outage Frequency*, *Latency*, *Response Time* and *Throughput*. Performance of each cloud providers are obtained using cloud monitoring tool¹ according to the selected service metrics. In our observation, service performance of 20 cloud providers are collected for 30 days: *Amazon S3*, *GMO Cloud*, *City Cloud*, *Google Cloud Storage*, *Gogrid Cloud*, *Rackspace Cloud*, *Centurylink Cloud*, *UpCloud*, *Softlayer Cloud*, *IBM Cloud*, *HP Cloud*, *Vault Network Cloud*, *Microsoft Azure Cloud*, *Digital Cloud*, *Elastic Host Cloud*, *Exoscale Cloud*, *Sigma Cloud*, *Cloud Central*, *Aruba Cloud* and *Baremetal Cloud*. The performance measurement is based on cloud storage services. It is considered that all cloud service users are located in Luxembourg. Service/availability regions of cloud providers are divided in different regions according to data center locations of the cloud providers. Collected service performance data from cloud providers are merged values of all service/availability regions of the cloud providers.

The analysis of performance of chosen cloud providers on selected metrics shows that *uptime*, *downtime*, *outagefrequency* of most cloud providers shows less variance, while *latency*, *responsetime* and *throughput* present more variance and hard to predict performance patterns easily. Figure 1 shows

the performance deviation of some cloud providers among our list of observations (*Gogrid Cloud*, *Microsoft Azure Cloud*, *Centurylink Cloud*, *Digital Cloud*, *Exoscale Cloud* and *Cloud Sigma*). These are some representative patterns of both less fluctuated values in performance parameters and highly fluctuated values in service performance parameters. For example, performance parameters: *uptime*, *downtime*, *outagefrequency* are stable near in *100%*, *0 second*, *0 times* in cloud provider *Gogrid Cloud* and *Microsoft Azure Cloud* respectively but these three service parameters were also fluctuating for cloud providers: *Centurylink Cloud*, *Digital Cloud*, *Exoscale Cloud* and *Sigma Cloud* together with other three service parameters: *latency*, *responsetime* and *throughput*. Service performance patterns in parameters: *uptime*, *downtime* and *outage frequency* in most of the cloud providers are very close in most of the time series and quite unpredictable in *latency*, *responsetime* and *throughput* (See, for instance, in (a) of Figures 2 to 4 of Exoscale Cloud and Figures 5 to 7 of Digital Cloud respectively). Due to space limit performance of few cloud providers on selected metrics are included in this paper, however, performance prediction accuracy of all cloud providers are presented in the Section VI. In the observation, performance parameters *latency*, *responsetime* and *throughput* followed high time dependent service variation than performance parameters *uptime*, *downtime* and *outagefrequency*. In the first column (a) of Figures 2 to 7 give the daily performance values of *Exoscale Cloud* and *Digital Cloud*.

V. SERVICE PERFORMANCE PREDICTION

The service patterns of all the selected cloud providers give the clear picture of cloud service performance of the cloud providers over a month. The performance of cloud providers collected over the month from multiple cloud providers neither followed specific pattern nor remained in stable seasonal patterns for specific time periods. Information of service delivery pattern of cloud providers of entire month (possibly delivery pattern of longer time period e.g. 3 months, 6 months, 1 year or more) gives the tentative impression of cloud service providers to choose appropriate cloud services for the cloud users. From monitoring cost, technical complexities and accessibility to the cloud providers through out the period, it is very challenging to collect the service pattern of cloud provider over the longer time period. Highly efficient prediction of data solves this problem to receive the service performance information for longer terms. To achieve the maximum accuracy in cloud service performance prediction, we applied automatic forecasting method to determine the appropriate time series prediction, particularly; ETS and ARIMA method for the cloud providers performance prediction.

The examples of predicted performance patterns of cloud providers using ETS and ARIMA methods are shown in (b) and (c) of Figures 2 to 7 of *Exoscale Cloud* and *Digital Cloud* respectively. We chose predicted patterns of *Exoscale Cloud* and *Digital Cloud* because their prediction patterns represented almost all the scenarios of predictions (best predictions and worst predictions). Red line gives the actual patterns of the

¹cloudharmony.com

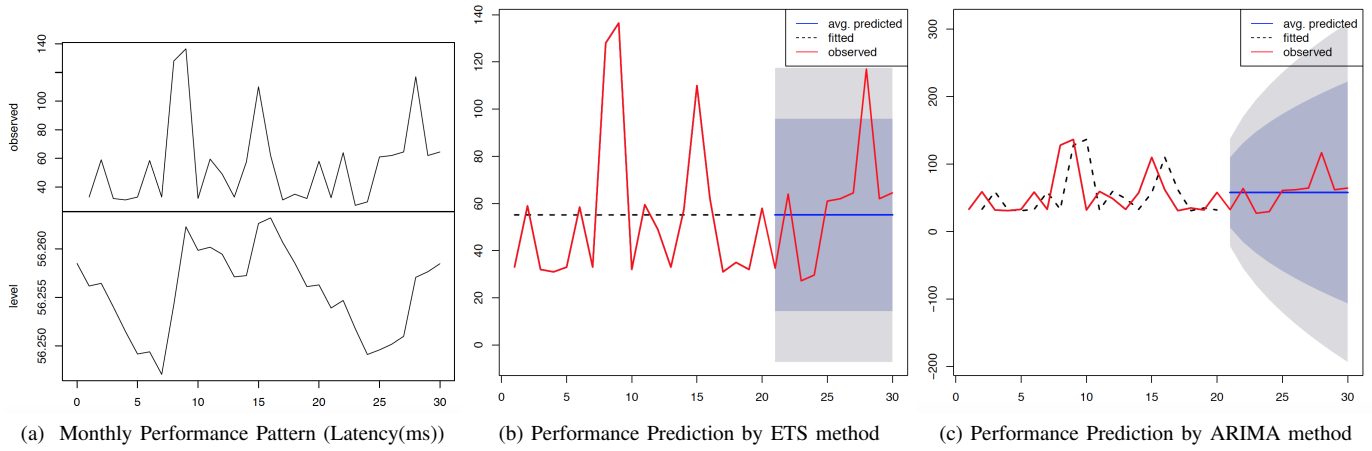


Fig. 2. Service Performance Prediction of Latency of Exoscale Cloud

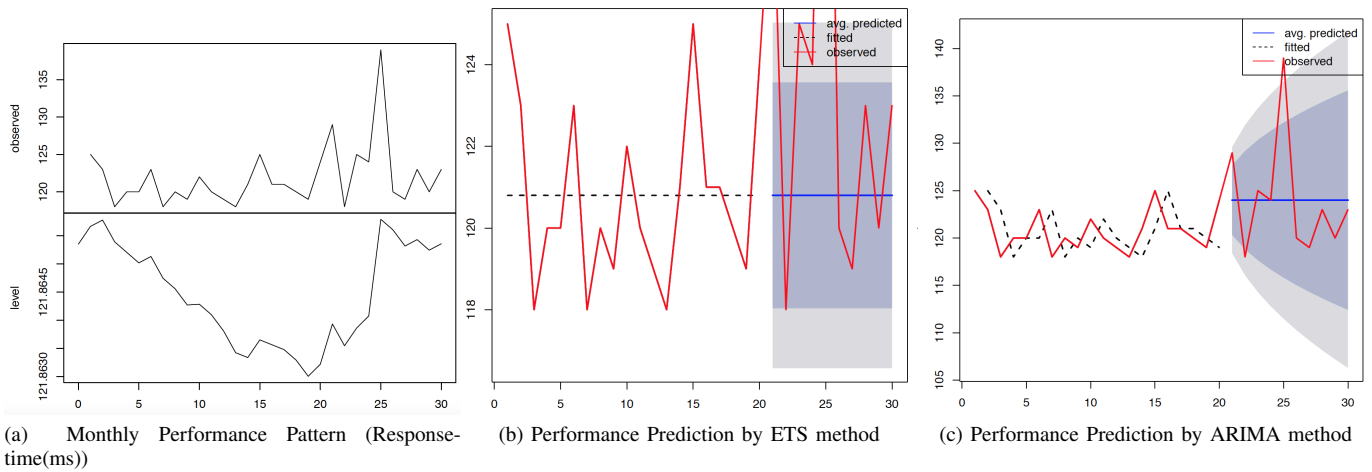


Fig. 3. Service Performance Prediction of Responsetime of Exoscale Cloud

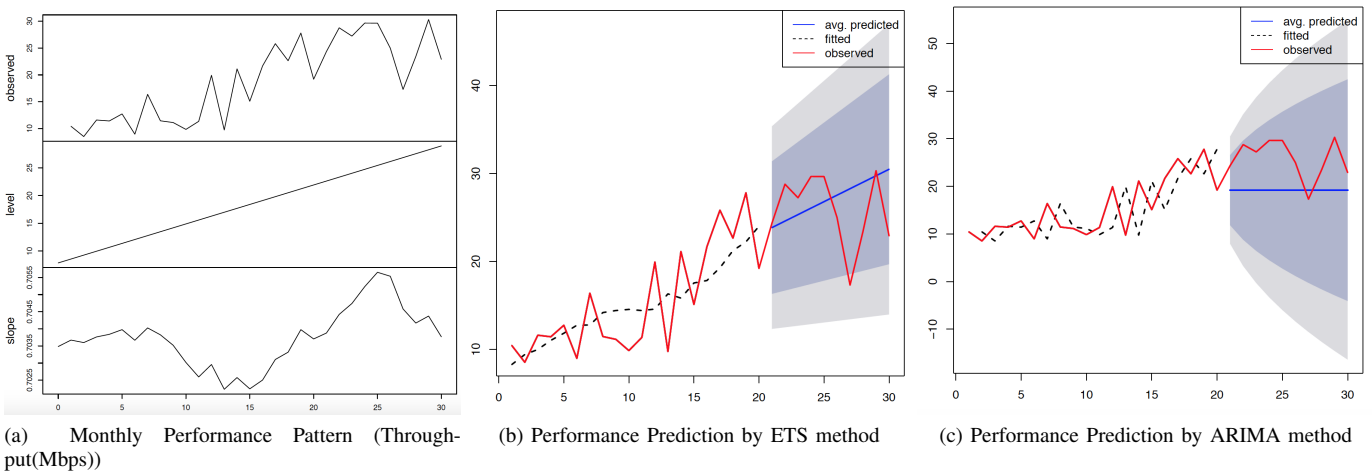


Fig. 4. Service Performance Prediction of Throughput of ExoscaleCloud

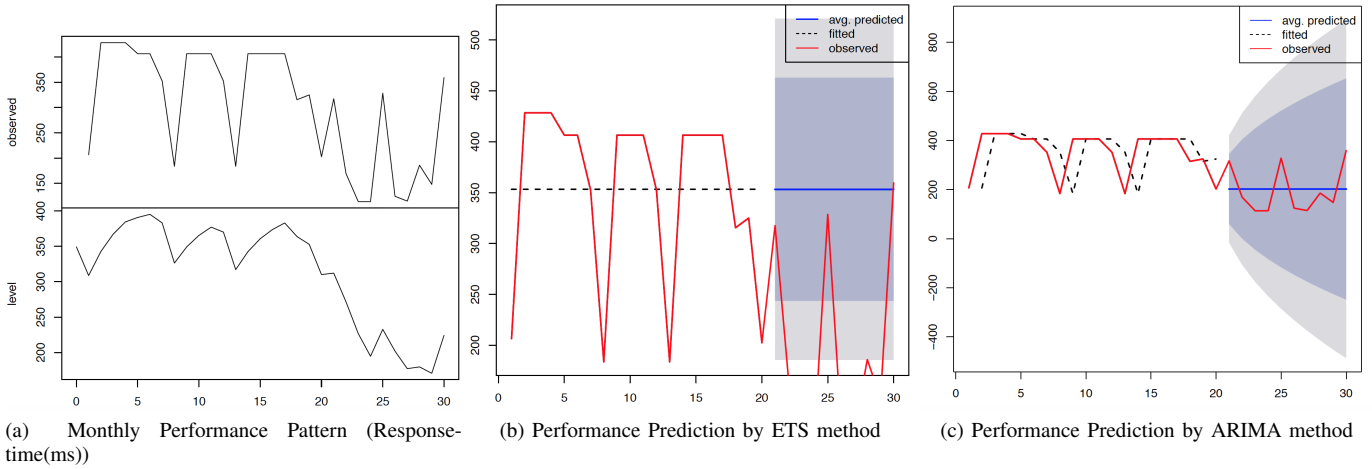


Fig. 5. Service Performance Prediction of Latency of Digital Cloud

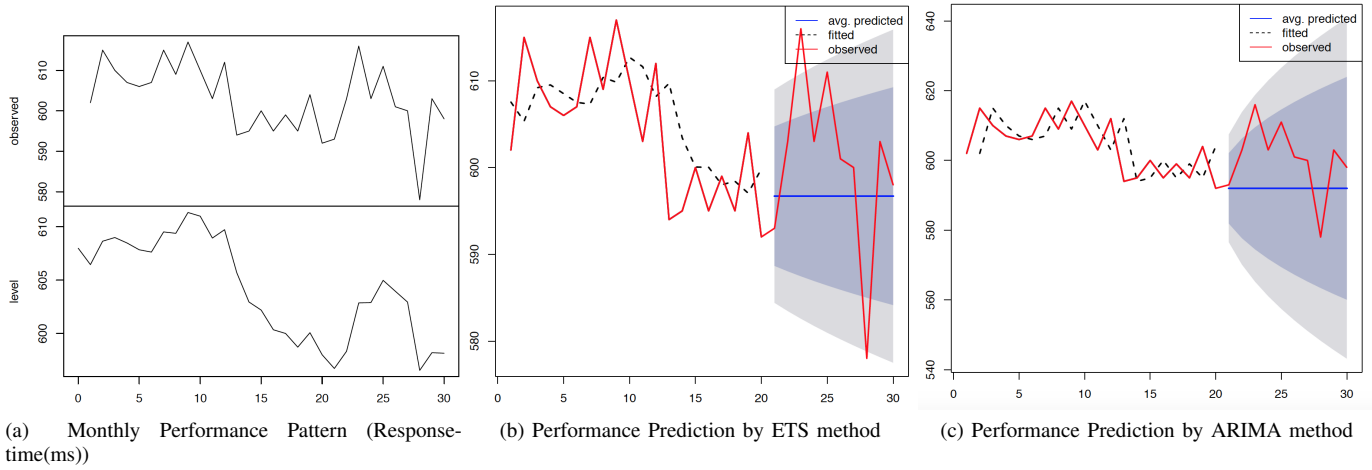


Fig. 6. Service Performance Prediction of Responsetime of Digital Cloud

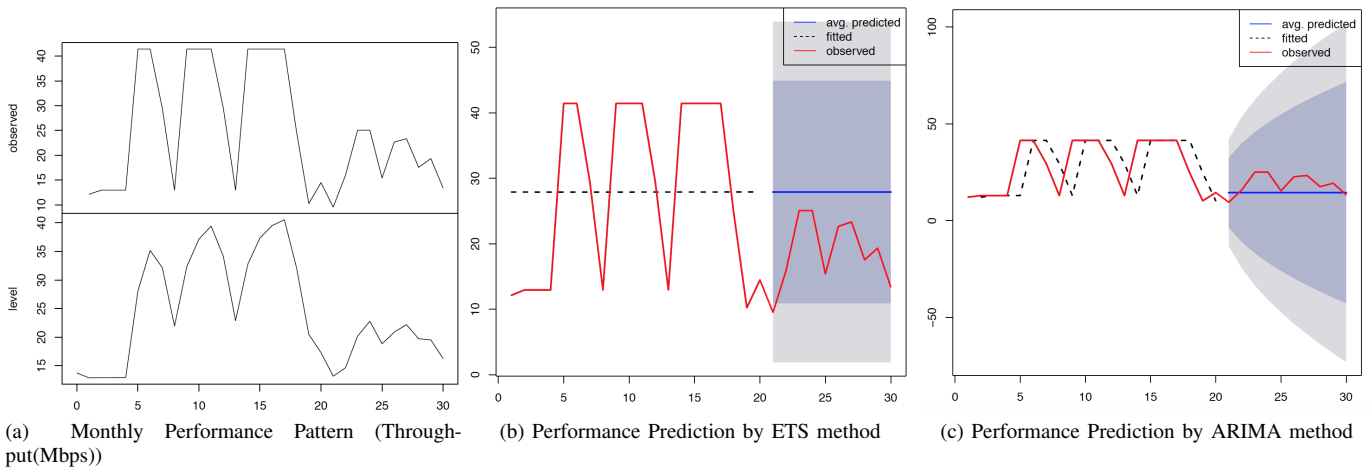


Fig. 7. Service Performance Prediction of Throughput of Digital Cloud

service performance, dark blue line gives the average predicted performance (point of forecast) and dotted black line gives the fitted patterns of actual performance measured of 20 days. The forecast intervals are at 80% to 95%. Lower forecast range at 80%(Low80) to high forecast range at 80%(High80) is shown in light blue background and 95% forecast interval (Low95 (lower forecast range at 95%) to High95 (high forecast range at 95%)) is shown in light black background give the minimum to maximum range of prediction of cloud provider performance. To evaluate the accuracy of the prediction of the CSPs' service performance, we trained the system by 20 days performance data collected from different cloud providers.

For instance, actual performance pattern of *latency* of *Exoscale Cloud* follows the lower range (Low80) to higher range(High95) of prediction in ETS prediction method (See in (b) of Figure 2) whereas observed performance pattern followed with average predicted pattern in ARIMA method (See in (c) of Figure 2). Observed performance pattern of *responsetime* of *Exoscale Cloud* follows in between Low80 to High95 in both prediction techniques (See in (b) and (c) of Figure 3). Likewise, in (b) and (c) of Figure 4, *throughput* of *Exoscale Cloud*, average predicted line (dark blue) exactly followed the fitted pattern (black dotted line) of observed data of 20 days in continuously increasing manner in ETS method but observed patterns followed the Low85 range of predicted pattern in ARIMA method.

In the analysis of performance prediction pattern of *Digital Cloud*, predicted pattern of *latency* by ARIMA method is more close to observed pattern than the predicted pattern in ETS method (See in (b) and (c) of Figure 5). Observed measurement of *responsetime* of *Digital Cloud* is out of prediction range for around 2-3 days in ETS method whereas observed performance pattern is within the range of Low80 in ARIMA method (See in (b) and (c) of Figure 6). See, for instance, in (b) and (c) of Figure 7 observed measurement pattern of *throughput* followed the almost same pattern as average predicted pattern by ARIMA method but observed measured value range is in between average predicted value to Low80 value in ETS method.

As explained, both prediction methods showed convincing predictions in each performance parameters. Service performance data of 20 days are trained for the performance prediction using ETS and ARIMA prediction method. Prediction of all the selected performance parameters of 20 cloud providers by both prediction methods produced the convincing results as in *Digital Cloud* and *Exoscale Cloud*. To analyze the accuracy of our prediction, different error measurements are calculated (See for instance Table II and I for *Exoscale Cloud* and *Digital Cloud* respectively).

VI. ERRORS MEASUREMENT AND PREDICTION ACCURACY

To check the accuracy of prediction method, different error parameters are considered. In the performance prediction, all error parameters do not return the exact accuracy of the prediction method because of different scale of the measured

data. So, scale dependent errors, percentage errors and scale independent errors are considered to check the accuracy of the prediction. Accuracy information correction parameters (AIC, BIC, AICc) are used to compare the prediction methods. Smaller the numerical value of information parameters signifies the better prediction method.

Error measures of all performance parameters of all 20-cloud providers are calculated. Prediction results of *uptime*, *downtime* and *outage frequency* are more stable in most of the cloud providers than the rest of the parameters: *latency*, *responsetime* and *throughput*. We chose *Digital Cloud* and *Exoscale Cloud*, to cover all the parameters prediction analysis, because they both have unstable service performance result in all the selected parameters. MASE value of in all predictions was less than or equal to 1 in both ETS and ARIMA prediction technique (except in parameter *throughput* in ETS predictions with slightly greater than 1) (See Tables I and II). It shows that our selected prediction methods are correct. Information correction parameters measured in ETS and ARIMA method shows that ARIMA method is more suited than ETS method (See for instance in Table II where all values of information correction (IC) have less value in ARIMA method than ETS), however, IC values in each prediction methods are very close to each other.

To evaluate our prediction accuracy, we calculated all error parameters as we have included in Table I and II for all 20 selected cloud providers in selected performance metrics. Error patterns of *RMSE*, *AIC*, *BIC* and *AICc* follow the similar pattern with minimum differences in error values in ARIMA and ETS prediction method whereas error patterns are significantly different in error measurement *MASE* and *MAPE* (See Figure 8).

However, majority of error parameters, for instance, *ME*, *MAE* are near to "zero", we considered *RMSE*, *MASE* and *MAPE* and information corrections parameters (AIC, BIC, AICc) to evaluate the accuracy of the prediction method because scaled and percentage error may not always return accurate information of prediction accuracy. We did not include full range of Y-axis in Figure 8 (to be fitted in the scale limit of the figure) as it only gives density of error values.

Most of the error values are near to zero in most of the prediction in *RMSE* (See in (a) of Figure 8). Overall information corrections parameters (AIC, BIC, AICc) are lower in ARIMA prediction method than ETS prediction method (See in (d),(e) and (f) of Figure 8). Similarly, *MASE* and *MAPE* have higher values in ETS method than ARIMA method (See in (b) and (c) of Figure 8). Prediction patterns and related errors calculated to evaluate the accuracy of prediction method shows that both prediction methods (ARIMA and ETS) are suitable to predict performance of cloud providers according to our data patterns collected from multiple cloud providers. However, in comparisons with ARIMA and ETS, ARIMA prediction method gives more smooth prediction than ETS method.

In both prediction methods, majority of *MAPE* values range from 0.4 to 1.4 (in ETS) and -0.4 to 2.5 (in ARIMA), however,

in very few worst scenario it reached to very high value. *MASE* values ranged in between -0.8 to 0.4 (in ARIMA) and -0.85 to 0.45 (in ETS).

VII. CONCLUSIONS AND FUTURE WORKS

The paper presents the service performance pattern analysis and prediction of commercially available cloud providers. It provides the forecasts of future service performance of cloud service providers. The input for the method is the previously observed performance. The method selected to produce the predictions are ETS and ARIMA. For evaluation, the real monitoring data was divided into training and test sets. Both prediction method returned the convincing results of performance prediction according to calculated errors in prediction; however, ARIMA method gave better performance prediction results than ETS method, as shown by the analysis of the errors of the prediction methods. The method presented in this paper summarizes the current and future service performance of cloud providers for the selected performance metrics. It helps cloud service users and brokers to choose cloud services according to their requirements. Predicted service performance results show that prediction is applicable for short duration prediction as well as long-term duration. The future work includes evaluation of service performance over longer time periods and applying any other prediction methods such as machine learning regression methods and/or neural networks. It is also planned to include more commercial cloud provider to analyze the service delivery patterns and their future service performance behavior based on observed value from cloud providers.

ACKNOWLEDGMENT

This work was partially funded from National Research Fund (FNR) ISHOP (POLLUX/13/IS/6466384) project. S. Wagle would also like to thank Law, Science and Technology (LAST-JD) project for financially supporting to perform this research.

REFERENCES

- [1] S. S. Wagle, M. Guzek, P. Bouvry, and R. Bisdorff, "An evaluation model for selecting cloud services from commercially available cloud providers," in *IEEE 7th International Conference on Cloud Computing*, Vancouver, Canada, Nov. 2015, pp. 107–114.
- [2] Z. Zheng, X. Wu, Y. Zhang, M. R. Lyu, and J. Wang, "Qos ranking prediction for cloud services," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1213–1222, June 2013.
- [3] X. Chen, Z. Zheng, and M. R. Lyu, *QoS-Aware Web Service Recommendation via Collaborative Filtering*. New York, NY: Springer New York, 2014, pp. 563–588. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-7518-7_22
- [4] R. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for r," *Journal of Statistical Software*, vol. 27, no. 1, pp. 1–22, 2008.
- [5] S. S. Wagle, M. Guzek, and P. Bouvry, "Cloud service providers ranking based on service delivery and consumer experience," in *IEEE 4th International Conference on Cloud Networking (CloudNet)*, Niagara Falls, Canada, Oct. 2015, pp. 202–205.
- [6] M. Li, J. Huai, and H. Guo, "An adaptive web services selection method based on the qos prediction mechanism," in *Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT '09. IEEE/WIC/ACM International Joint Conferences on*, vol. 1, Sept 2009, pp. 395–402.
- [7] M. Godse, U. Bellur, and R. Sonar, "Automating qos based service selection," in *Web Services (ICWS), 2010 IEEE International Conference on*, July 2010, pp. 534–541.
- [8] P. N. Nobile, R. R. F. Lopes, C. E. Moron, and L. C. Trevelin, "Qos proxy architecture for real time rpc with traffic prediction," in *Distributed Simulation and Real-Time Applications, 2007. DS-RT 2007. 11th IEEE International Symposium*, Oct 2007, pp. 261–267.
- [9] S. Zhu, Z. Du, Y. Chen, X. Chai, and B. Li, "Qos enhancement for pdes grid based on time series prediction," in *Proceedings of the Sixth International Conference on Grid and Cooperative Computing*, ser. GCC '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 423–429. [Online]. Available: <http://dx.doi.org/10.1109/GCC.2007.99>
- [10] L. Zeng, C. Lingenfelder, H. Lei, and H. Chang.
- [11] A. Amin, A. Colman, and L. Grunske, "An approach to forecasting qos attributes of web services based on arima and garch models," in *Web Services (ICWS), 2012 IEEE 19th International Conference on*, June 2012, pp. 74–81.
- [12] Y. Zhang, Z. Zheng, and M. R. Lyu, "Exploring latent features for memory-based qos prediction in cloud computing," in *Reliable Distributed Systems (SRDS), 2011 30th IEEE Symposium on*, Oct 2011, pp. 1–10.
- [13] —, "Real-time performance prediction for cloud components," in *Object/Component/Service-Oriented Real-Time Distributed Computing Workshops (ISORCW), 2012 15th IEEE International Symposium on*, April 2012, pp. 106–111.
- [14] J. Panneerselvam, L. Liu, N. Antonopoulos, and Y. Bo, "Workload analysis for the scope of user demand prediction model evaluations in cloud environments," in *Utility and Cloud Computing (UCC), 2014 IEEE/ACM 7th International Conference on*, Dec 2014, pp. 883–889.
- [15] A. Biswas, S. Majumdar, B. Nandy, and A. El-Haraki, "Predictive auto-scaling techniques for clouds subjected to requests with service level agreements," in *2015 IEEE World Congress on Services*, June 2015, pp. 311–318.
- [16] —, "An auto-scaling framework for controlling enterprise resources on clouds," in *Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on*, May 2015, pp. 971–980.
- [17] Y. Syu, Y. Y. Fanjiang, J. Y. Kuo, and S. P. Ma, "Applying genetic programming for time-aware dynamic qos prediction," in *2015 IEEE International Conference on Mobile Services*, June 2015, pp. 217–224.
- [18] R. N. Calheiros, E. Masoumi, R. Ranjan, and R. Buyya, "Workload prediction using arima model and its impact on cloud applications x2019; qos," *IEEE Transactions on Cloud Computing*, vol. 3, no. 4, pp. 449–458, Oct 2015.
- [19] J. B. Guerard, *An Introduction to Time Series Modeling and Forecasting*. New York, NY: Springer New York, 2013, pp. 47–72. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-5239-3_3
- [20] E. J. HANNAN and J. RISSANEN, "Recursive estimation of mixed autoregressive-moving average order," vol. 69, no. 1, pp. 81–94, 1982.
- [21] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, pp. 679–688, 2006.
- [22] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, Dec 1974.
- [23] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, no. 2, pp. 461–464, 03.
- [24] C. M. HURVICH and C.-L. TSAI, "Bias of the corrected aic criterion for underfitted regression and time series models," vol. 78, no. 3, pp. 499–509, 1991.