Accepted Manuscript

Performance prediction model for cloud service selection from smart data

Abdullah Mohammed Al-Faifi, Biao Song, Mohammad Mehedi Hassan, Atif Alamri, Abdu Gumaei

PII: S0167-739X(18)30014-1
DOI: https://doi.org/10.1016/j.future.2018.03.015
Reference: FUTURE 4027

To appear in: *Future Generation Computer Systems*

Received date : 2 January 2018
Revised date : 23 February 2018
Accepted date : 6 March 2018

Please cite this article as: A.M. Al-Faifi, B. Song, M.M. Hassan, A. Alamri, A. Gumaei, Performance prediction model for cloud service selection from smart data, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.03.015

Performance Prediction Model for Cloud Service Selection from Smart Data

Abdullah Mohammed Al-Faifi[a, b], Biao Song[b], Mohammad Mehedi Hassan[a, b,1], Atif Alamri[a, c] and Abdu Gumaei[d]

[a] *Chair of Pervasive and Mobile Computing, Collage of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia*

[b] *Department of Information Systems, King Saud University, Riyadh 11543, Saudi Arabia*

[c] *Department of Software Engineering, King Saud University, Riyadh, 11543, Saudi Arabia*

[d] *Department of Computer Science, King Saud University, Riyadh, 11543, Saudi Arabia*

Email: abdullah_ah50@yahoo.com, {bsong, mmhassan, atif}@ksu.edu.sa, abdugumaei@gmail.com

Abstract

Cloud computing is a computing model that has experienced significant growth in the world in contemporary time. Cloud providers offer services to consumers at different levels of performance, costs, and configurations. Many enterprises and organizations are planning to move their services to a cloud platform. The most challenging issue for them is choosing the most appropriate services that meet their requirements. In this paper, we try to tackle this challenge by automating the selection process based on actual workload pattern from Smart data and resource demand acquired from existing service history data. An automatic performance prediction model based on Naïve Bayes classifiers is proposed to predict the performance metrics of cloud nodes with respect to different options for configuration of their resources. We examined Naïve Bayes classifier along with kernel density estimation to solve the zero variance of feature distribution and enhance the accuracy of predictions. We also evaluated our model using a detailed one-year dataset from a realistic environment with thousands of records and hundreds of machines. A simulation on the MATLAB was performed and the results showed that the proposed model indicates how naïve Bayes can provide accurate and efficient results.

Keywords— Cloud service selection, Smart data, Bayes classifier, Performance prediction

1. Introduction

Cloud computing is emerging as a new and popular paradigm in computing and storage resource allocation technology over the Internet. This paradigm of computing offers significant benefits to businesses, government agencies, and the general population by relieving them of low-level tasks related to setting up IT infrastructure as well as allowing organizations to begin small and increase resources only when there is a high demand for service, thus enabling more time for innovation and the creation of business value [18]. Cloud computing is a model for offering a pool of services like software as a service, infrastructure as a service, and platform as a service, all of which are provided on demand.

[1]Corresponding Author
Email address: **mmhassan@ksu.edu.sa** (Mohammad Mehedi Hassan)