The Journal of Systems and Software 134 (2017) 279–303

Contents lists available at ScienceDirect

The Journal of Systems and Software

journal homepage: www.elsevier.com/locate/jss

Time series forecasting for dynamic quality of web services: An empirical study

Yang Syu[a], Jong-Yih Kuo[b], Yong-Yi Fanjiang[c,*]

[a] Institute of Information Science, Academia Sinica, Taipei 115, Taiwan, ROC
[b] Department of Computer Science and Information Engineering, National Taipei University of Technology, Taipei 106, Taiwan, ROC
[c] Department of Computer Science and Information Engineering, Fu Jen Catholic University, New Taipei City 24205, Taiwan, ROC

ARTICLE INFO

Article history:
Received 15 May 2016
Revised 24 August 2017
Accepted 12 September 2017
Available online 15 September 2017

Keywords:
Web service
QoS prediction
Time-series forecasting

ABSTRACT

Web Services (WSs) constitute a critical component of modern software development. Knowing the dynamic qualities of WSs is mandatory during use, and these qualities vary continuously over time. However, in most cases, the quality information furnished by service providers is static and does not consider dynamic variations in quality over time. Thus, it is necessary to determine a method for acquiring accurate quality values for WSs. The motivation for this research is that the most suitable time-series method for dynamic quality prediction of WSs remains unknown because no comprehensive empirical comparison of the representative time-series methods has been performed. Therefore, in this paper, we implement all the representative time-series methods and compare their dynamic quality predictions for WSs using a real-world quality dataset. For empirical comparison, we have ensured that our study is reproducible and referenceable by providing diverse specifics and evaluating their validity in detail. The experimental results and diverse discussions presented in this paper may act as a valuable reference to both academic researchers and WS consumers and providers in industry because they can depend on the results when selecting the most suitable time-series method for direct use or as a starting point for further modifications. Based on our experimental results, among the included time-series forecasting approaches, genetic programming (GP) can achieve the highest quality of service (QoS) forecasting accuracy (in our experiments, the average mean absolute error and mean absolute percentage error are 1258 and 20%, respectively); however, this approach also requires the longest time to produce a QoS predictor (67.7 s on average). Though auto-regressive integrated moving average (ARIMA, with average error measures of 1343 and 25.4%) and exponential smoothing (ES, with average error measures of 1354 and 25.7%) present slightly worse accuracy than GP, ARIMA and ES require much less time to generate a predictor than GP (on average, 0.1612 and 0.1519 s, respectively); thus, these approaches might represent a compromise between forecasting accuracy and cost.

1. Introduction

Web Services (WSs) include both those based on SOAP/WSDL and those based on REST (Lee et al., 2014). WSs are integral components of modern software engineering and application/information system development and have become increasingly popular and prevalent in software development because of their numerous advantages, including interoperability, reusability, composability, and near global accessibility (they are available on and accessible via the Web) (Fanjiang et al., 2015). These advantages have led diverse organizations (e.g., banks) to expose their businesses and provide services through WSs. Moreover, many large IT companies, such as *Google* and *Facebook*, also provide their rich services and applications in the form of WSs. Given these circumstances, it is little wonder that a large number of WSs are currently available on the Web.

The large number of available WSs leads to the phenomenon that when searching for a web service (WS) with a specific functionality, there will often be multiple functionality-equivalent WSs that match and satisfy the specified functional requirements. Therefore, in both research and industry, the convention is to select the most qualified WS based on other, non-functional properties. Thus far, when performing such selection, many different non-functional properties of WSs have been considered and deeply studied including those reviewed in Syu et al. (2012), Syu and Fanjiang (2013), Zhovtobryukh (2007), and Yang et al. (2011). Among

* Corresponding author.
E-mail addresses: a29066049@gmail.com (Y. Syu), jykuo@ntut.edu.tw (J.-Y. Kuo), yyfanj@csie.fju.edu.tw, yyfanj2@gmail.com (Y.-Y. Fanjiang).

http://dx.doi.org/10.1016/j.jss.2017.09.011
0164-1212/

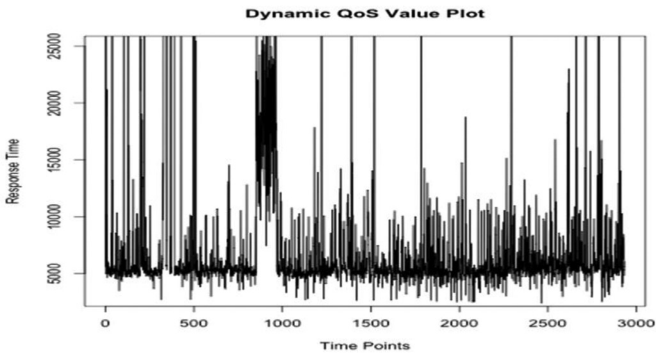

Fig. 1. The response time of a real-world commercial Amazon WS at different time points.

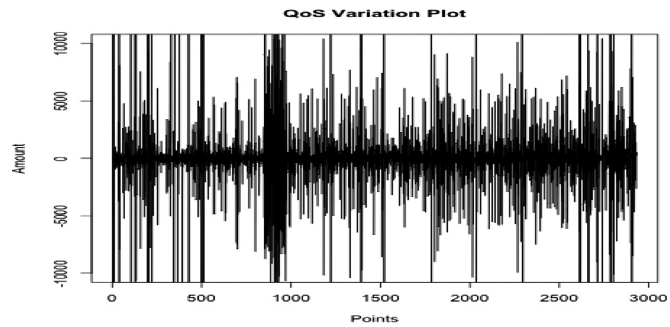

Fig. 2. The first-order difference of the time series.

the various properties studied, however, the most widely considered and adopted property is *Quality of Service* (QoS).

QoS represents the quality instead of the functional features of a service (in this case, a WS). QoS is actually a conceptual abstract term that includes numerous concrete QoS attributes. A more comprehensive introduction and definition of QoS can be found in Section II of Kritikos and Plexousakis (2009), Section III of Lijuan et al. (2012), and Section II of Shi and Chen (2011). Diverse QoS classifications have been proposed and employed to categorize the large number of QoS attributes, including the domain-independent and domain-dependent QoS attributes categories used in Kritikos and Plexousakis (2009). In this paper, however, based on our research objective, we simply categorize QoS attributes into two top-level classes: *static* QoS attributes and *dynamic* QoS attributes. The latter are the focus of this study. Static QoS attributes are defined as those whose actual value does not change; inversely, dynamic QoS attributes are those whose values vary due to a number of factors. Based on our long-term study, we find two different factors that can cause the dynamic QoS values of a WS as perceived by its users (service consumers) to vary: *disparate WS invocation locations* (consumer locations) and *different WS calling times*. A number of remarkable studies and forecasting approaches, such as Zheng and Lyu (2010, 2013) and Zheng et al. (2011, 2013) are already available to explain and predict QoS variations caused by different WS invocation locations. Therefore, this paper focuses on the second factor, namely, QoS changes over time.

The diverse QoS values of a WS are usually published by its provider; however, publisher-provided QoS information usually ignores the fact that some dynamic QoS attribute values vary over time. In publisher-provided QoS information, the values of diverse QoS attributes–whether static or dynamic–are always fixed and do not change (or are updated only infrequently). In other words, in terms of dynamic QoS attributes (e.g., the response time and throughput of WSs), the static QoS values provided by service providers are unreliable and sometimes even useless because they are inaccurate. To demonstrate the QoS variation of WSs over time based on the real-world WS QoS dataset provided by Cavallo et al. (2010), Fig. 1 presents the response time of a commercial WS at different times (a time-series plot of an Amazon-provided WS from 6 July 2006 to 19 November 2006). Moreover, Fig. 2 shows the amount of change (the first-order difference, i.e., $X_i = Y_i - Y_{i-1}$, where Y_i denotes the QoS value at time i) of this response-time time series to help readers fully understand that dynamic QoS value can vary dramatically over time (Fig. 1) and that the scale of change is quite large (Fig. 2). More evidence and examples of such WS QoS fluctuations over time can be found in Section II of Amin et al. (2012a), Section 3.2 of Amin et al. (2012b), Section 4.2 of Godse et al. (2010), Section 2 of Zadeh and Seyyedi (2010), Section 1 of Ye et al. (2014), Section 3.3 of Zheng et al. (2012), Section 8 of Xia et al. (2013), and Section III.B of Yilei et al. (2011). The authors of various studies provide their own reasons why the real values of dynamic WS QoS attributes vary with time in Section 1 of Mu et al. (2009), Section 1 of Xia et al. (2013), Section 1 of Godse et al. (2010), Section 3.3 of Zheng et al. (2012), and Section 2 Senivongse and Wongsawangpanich (2011). For example, in Section 1 of Mu et al. (2009), the authors pointed out that QoS attributes are dynamic due to the characteristics of WSs, namely, their distributed, heterogeneous, and autonomic features. WSs are distributed on the Web, owned and provided by different organizations, and run on different platforms; thus, their quality is characterized by uncertainty. In Section 1 of Xia et al. (2013), the authors mentioned that actual QoS values constantly fluctuate because of the dynamics of their environment, such as Internet congestion and unexpected connection latency. Godse et al. (2010) considered that there are significant QoS variations in execution time and other dynamic QoS attributes because the number of demands for WSs (their work loads) and the resources available for them (e.g., CPU cycles) are dynamic. Finally, Senivongse and Wongsawangpanich (2011) found that the variability in real-world QoS values for WSs is caused by varying numbers of clients, disparate network conditions, and service maintenance operations.

From the above illustrations and discussion, as noted by Godse et al. (2010), it would clearly be unwise to depend on static QoS information provided by service providers to perform diverse WS operations, such as ranking, selection, and composition. Thus, dynamic WS QoS values must be acquired in a manner that provides accurate and reliable QoS data. Based on our investigation, to address the problem of *obtaining accurate dynamic QoS values that change over time*, most existing studies model the problem as an instance of time-series forecasting. Then, they use or propose improvements to various time-series methods to solve the forecasting problem. Time-series forecasting is a well-developed and widely studied research field with a long history that reaches into both statistics and economics; thus, it is reasonable and wise to take advantage of the forecasting methods developed in this field.

The existing primary time-series forecasting methods are thoroughly introduced and described in Section II of Wagner et al. (2007); however, although some studies already use time-series methods to predict future QoS values, the critical and meaningful question of *which time-series method/model is most suitable for forecasting dynamic WS QoS attributes* is still unsolved and cannot be answered without a comprehensive, empirical study comparing the performance of diverse time-series methods in QoS forecasting, which has not yet been performed. The reviewed QoS forecasting studies primarily compare either none or only a few of the available time-series methods to their proposed approach (which is also generally based on a particular time-series method). However, without a comprehensive performance comparison (an empirical study) covering all the common time-series methods,