The Indonesian Journal of Computer Science
www.ijcs.net
Volume 13, Issue 4, August 2024
https://doi.org/10.33022/ijcs.v13i4.4257

A Review to Identify Adequate Data Analytic Frameworks for Managing Cloud Computing Resource Allocation

Sello Prince Sekwatlakwatla[1], Vusumuzi Malele[2]
sek.prince@gmail.com[1], vusi.malele@nwu.ac.za[2]
[1,2]Unit for Data Science and Computing School of Computer Science and Information Systems North-West University Vanderbijlpark, South Africa

Article Information

Received : 5 Jul 2024
Revised : 31 Jul 2024
Accepted : 8 Aug 2024

Keywords

Resource allocation,Data analytic frameworks, Traffic prediction

Abstract

Organizations need to understand the importance of resource allocation and traffic forecasting for cloud computing more than ever because of the increasing demand for online services, data storage and remote work, which makes it challenging to estimate traffic and distribute resources. In cloud computing, there is still opportunity for increasing the model's forecast accuracy. The more accurate the traffic flow, the better the resource allocation. Therefore, this study investigate the adequate data analytic frameworks for managing cloud computing resource allocation in the organisation, According to the findings, six technique were used just for prediction, while one was utilized solely for resource allocation. In this sense, cloud computing resource allocation and prediction may be achieved by combining several techniques.this paper contributes the review to identify adequate data analytic frameworks for managing cloud computing resource allocation.

A. Introduction

Cloud computing is one of the key technologies that allows organizations of all sizes to compete fairly[1]. Most businesses use cloud integration because it satisfies technological requirements, encourages growth and agility and facilitates more seamless communication between IT and other business divisions. However, the majority of service providers are getting complaints from clients about sluggish system performance, application timeouts, and the inability to access specific services.Cloud computing traffic prediction and analysis provide various useful advantages, including quick execution, continuous web presence, improved cooperation, and cheaper subscription costs [2]. These advantages make it the ideal motivator for businesses to migrate their data to the cloud and examine the prospect of forecasting the future in order to manage and supply consistent resources to their consumers.Organizations need to understand the importance of resource allocation and traffic forecasting for cloud computing more than ever because of the increasing demand for online services [3], data storage, and remote work, which makes it challenging to estimate traffic and distribute resources.Numerous businesses might collapse due to the lack of a reliable technique for predicting cloud computing traffic [3-4].

Adoption of ICTs has the potential to significantly impact Africa's development and sustainability, on the one hand. This is because information and communication technology (ICT) may be used to monitor and manage natural resources, promote sustainable behaviors, reduce energy consumption, and mitigate the consequences of climate change[5].Data analytics is the process of analyzing raw data to acquire actionable insights [6], which are then utilized to make informed business decisions. Rather than speculating about the future, data analytics allows you to analyze the past and forecast future patterns and behaviors [7-8]. In a cloud computing environment, the uneven flow of network traffic makes it challenging to forecast the network resources that will be adequate to satisfy the demands of every network client at any given moment [9].

Cloud computing as an infrastructure necessitates appropriate and accurate traffic forecasts in order to satisfy consumer requests and professionally support enterprises [10-11].A study of the literature was done in this respect to look at the data analytic frameworks and tools that have been utilized to predict traffic flow. The review's objective was to find frameworks, models, or algorithms that, are adequate for managing cloud computing resource allocation.

The following research databases were used: the Association for Computing Machinery (ACM), the Institute of Electrical and Electronics Engineering (IEEE), and Scopus.

- The inclusion criteria comprised all articles published between 2020 and 2023.
- The exclusion standards: any articles released before 2018 were not included.

The Systematic Literature Review (SLR) findings are shown in Table 1. It displays the chosen approach, together with its benefits and drawbacks about the paper's topic.

Table 1. The SLR On Predicting Traffic Flow.

Title	Method	Merits	Demerits
Prediction of Bus Passenger Traffic using Gaussian Process Regression	Gaussian Process Regression (GPR) [2]	Even with insufficient data, prediction accuracy may be achieved, and the training and prediction processes are considerably simpler and faster.	Not appropriate for interactive, real-time applications
Supervised Wireless Communication: An Analytic Framework for Real-Time Model Inference in the 5G Core Network	Fifth Generation (5G) wireless communication's real-time prediction framework [1]	forecasting with great accuracy and efficiency during real-time operation	In the future, the planned study would handle real-time resource scaling and usage based on live forecasting.
Short-term traffic flow prediction: An ensemble machine learning approach	An ensemble short-term traffic flow prediction method based on optimized variational mode decomposition (OVMD) and combined long short-term memory network (LSTM)[3]	Enhance the accuracy of forecast outcomes.	The suggested approach does not take into account the influence of multivariate factors.
A network traffic forecasting method based on SA optimized ARIMA-BP neural network.	Simulated Annealing optimized Autoregressive Integrated Moving Average Model Back Propagation Neural Network (SA optimized ARIMA–BPNN) [4]	a number of advancements in network traffic prediction accuracy	There is no clear indication of which field will be responsible for optimizing network infrastructure architecture in the future.
DAGCRN: Graph convolutional recurrent network for traffic forecasting with dynamic adjacency matrix	Traffic factor state network (TFSN) framework [5]	Increase the precision of traffic flow forecasting.	It is difficult to completely investigate and use the correlation mechanism between traffic metrics.
Application of ARIMA-RTS optimal smoothing algorithm in gas well production prediction	ARIMA-RTS (Rauch Tung Striebel) [6]	The model is capable of removing the effects of asynchrony and hysteresis.	Reduce the inaccuracy caused by stimulation while forecasting.
Prediction of Monetary Fund Based on the ARIMA Model	ARIMA Model [8]	Boost the accuracy of traffic flow forecasts	There are still some discrepancies between the expected and tested results.

Predicting daily streamflow with a novel multi-regime switching ARIMA-MS-GARCH model	ARIMA-GARCH model [7]	enhance the model's daily runoff forecasting accuracy ARIMA-GARCH	When traffic grows, the peak values are overestimated.
Traffic Flow Prediction Using SPGAPSO-CKRVM Model	Convolutional neural network (CNN)-based bidirectional spatial–temporal network (CNN-BDSTN) CNN-BDSTN model [9]	The adaptability of incorporating outside elements,	Reduced accuracy
Transferability improvement in short-term traffic prediction using stacked LSTM network	Traffic factor state network (TFSN) framework [10]	Boost the predictability of traffic flow.	It is challenging to completely understand and make use of the traffic parameter correlation process.

This article adds to the current doctorate study with the following goals in mind: To use systematic literature review to identify adequate data analytic frameworks for managing cloud computing resource allocation. The rest of this paper is divided into three sections: section (b) deals with research method, and section (c) deals with Results. Section (e) Conclusion.

B. Research Method

On 05 July 2024, the search term "data analytic frameworks for traffic prediction" was used to extract the information from Scopus, Web of Science, and litmaps. The authors processed and analyzed the data using an online Litmaps tool and a C++ program built using RStudio.

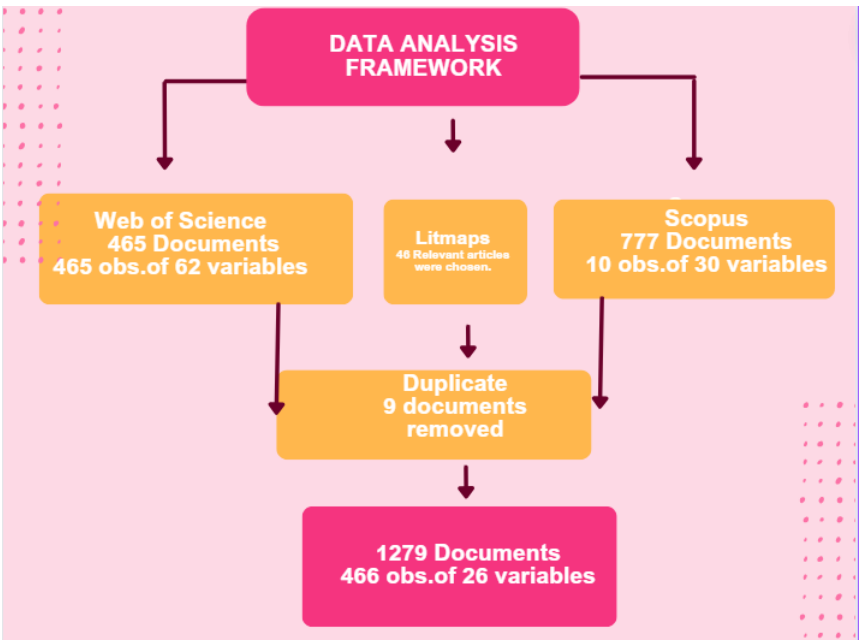

Figure 1. Method analysis

Data is gathered from the Web of Science (WOS), which contains 465 papers and 777 documents from the Scopus database, as shown in Figure 1. The steps to integrate the Scopus database with the Web of Science are as follows: The first phase involves creating bibliographies from the WOS and Scopus databases. Verifying that the tag fields in the WOS file and the Scopus bibtex file match is the second step. The final phase involves merging the two datasets in Rstudio and removing duplicates (there are nine duplicate records in this research). The last stage is to run bibliometric analysis concurrently using Litmaps to identify the most pertinent publications and suitable data analytic frameworks for allocating cloud computing resources.

C. Result and Discussion

1. Bibliometric Analysis

The data record results span the years 2000–2024 and were accessed on July 05, 2024, using the search string "data analytic frameworks for traffic prediction." The article published for the data analytic frameworks has an annual growth rate of 0%, and the international co-authorship is 30.47%, with 226 sources coming from the Scopus and WOS databases. As Figure 2 illustrates below, the average document age was 2.81, with an average of 31, 52 citations per document.

Figure 2. Result Analysis

The main objective of the bibliometric analysis in the paper is to identify adequate data-analytic frameworks for managing cloud computing resource allocation. Which techniques are associated with the data analytic tool? Is there any proposed method for predicting traffic in cloud computing that is considered in the search? As illustrated in figure 3, The Ensemble method and neural network algorithms are indicated as the most dominant prediction tools in cloud computing.

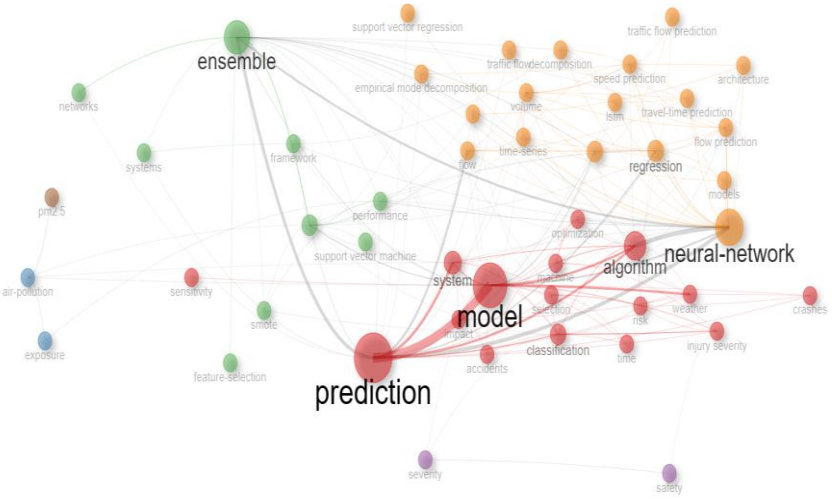

Figure 3. Adequate data-analytic frameworks

2. Litmaps

Litmaps search and review as shown in Figure 4, 2021 dominated the search for data analytic tool techniques for traffic prediction in cloud computing. The search verified the authors of articles published between 2010 and 2022 that are connected to the search for analytic frameworks for managing cloud computing resource allocation.

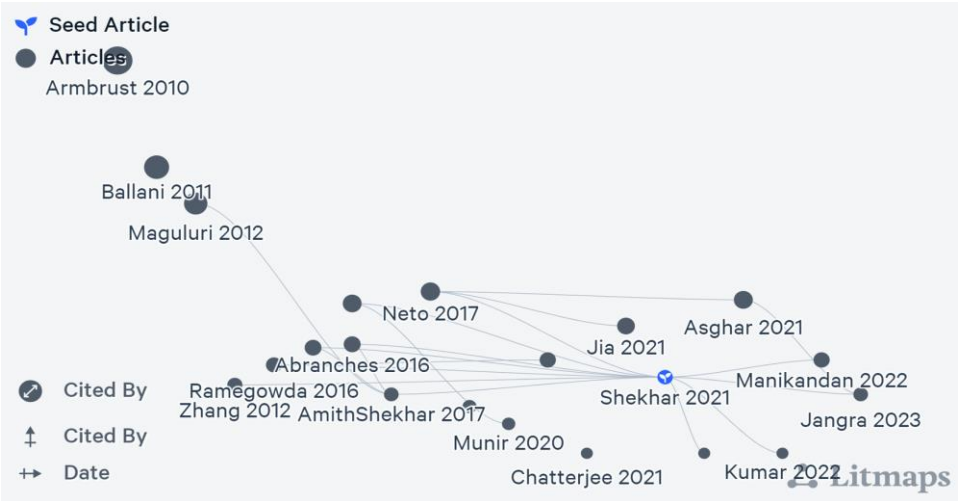

Figure 4. Litmaps Adequate data analytic frameworks.

By increasing the data to further develop research about data analytic frameworks, there has been little recent research to nothing that has shown that researchers are employing data analytic tools to look for traffic prediction, as shown in the figure 5.

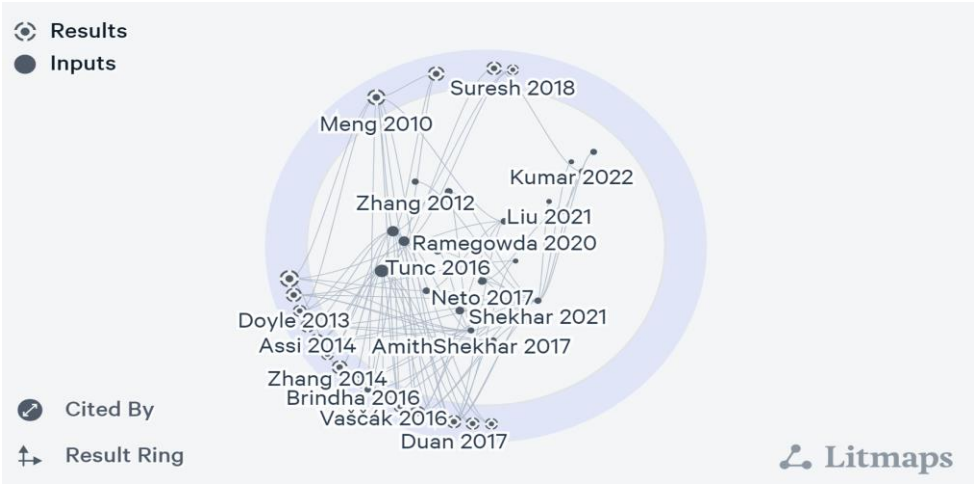

Figure 5. Litmaps Results Ring

An overview of appropriate data analytic frameworks for traffic prediction in cloud computing resource allocation from Litmabs and bibliometric analysis, with models created from simulation results shown in Table 2, the results only focus on the last two years, from 2022 to 2023.

Table 2. Litmabs and bibliometric Data analytic analysis

Method	Year	Litmabs	Bibliometric analysis	Prediction	Resources allocation
Java Virtual Machine [12]	2022	✓	×	✓	×
load balancing framework (GA, SARSA, and Q-learning) [13]	2023	✓	×	✓	✓
Black widow optimization (BWO) algorithm [14]	2022	✓	×	×	✓
LightGBM (Light Gradient Boosting Machine) [15]	2023	×	✓	✓	×
visual information analytics framework [16]	2023	×	✓	✓	✓
Spatio-Temporal Graph Convolution Network (STGCN) [17]	2022	×	✓	✓	×
Long Short-Term Memory (LSTM)-based predictive collision risk (PCR) [18]	2023	×	✓	✓	×
Naive Bayes, Random Forest, Logistic Regression, and Artificial Neural Networks [19]	2023	×	✓	✓	×

As shown in Table 2, two methods predicted traffic and allocated resources, whereas six methods were used for prediction only and one method was used for

resource allocation only. In this regard, the combination of this technique can be used for prediction and resource allocation in cloud computing.

D. Conclusion

The study reviews and identifies adequate data analytic frameworks for managing cloud computing resource allocation by proposing the Systematic Literature Review (SLR) to identify related work in data analytics. The study further investigates related techniques by using Litmaps and Bibliometric simulation. The search term used was "data analytic frameworks for traffic prediction." Data is extracted from Scopus, Web of Science, and Litmaps. According to the findings, six technique were used just for prediction, while one was utilized solely for resource allocation. In this sense, cloud computing resource allocation and prediction may be achieved by combining several techniques.this paper contributes the review to identify adequate data analytic frameworks for managing Cloud computing resource allocation.

E. Acknowledgment

The article is a component of the PhD program in Computer and Information Sciences with Information Technology at North-West University.

F. References

[1] R. Reddy, Y. Munoz, C. Lipps and H. D. Schotten, "Supervised Wireless Communication: An Analytic Framework for Real-Time Model Inference in the 5G Core Network," 2023 International Balkan Conference on Communications and Networking (BalkanCom), İstanbul, Turkiye, 2023, pp. 1-5, doi: 10.1109/BalkanCom58402.2023.10167924.

[2] R. Reddy, Y. Munoz, C. Lipps and H. D. Schotten, "Supervised Wireless Communication: An Analytic Framework for Real-Time Model Inference in the 5G Core Network," 2023 International Balkan Conference on Communications and Networking (BalkanCom), İstanbul, Turkiye, 2023, pp. 1-5, doi: 10.1109/BalkanCom58402.2023.10167924.

[3] G.S.Vidya and V.S.Hari,"Prediction of Bus Passenger Traffic using Gaussian Process Re-gression". Journal of Signal Processing Systems vol. 9, 281–292(2023),[Online].Available: https://doi.org/10.1007/s11265-022-01774-3.

[4] D.Guowen, J.Tang , W.Luo ," Short-term traffic flow prediction " An ensemble machine learning approach,Alexandria Engineering Journal 74(2023),[Online].Available: https://doi.org/10.1016/j.aej.2023.05.015

[5] Z.Shi et al.,"Graph convolutional recurrent network for traffic forecasting with dynamic adjacency matrix". Expert Systems with Applications, Vol 227(2023), [Online].Available: https://doi.org/10.1016/j.eswa.2023.120259.

[6] Zhang et al.," Security computing resource allocation based on deep reinforcement learning in serverless multi-cloud edge computing, journal of Future Generation Computer Systems, Volume 151, February 2024, pp 152-161. [Online].Available: https://doi.org/10.1016/j.future.2023.09.0

[7] Y.Duan et al, "Application of ARIMA-RTS optimal smoothing algorithm in gas well production prediction".Journal of Petroleum, Volume 8, Issue 2, June

2022, Pages 270-277, [Online].Available: https://doi.org/10.1016/j.petlm.2021.09.001

[8] H.Wang , S.Song , G.Zhang , O.Ayantoboc,"Predicting daily streamflow with a novel multi-regime switching ARIMA-MS-GARCH model".Journal of Hydrology: Regional Studies Volume 47, June 2023, 101374.[Online].Available: https://doi.org/10.1016/j.ejrh.2023.101374

[9] W.Fan.: Prediction of Monetary Fund Based on ARIMA Model,Procedia Computer Science Volume 208, 2022, Pages 277-285,[Online].Available: https://doi.org/10.1016/j.procs.2022.10.040

[10] L.Hao,L.Li,H.Wang,Y.Wang,Z.Ma,"Traffic flow prediction using SPGAPSO-CKRVM model",Revue d'Intelligence Artificielle Vol. 34, No. 3, June, 2020, pp. 257-265.June 2020 DOI:10.18280/ria.340303

[11] L.Junyi et al., , "Transferability improvement in short-term traffic prediction using stacked LSTM network.Transportation Research Part C" journal of Emerging Technologies Volume 124, March 2021, 102977. [Online].Available: https://doi.org/10.1016/j.trc.2021.102977

[12] R. Liu et al., "Challenges and Opportunities of a Vastly Distributed Cloud Computing Infrastructure – In the Context of the Dong Shu Xi Suan (DSXS) Project of China," 2023 23rd International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, 2023, pp. 1-4, doi: 10.1109/ICTON59386.2023.10207316.

[13] R.Kumaret al.,Generalized and Simulated Architecture for Seamless Experiment Conditions in Cloud Computing Proceedings of the International Conference on Computer Vision, High Performance Computing, Smart Devices and Networks pp 163–172,doi:10.1007/978-981-19-4044-6_17

[14] A.Jangra and N.Mangla,"An efficient load balancing framework for deploying resource schedulingin cloud based communication in healthcare", Journal of Measurement: Sensors Volume 25, February 2023, 100584,[Online].Available: https://doi.org/10.1016/j.measen.2022.100584

[15] N. Manikandan, P. Divya and S.Janani, "BWFSO: Hybrid Black-widow and Fish swarm optimization Algorithm for resource allocation and task scheduling in cloud computing", journal of Materials Today: Proceedings,62 (2022)[Online].Available: https://doi.org/10.1016/j.matpr.2022.03.535

[16] L.Sobrie,M.Verschelde and B.Roets,"Explainable real-time predictive analytics on employee workload in digital railway control rooms",European Journal of Operational Research 72 (2023), [Online].Available: https://doi.org/10.1016/j.ejor.2023.09.016

[17] R. Raman et al, "An investigation on the role of artificial intelligence in scalable visual data analytics," 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023, pp. 666-670, doi: 10.1109/AISC56616.2023.10085495.

[18] R. Chattopadhyay and C. -K. Tham, "Mixture of Experts based Model Integration for Traffic State Prediction," 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 2022, pp. 1-7, doi: 10.1109/VTC2022-Spring54318.2022.9860682.

[19] B.Noh, H.Park and H.Yeo,Analyzing vehicle–pedestrian interactions: Combining data cube structure and predictive collision risk estimation model

journal of Accident Analysis & Prevention Volume 165, February 2022, 106539,[Online].Available:https://doi.org/10.1016/j.aap.2021.106539

[20] I.C.Obasi and C.Benson."Evaluating the effectiveness of machine learning techniques in forecasting the severity of traffic accidents" journal of Heliyon Volume 9, Issue 8, August 2023, e18812 ,[Online].Available: https://doi.org/10.1016/j.heliyon.2023.e18812

[21] L.Wenqi et al.,"Traffic speed forecasting for urban roads: A deep ensemble neural network model",JOURNAL OF Physica A 593 (2022) 126988. [Online].Available:.https://doi.org/10.1016/j.physa.2022.126988

[22] D.Toratani,T.Yoshihara&A.Senoguchi,"Support algorithm for air traffic controllers' arrival spacing: Improvement of trajectory estimation using Gaussian Process Regression" .journal of Control Engineering Practice,128 (2022).[Online].Available:https://doi.org/10.1016/j.conengprac.2022.105343