Indonesian Journal of Data and Science

Volume 5 Issue 2
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v5i2.136

Research Article

A Literature Review to Investigate Data Analytics Tools for The Allocation of Resources in Cloud Computing

Sello Prince [1,*]; Vusumuzi Malele [2]

[1] Department of Computer Engineering, Gumushane University, Turkey, musa.ju2002@gmail.com
[2] North-West University, Afrika, Vusi.Malele@nwu.ac.za

Correspondence should be addressed to Sello Prince; musa.ju2002@gmail.com

Received 05 June 2024; Accepted 28 July 2024; Published 31 July 2024

Abstract:

To ensure efficient operations and cost-effectiveness, resource management in cloud computing entails managing cloud resources to satisfy application needs, financial restrictions, and security. In this regard, utilizing data analytics tools for the allocation of resources in cloud computing to efficiently predict, track, allocate, and monitor resources enables businesses to make informed decisions based on real-time data, which plays a crucial role in resource allocation. Organizations adopting cloud computing services face increased network traffic, limiting traffic routing flexibility and causing excess traffic to reach unprepared physical nodes due to an inability to adjust to real-time traffic changes. This paper uses a systematic literature review to investigate the data analytics techniques used for resource allocation in cloud computing. It uses data from 2019 to 2024, sourced from different research databases. The results show that the majority of data analytics tools, including ARIMA and SVM, are employed for resource allocation in cloud computing. This study offers guidance to organizations regarding data analytics tools for the allocation of resources in cloud computing, and the recommendations can be utilized for the enhancement of the results in cloud computing, as well as to scholars by suggesting techniques to further investigate resource allocation to address the current gaps in cloud computing.

Keywords: Resource Management, Cloud Computing, Data Analytics Tools.
Dataset Link: -

1. Introduction

The rapid growth of cloud computing services in the world has pushed most organizations seeking better solutions in their organizations to align with the ongoing increase and updated technology in cloud computing [1]–[3]. By enabling agility, scalability, and the provision of resources on demand, cloud computing revolutionizes business operations and enables organizations to quickly adapt to market demands. Most organizations, after adopting cloud computing services, are experiencing more network traffic, which limits the flexibility of traffic routing [4]. Existing software-defined network techniques are unable to adjust to changes in traffic in real time, which results in an excess of traffic reaching unprepared physical nodes [5].

To address this issue, this paper investigates data analytics tools for resource allocation in cloud computing. The ineffective distribution of resources can result in higher expenses and decreased efficiency, making the application of cutting-edge technology and analytical methods essential for efficient resource management [6]–[9]. Cloud analytics uses scalable computing and powerful analytic software to identify patterns in data and extract new insights, delivering results of interest [10].

It is also a process that uses a variety of tools and technology to turn unstructured data into actionable insights that may be used to spot patterns, resolve issues, and promote company expansion [11]. For the majority of cloud computing organizations, data analytics has grown in importance as a tool [12]. By evaluating data on system responses, real-time transactions, and other aspects, businesses may enhance their bottom line and optimize their operations in cloud computing.

İn this regards by continually monitoring workload, monitoring tools may uncover trends in resource utilization, spot bottlenecks, and distribute resources efficiently, assuring performance without sacrificing performance [13], [14].

Organizations may anticipate changes in workload and dynamically distribute resources by utilizing predictive analytics to estimate resource demands [15]. By proactively allocating resources based on past data and trends, these strategies increase overall efficiency. Despite this introduction, this article includes methodology, findings and discussion, then collude with conclusion and future studies.

2. Method:

This study employs the systematic literature review (SLR) approach (as illustrated in Figure 1). First, relevant data are gathered from journal articles and literature studies in order to formulate research questions. This study used a modified version of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist to assess the publications see in **Figure 1**.

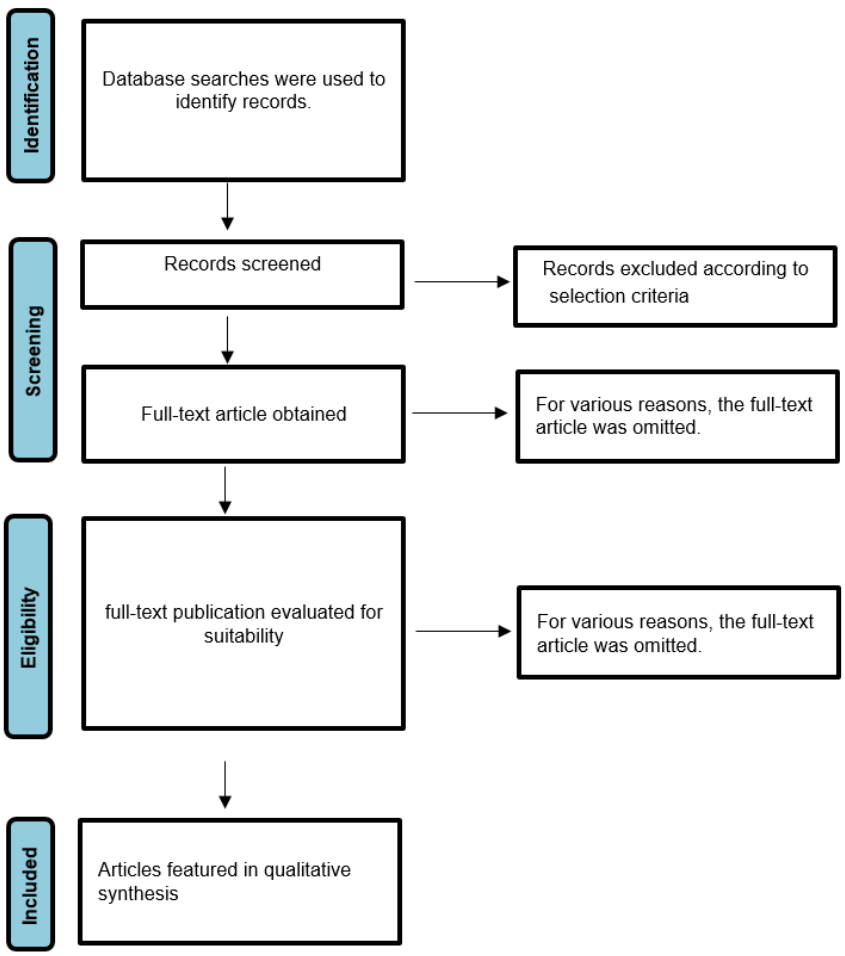

Figure 1. PRISMA Flowchart

a. Inclusion and exclusion criteria

Articles written in English, duplicates, full-text citations, and the date the articles were published are the filter criteria for this study.

b. Search criteria

The following search parameters were used to look for research-related publications on Scopus, Web of service and IEEE that were published after 2020.

- Data analytics tools for the allocation resources

- Allocation of resources in cloud computing using data analytics tools

3. Results and Discussion

Many data analytics tools can be used for the allocation of resources in cloud computing [16]. Ding et al.'s proposed framework supports hybrid analytical models of batch and stream computing, enabling effective collaboration between cloud and edge resources, results show optimizing resource utilities in different conditions, group optimization for batch tasks, and adaptive task scheduling for stream tasks when workloads burst or shrink [17]. Vhatkar and Girish introduce the Whale Random Update-Assisted Lion Algorithm (WR-LA) framework, a novel container allocation technique that reduces network overhead, improves load balancing, and decreases microservice failures by reducing distances between containers and microservices [18], [19]. A fault-tolerant hybrid resource allocation model is proposed in cloud computing for resource allocation. Results show that these methods outperform traditional minimum completion time (MCT) in managing dynamic compute-intensive grid system failures and ensuring task execution interruptions [20].

The ARIMA-RTS (Rauch-Tung Striebel) resource allocation model is presented to effectively minimize the inaccuracy brought on by stimulation while allocating resources in an efficient manner [21]. The allocation of resources from Activity Stream was verified and came in at 85% of what was expected. Implementing a system that uses an application framework to allocate resources, however, has not advanced [21]–[24]. Since cloud computing is an ongoing activity, it is very hard to anticipate with 100% accuracy [25].

By making use of cloud computing's capacity to efficiently handle, comprehend, and use enormous volumes of data, cloud analytics improves an organization's ability to use and analyse this data [26]. A burstable resource-oriented scheduling technique is proposed for heterogeneous clouds that is based on limited space and delta solution-based enhancements. The method outperforms other algorithms, according to experimental data, and the method improves the allocation of resources by over 40% [27]–[29]. The present problem of resource allocation is not unique among cloud computing architectures. Rather than causing CPU or storage constraints, these frameworks may cause system bottlenecks since they transfer a lot of data via the cloud [30].

The use of cloud computing in the telecom sector has developed over the past several years from a new technology to a reliable networking solution that is being widely used [31]–[34]. Network traffic is expected to rise as a result of the cloud computing industry's rapid expansion, causing network congestion, and resource allocation will be affected. In this regard, several scholars have put forth several strategies and carried out tests to improve resource allocation accuracy above conventional techniques. The findings displayed in **Tables 1**, **2**, are derived from the selected publications for this investigation.

Table 1 shows the distribution of the 38 publications for this SLR by publication year. According to the data presented, research on the use of data analytics in cloud computing and resource allocation strategy climbed by 75% in 2020 compared to 2019, increased by 5% in 2021, and decreased by 15% in 2022. Between 2023 and 2024, a consistent volume of research was carried out, increasing threefold, resulting in the publication of 23 research papers during that time.

Table 1. Distribution of Papers by Year

Year	Paper included	Reference
2019	3	[31], [34], [39]
2020	4	[33], [36], [38], [40]
2021	5	[5], [28], [30], [32], [37]
2022	3	[22], [27], [29]
2023	13	[3], [4], [6], [7], [8], [11], [14], [17], [18], [20], [24], [25], [35]
2024	10	[4], [10], [12], [13], [15], [16], [19], [21], [23], [26]

Table 2. Data analytics tools for the allocation of resources

Year	Techniques				
	Autoregressive Integrated Moving Average	**Support Vector Machine**	**Spatio-Temporal Analysis**	**Edge-Cloud Collaborative Scheme**	**Linear Regression**
2019	[39]				[31], [34]
2020	[38]	[40]			[33], [36]
2021	[5], [37]			[28], [30]	[32]
2022			[22]	[27], [29]	
2023	[3], [4], [6], [7], [8], [11]	[14], [17], [18]	[20]	[24], [25]	[35]
2024	[4], [10]	[12], [13], [15], [16]	[19], [21], [23]	[26]	
Total	12	8	5	7	6

In this research, we analysed and compared the most popular data analytics tools for cloud computing resource allocation. According to this study, ARIMA is the most widely utilized approach, and its difference from other methods is rather considerable. Interestingly, 2023 has the most academic publications covering ARIMA six total than any other year. 2021 and 2024, each with two papers, the year 2023 had the greatest number of publications published in this research, with 13 articles. The second method is SVM with 8 published papers. The SVM saw an increase in published papers between 2023 and 2024 with the 3 and 4 published papers. Edge-cloud collaborative scheme also significantly contributed to the allocation of resources in cloud computing at number 3, with seven published papers. Followed by spatial-temporal analysis and linear regression, with 5 and 6 papers published (see table 2).

4. Conclusion

In conclusion, data analytics tools help organizations that are running their platform in cloud computing with the allocation of resources in cloud computing. aids businesses in utilizing data to make choices, streamline operations, and obtain a competitive edge more effectively than with conventional methods, as data analytics tools can now handle vast volumes of data more quickly and accurately. Cloud computing's use of data analytics tools reached its peak in 2021 and then fell down again in 2022; however, the 2023–2024-time frame produced the greatest number of papers (23). Numerous data analytics tool techniques provide very accurate resource allocation in cloud computing. To increase accuracy, some suggest creating unique machine learning models. While linear regression, spatiotemporal analysis, and the edge-cloud collaboration scheme are also utilized, the majority of data analytics tools, including ARIMA and SVM, are employed for resource allocation in cloud computing. We anticipate that these results will aid in the advancement of scholars' comprehension of the best method for allocating resources in cloud computing.

Acknowledgments:

We sincerely thank the North-West University (NWU), Vanderbijlpark, for their outstanding resources, support, and financing for study. Their help has been essential to propelling our research to new heights. Without their assistance, the achievements and developments in this study would not have been feasible. We are appreciative of NWU for the opportunities they have provided us with as well as their dedication to advancing innovation and knowledge.

References:

[1]. N. Kapil,V.Girish,P. Bhole, "Optimal container resource allocation in cloud architecture: A new hybrid model ",Journal of King Saud University –Computer and Information Sciences,Vol.34, 2022.https://doi.org/10.1016/j.jksuci.2019.10.009

[2]. W.Ding,Z.Zhao,J.Wang,H.Li,"Task Allocation in Hybrid Big Data Analytics for Urban IoT Applications ",ACM/IMS Transactions on Data Science (TDS),Vol.1,2020.https://doi.org/10.1145/3374751

[3]. Y.Lin, C.Wan,S.Li,S.Xie, Y.Gan, Y.Lu ,"Prediction of women and Children's hospital outpatient numbers based on the autoregressive integrated moving average model", journal of Heliyon. Vol.9,2023.https://doi.org/10.1016/j.heliyon.2023.e14845

[4]. S.Y.Ilu, R.Prasad."Improved autoregressive integrated moving average model for COVID-19 prediction by using statistical significance and clustering techniques",journal of Heliyon. Vol.9,2023 [Online].Available https://doi.org/10.1016/j.heliyon.2023.e13483

[5]. J.Yuan,D.Li, "Epidemiological and clinical characteristics of influenza patients in respiratory department under the prediction of autoregressive integrated moving average model",journal of Results in Physics Vol.24,2021 [Online].Available https://doi.org/10.1016/j.rinp.2021.104070

[6]. J.Wang,Z. Kacie Pei,Y.Wang,Z.Qin,"An investigation of income inequality through autoregressive integrated moving average and regression analysis", journal of Healthcare Analytics Vol.5,2024 [Online].Available https://doi.org/10.1016/j.health.2023.100287

[7]. Ahmad Hauwa Amshi, Rajesh Prasad,Time series analysis and forecasting of cholera disease using discrete wavelet transform and seasonal autoregressive integrated moving average model,journal of Scientific African. https://doi.org/10.1016/j.sciaf.2023.e01652

[8]. L.Yao, R.Ma,H.Wang,"Baidu index-based forecast of daily tourist arrivals through rescaled range analysis, support vector regression, and autoregressive integrated moving average,"Alexandria Engineering Journal. Vol.60,2021.https://doi.org/10.1016/j.aej.2020.08.037

[9] A.Durand,F.Roueff, "Hilbert space-valued fractionally integrated autoregressive moving average processes with long memory operators", Journal of Statistical Planning and Inference, Vol.231,2024 [Online].Available .https://doi.org/10.1016/j.jspi.2024.106146

[10] J.Zhuang ,Y.Cao,Y.Ding,M.Jia ,K.Feng,"An autoregressive model-based degradation trend prognosis considering health indicators with multiscale attention information ". journal of Engineering Applications of Artificial Intelligence Vol.131,2024.https://doi.org/10.1016/j.engappai.2024.107868

[11]. X. Xu , X.Jin , D.Xiao , C. Ma, S.C. Wong,"A hybrid autoregressive fractionally integrated moving average and nonlinear autoregressive neural network model for short-term traffic flow prediction",Journal of Intelligent Transportation System,Vol.27,2023.https://doi.org/10.1080/15472450.2021.1977639

[12]. C.Lou, X.Xie,"Multi-view universum support vector machines with insensitive pinball loss",journal of Expert Systems With Applications, Vol.228,2024.https://doi.org/10.1016/j.eswa.2024.123480

[13]. M.Zhao, B.Xue, L.Bohan, J.Zhu, W.Song,"Ensemble learning with support vector machines algorithm for surface roughness prediction in longitudinal vibratory ultrasound-assisted grinding",journal of Precision Engineering.Vol.88,2024.https://doi.org/10.1016/j.precisioneng.2024.02.018

[14]. S.Yang,Z.He,J.Chai ,D.Meng ,W.Macek,R.Branco, S.Zhu,"A novel hybrid adaptive framework for support vector machine-based reliability analysis: A comparative study",journal of Structures Vol.58,2023.https://doi.org/10.1016/j.istruc.2023.105665

[15]. H.Moosaei, F.Bazikar, M.Hladík,"Multi-task twin support vector machine with Universum data",journal of Engineering Applications of Artificial Intelligence Vol.132,2024.https://doi.org/10.1016/j.engappai.2024.107951

[16]. W.Dudzik, J.Nalepa, M.Kawulok,"Ensembles of evolutionarily-constructed support vector machine cascades,"journal of Knowledge-Based Systems Vol.288,2024.https://doi.org/10.1016/j.knosys.2024.111490

[17]. W.Yin, H.Xia ,X.Huang, J.Zhang, M.E. Miyombo,"A fault diagnosis method for nuclear power plant rotating machinery based on adaptive deep feature extraction and multiple support vector machines",journal of Progress in Nuclear Energy , Vol.164,2023 [Online].Available https://doi.org/10.1016/j.pnucene.2023.104862

[18]. M.S.Chowdhury,"Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting",journal of Environmental Challenges Vol.14,2024.https://doi.org/10.1016/j.envc.2023.100800

[19]. X.Luo,C.Liu,H.Zhao,"Modeling and spatio-temporal analysis on CO2 emissions in the Guangdong-Hong Kong-Macao greater bay area and surrounding cities based on neural network and autoencoder",journal of Sustainable Cities and Society Vol.103,2024.https://doi.org/10.1016/j.scs.2024.105254

[20]. P.S.Thakur,O.Krejcar,V.Bhatia,S.Prakash."Deep learning based processing framework for spatio-temporal analysis and classification of laser biospeckle data",journal of Optics and Laser Technology. Vol.169,2024. https://doi.org/10.1016/j.optlastec.2023.110138

[21]. T.T.Zeleke,A.Zakaria, W.Lukwasa,K.T.Beketie,D.Y.Ayal,"Analysis of spatio-temporal precipitation and temperature variability and trend over Sudd-Wetland, Republic of South Sudan",journal fo Climate Services Vol.34,2024 [Online].Available. https://doi.org/10.1016/j.cliser.2024.100451

[22]. M.Wu, R.Long,F.Chen, H.Chen,Y.Bai,K.Cheng,H.Huang,"Spatio-temporal difference analysis in climate change topics and sentiment orientation: Based on LDA and BiLSTM model", journal of Resources, Conservation & Recycling. Vol.188,2023. https://doi.org/10.1016/j.resconrec.2022.106697

[23]. P.S.Thakur, O.Krejcar,V.Bhatia,S.Prakash,"Deep learning based processing framework for spatio-temporal analysis and classification of laser biospeckle data",journal of Optics and Laser Technology Vol.169,2024 [Online].Available. https://doi.org/10.1016/j.optlastec.2023.110138

[24]. W. Li, J.Ou,"Machine scheduling with restricted rejection: An Application to task offloading in cloud–edge collaborative computing,European"Journal of Operational Research. Vol.314,2024. https://doi.org/10.1016/j.ejor.2023.11.002

[25]. K.Zhou,F.Gao,Z.Hou,J.Liu, X.Meng,"Cloud-edge collaborated dust deposition degree monitoring for distributed photovoltaic systems ",International Journal of Electrical Power and Energy Systems. Vol.153,2023.https://doi.org/10.1016/j.ijepes.2023.109298

[26]. H.Yang, C.Wang, K.Zhang, S.Dong,"End-edge-cloud collaborative learning-aided prediction for high-speed train operation using LSTM",journal of Transportation Research Part C, Vol.160,2024[Online].Available .https://doi.org/10.1016/j.trc.2024.104527

[27]. W.Zhang, N.Wang, L.Li, T.Wei,"Joint compressing and partitioning of CNNs for fast edge-cloud collaborative intelligence for IoT",Journal of Systems Architecture. Vol.125,2022. [Online]. Available https://doi.org/10.1016/j.sysarc.2022.102461

[28]. Z.Tong, X.Deng, J.Mei, B.Liu, K.Li,"Response time and energy consumption co-offloading with SLRTA algorithm in cloud–edge collaborative computing",journal of Future Generation Computer Systems. Vol.129,2022.. https://doi.org/10.1016/j.future.2021.11.014

[29]. B.Yi, J.Lv, X. Wang, L.Ma, M.Huang,"Digital twin driven and intelligence enabled content delivery in end-edge-cloud collaborative 5G networks", journal of Digital Communications and Networks.[Online]. Available. https://doi.org/10.1016/j.dcan.2022.09.014

[30]. F.Xu,Y.Xie, Y.Sun, Z. Qin, G. Li,Z.Zhang,"Two-stage computing offloading algorithm in cloud-edge collaborative scenarios based on game theory",journal of Computers and Electrical Engineering. Vol.97, 2022. https://doi.org/10.1016/j.compeleceng.2021.107624

[31]. L.Ruana, C.Lia, Y. Zhang, H.Wangc,"Soft computing model based financial aware spatiotemporal social network analysis and visualization for smart cities", journal of Computers, Environment and Urban Systems. Vol.77, 2019.. https://doi.org/10.1016/j.compenvurbsys.2018.07.002

[32]. F.Fouedjio,"Exact Conditioning of Regression Random Forest for Spatial Prediction, journal of Artificial Intelligence in Geosciences. Vol.1, 2020.. https://doi.org/10.1016/j.aiig.2021.01.001

[33]. J.Le, G.L.Salle, J.Badosa, M.David, P.Pinson, P.Lauret,"journal of Added-value of ensemble prediction system on the quality of solar irradiance probabilistic forecasts", journal of Renewable Energy, Vol.162, 2020.. https://doi.org/10.1016/j.renene.2020.07.042

[34]. F. Wang,"The Journey of Cloud Computing with Open Source", UCC '19 Companion: Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing CompanionDecember,2019. https://doi.org/10.1145/3368235.3369378

[35]. T.He, R.Buyya,"A Taxonomy of Live Migration Management in Cloud Computing", ACM Computing Surveys (CSUR), Vol.56, 2023. https://doi.org/10.1145/3615353

[36]. T.Hagemann, K.Katsarou,"A Systematic Review on Anomaly Detection for Cloud Computing Environments, AICCC '20: Proceedings of the 2020 3rd Artificial Intelligence and Cloud Computing Conference. Vol.56, 2020. https://doi.org/10.1145/3442536.3442550

[37]. C.K.Chin,D.Azra, B.Mat,A.Y. Saleh, Skin Cancer Classification using Convolutional Neural Network with Autoregressive Integrated Moving Average,ICRSA '21: Proceedings of the 2021 4th International Conference on Robot Systems and Applications 2021. https://doi.org/10.1145/3467691.3467693

[38]. Y.E.Sutoyo, A.Musnansyah,"A Hybrid of Seasonal Autoregressive Integrated Moving Average (SARIMA) and Decision Tree for Drought Forecasting, ICONETSI '20", Proceedings of the 2020 International Conference on Engineering and Information Technology for Sustainable Industry, 2020. https://doi.org/10.1145/3429789.3429870

[39]. K. Kandananond," Electricity demand forecasting in buildings based on ARIMA and ARX models,IEEA '19: Proceedings of the 8th International Conference on Informatics, Environment, Energy and Applications March 2019. https://doi.org/10.1145/3323716.3323763

[40]. R.Liu, "Stock Selection Strategy Based on Support Vector Machine, MLMI ",'20: Proceedings of the 3rd International Conference on Machine Learning and Machine Intelligence September 2020. https://doi.org/10.1145/3426826.3426829

[41] Yonggang Duan,Huan Wang, Mingqiang Wei,Linjiang Tan,Tao Yue,Application of ARIMA-RTS optimal smoothing algorithm in gas well production prediction, journal of Petroleum, Vol.8, 2022. https://doi.org/10.1016/j.petlm.2021.09.001