Received: 26 June 2018 | Accepted: 5 July 2018
DOI: 10.1002/spe.2628

EDITORIAL

WILEY

Metaheuristics in cloud computing

1 | INTRODUCTION

The cloud computing paradigm is increasingly becoming mainstream and a growing number of companies and research organizations seek to gain value from its unique characteristics, service models, and deployment forms. According to recent estimates of Gartner[1] and IDC,[2] the public cloud services market is projected to grow at a compound annual growth rate of 16.5% and 21.5%, respectively, from 2015-2020. Gartner expects the highest growth rates in spendings on public cloud infrastructure services. This implies not only a high relevance of approaches supporting the management of cloud infrastructures, but also the need for approaches facilitating and improving the use of cloud infrastructures from the consumer's point of view. The realization of the computing utility vision, commoditized and delivered in a manner similar to traditional utilities like water and electricity,[3] however, currently fails because of the huge complexity in cloud environments where many complex decisions have to be made, as recently reported in one of the largest global IT leadership surveys.[4] In addition, due to the environmental impact of cloud computing and the relation between energy consumption and costs, a sustainable use of resources is of considerable importance and requires the development of new policies and techniques. As a result, the success of cloud computing adoption and provision is highly dependent on efficient and intelligent decision making.

This gives rise to many complex combinatorial optimization problems both from the consumers' and providers' perspective.[5] The implications of decisions need to be quantifiable in terms of different dimensions (eg, economics, sustainability, performance) to allow a comparison of alternatives and to decide on the right course of action. Cloud service providers aim to reduce operational expenses and improve economies of scale to maximize profits and provide competitive prices in a highly competitive market. Therefore, cloud service providers are particularly concerned with a cost-effective, energy-efficient, fault-tolerant utilization and management of physical and virtual computing resources, whereby the quality-of-service (QoS) requirements of consumers have to be satisfied. From the consumers' point of view, particularly an efficient selection and utilization of cloud providers and cloud services, as well as the management thereof, is of interest. Several requirements in terms of costs, service quality, and legal constraints need to be taken into account when placing applications into the cloud (for extensive overviews, the interested reader is referred to Heilig et al[6,7]). This also involves specific characteristics of cloud computing, such as the options of scalability, load balancing, and replication. Moreover, cloud-related application areas and research directions have been expanded to novel paradigms in recent years including fog computing, serverless computing, and internet of things (IoT), leading to new open issues and optimization challenges regarding resource management and scheduling, reliability, scalability and elasticity, and sustainability.[8]

The application of metaheuristics for solving combinatorial problems in the area of cloud computing is a growing field of research. Metaheuristics are known as solution methods that employ higher-level strategies to guide local improvement procedures in order to efficiently and effectively explore a solution space (see, eg, Glover and Kochenberger[9]). As such, the application of metaheuristics plays an important role in supporting fast decision making in highly dynamic cloud environments, especially when exact algorithms cannot be used because of a high computational complexity. Given the many applications of heuristics for solving various combinatorial problems in the area of cloud computing, we further see the potential for enhancing the research efforts by incorporating heuristics into metaheuristic frameworks. This further involves the development of hybrid metaheuristics (see, eg, the works of Blum et al[10] and Caserta and Voß[11]) and matheuristics (see, eg, the work of Maniezzo et al[12]), as well as the integration of machine learning approaches.

The special issue *Metaheuristics in Cloud Computing* compiles eight contributions that enhance the state of the art of decision support in cloud computing by applying advanced combinatorial optimization techniques including mathematical programming, heuristics, and metaheuristics. The special issue is divided into works addressing the perspective of both cloud providers and consumers. From the providers' perspective, the first part of this special issue is particularly concerned with the management of cloud resources by means of optimization methods. The first two articles propose models and methods to support energy-aware resource provisioning[13] and virtual machine (VM) consolidation[14] in cloud data centers. Novel power and cooling strategies using metaheuristics[15] and a numerical algorithm for optimal load distribution[16] to optimize the performance and energy consumption in cloud data centers are discussed in the third and fourth article, respectively. The second part of this special issue covers optimization problems from the cloud consumers' point of view. First, methods for aiding the selection of VMs for task scheduling in geographically distributed clouds are presented.[17] The subsequent two articles focus on composite software-as-a-service (SaaS) placement optimization[18] and web service composition.[19] Finally, the last article of this special issue addresses an industrial IoT application by proposing metaheuristic-based methods for optimizing the placement of software-defined networking controllers and sink nodes in cloud environments.[20] In the remainder of this editorial, we summarize the scope and contribution of the articles included in this special issue.

2 | SUMMARY OF THE CONTRIBUTIONS

The first article of this special issue, entitled "An energy-aware resource provisioning scheme for real-time applications in a cloud data center", was presented by Faragardi et al[13] and addressed the resource provisioning for real-time applications in cloud data centers with the goal of minimizing energy consumption through the reduction of used servers. The authors formalized the tackled problem as an integer linear optimization problem. Due to the complexity of the problem, the authors proposed a two-level solution approach composed of a parallel ant colony system leveraged by simulated annealing (PACSA). The selection of such algorithm is based on its good performance in related problems, such as task allocation problems on multicores and placement problems in IoT systems. To assess the performance of this solution approach for the problem at hand, the work provides a well-defined comparison for different problem size scenarios, ranging from 15 to 100 servers, while incorporating various algorithms to evaluate the contribution of their approach. In this regard, the results show that using the proposed metaheuristic approach allows a reduction of around 16% of power consumption, accompanied with a relevant increase of the servers' utilization. This work contributes to the state of the art of metaheuristic applications in cloud computing while also providing real-time decision support in cloud environments.

The article of Yousefipour et al[14] "Energy and cost-aware virtual machine consolidation in cloud computing" aims at optimizing the consolidation process for minimizing the total number of active physical servers with the goal of reducing power consumption and costs in cloud data centers. The authors proposed a mixed-integer nonlinear programming model and a genetic algorithm–based algorithm, referred to as ECVMC. Since the proposed model aims to find a trade-off between power consumption and cost, the objective function minimizes the sum of different energy consumers, namely, the idle power consumption of servers, the power consumption of VMs, and the power consumption for switching the server on or off. Their proposed population-based algorithm shows superior performance compared with that of other placement algorithms from related literature. In this sense, ECVMC is able to reduce the number of active servers by maximizing their utilization and thus leads to a reduction of the total power consumption. In the worst case, ECVMC reduces 24% of the overall power consumption with respect to other algorithmic alternatives, which recommends applying such type of methods and continuing this research direction for addressing important sustainability issues.

The article of Arroba et al,[15] entitled "Heuristics and metaheuristics for dynamic management of computing and cooling energy in cloud data centers", explores a set of energy-aware VM allocation strategies that take into account the impact of thermal conditions in cloud data centers under variable conditions. For this purpose, the authors provided a multi-step procedure and a corresponding optimization framework. The framework incorporates a set of single-objective and multi-objective best-fit decreasing allocation policies and a simulated annealing-based allocation policy. Moreover, a novel cooling strategy for controlling the set point temperature of server rooms, maintaining a safe temperature in the data center, is proposed and evaluated as an essential part of the allocation strategies. To evaluate the different methods for realistic cloud environments, the authors provided extensions for the CloudSim toolkit (see, eg, the work of Calheiros et al[21]), such as to support dynamic voltage and frequency scaling management and thermal models. Based on real data covering scenarios with different levels of workload variability, several experiments are conducted to evaluate the performance of the proposed allocation policies with and without a cooling strategy (ie, fixed set-point temperatures). The authors showed that the simulated annealing algorithm together with the proposed cooling strategy allows to improve the energy efficiency in scenarios with high workload variability and generally indicated that the metaheuristic approach can achieve considerable energy savings up to 21.7% compared with heuristic approaches. This demonstrates the potential of applying metaheuristics for finding improved solutions in the context of cloud resource management.

The fourth article of the special issue is entitled "Optimal load distribution for multiple classes of applications on heterogeneous servers with variable speeds" and presented by Li.[16] The work adapts means of variable and task-type–dependent server speed management that are used to optimize the server performance and to minimize the power consumption of a server with mixed applications. More specifically, the author tackled the problem of finding an optimal load distribution and optimal server speed setting for multiple classes of applications on heterogeneous servers with variable speeds. The objective is to minimize the average task response time while ensuring that a certain power supply is not exceeded. The power-constrained performance optimization problem is studied by means of a M/G/1 queueing system, representing the physical servers, with mixed classes of tasks. The author proposed a numerical algorithm for solving the optimization problem by solving a system of nonlinear equations. The effectiveness of the proposed model and method is demonstrated. As such, the article addresses a critical issue in the context of cloud energy management and further provides an approach to ensure a high QoS while consuming a certain amount of available power in modern cloud data centers.

Ziafat and Babamir[17] addressed the selection of VMs for resource task scheduling in geographically distributed clouds by means of fuzzy c-mean (FCM) and multi-objective linear programming (MOLP). In this problem, providers of distributed clouds have to schedule tasks (clients' requests) in such a way that the clients' objectives (cost, time, reliability, etc.) are optimized. To properly tackle this problem, it is formulated as a multi-objective problem with conflicting objectives, and the authors proposed a combination of a clustering method based on FCM and a MOLP method for providing high-quality solutions. It should be noted that, in this approach, FCM is used for clustering geographically distributed data centers, whereas the MOLP method is used for selecting the VMs matching the users' requests. The performance assessment of the clustering-based MOLP covers scenarios with a fixed number of users and a variable number of requests, and vice versa. Additionally, the authors compared their approach with other well-known multi-objective algorithms such as NSGA-II. According to the reported computational results, the clustering-based MOLP leads to better results within less computational time when compared with other approaches. These results and findings support the implementation of this algorithm for this class of problems.

The article entitled "A multi-swarm for composite SaaS placement optimization based on PSO" by Chainbi and Sassi[18] tackled the dynamic SaaS placement problem occurring in cloud environments. This problem seeks to determine how SaaS components should be placed into the cloud considering operational cloud server requirements and nonoperational ones given by predefined criteria like cost, energy, time, etc. For addressing this problem, a multiswarm particle swarm optimization (PSO) was proposed. The authors evaluated the contribution of their multipopulation-based approach by comparing it with a genetic algorithm and a standard PSO. The reported results indicate that the multiswarm PSO algorithm performs better than the other solution approaches suggesting further investigations of this type of approaches in cloud-related optimization problems.

Ghobaei-Arani et al[19] proposed a moth-flame optimization (MFO) algorithm for web service composition in cloud computing environments. Based on users' needs and requests, the optimization of web service compositions aims at satisfying users' QoS requirements considering resource capacities of geo-distributed cloud data centers. The QoS criteria tackled in this work consider cost, response time, availability, and reliability. The MFO algorithm is tested on a benchmark suite covering scenarios with up to 90 services and 900 atomic services. In addition, the contribution of MFO is put into context by including a comparison with state-of-the-art metaheuristic approaches. In this regard, the MFO approach shows a competitive convergence speed and a relevant improvement in all QoS criteria while outperforming the other algorithms. These results acknowledge the appropriateness of this method and suggest its use in real-world web service composition applications.

The final article of the special issue, entitled "An efficient placement of sinks and SDN controller nodes for optimizing the design cost of industrial IoT systems", is presented by Faragardi et al.[20] For satisfying performance requirements in time-critical IoT systems while reducing deployment costs in underlying software-defined networks, it is important to select a proper number of controllers and sink nodes and locate them in a way that they facilitate an efficient coverage. The article addressed the multiple-sink/controller problem (MSCP), supporting decisions regarding the type and location of controllers and sinks while satisfying reliability and timeliness requirements of a sensor network. The authors presented an integer programming model and proposed a metaheuristic algorithm, PACSA-MSCP, hybridizing a parallel version of the max-min ant system with simulated annealing. In an extensive set of computational experiments, the authors demonstrated the performance of the method against other well-known metaheuristic approaches including simulated annealing, genetic algorithms, and greedy randomized adaptive search procedures. The computational results indicate that the approach outperforms respective methods by reducing the total deployment cost by up to 19%. Moreover, the method's solution gap to the optimal solution, given by the general-purpose solver CPLEX, is less than 2.7%. As a consequence, the work efficiently addresses a critical issue in the context of IoT and fog computing.

Overall, the presented articles in the special issue cover different areas, methodological approaches, and perspectives for advancing the area of cloud computing research with novel approaches in the context of optimization and decision support. We acknowledge the authors for their expertise, dedication, and patience, particularly their effort to produce high-quality articles during multiple rounds of peer reviews. Worthwhile emphasizing is the contribution of the reviewers by making available their expertise in different fields along with invaluable insights and recommendations for improving the articles presented in the special issue. Finally, we hope that the special issue helps to promote future research in applying metaheuristics to complex combinatorial problems in the area of cloud computing research and are looking forward to novel and innovative solution approaches.

ORCID

Leonard Heilig http://orcid.org/0000-0003-0482-611X
Eduardo Lalla-Ruiz https://orcid.org/0000-0002-7286-9501
Stefan Voß https://orcid.org/0000-0003-1296-4221
Rajkumar Buyya https://orcid.org/0000-0001-9754-6496

Leonard Heilig[1]
Eduardo Lalla-Ruiz[1,2]
Stefan Voß[1]
Rajkumar Buyya[3]

[1]*Institute of Information Systems, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany*

[2]*Department of Industrial Engineering and Business Information Systems, University of Twente, 7500 AE Enschede, The Netherlands*

[3]*Cloud Computing and Distributed Systems (CLOUDS) Lab, School of Computing and Information Systems, The University of Melbourne, Victoria 3010, Australia*

Correspondence

Leonard Heilig, Institute of Information Systems, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany.
Email: leonard.heilig@uni-hamburg.de

REFERENCES

1. Gartner. Forecast: Public Cloud Services, Worldwide, 2014-2020, 4Q16 Update. https://www.gartner.com/doc/3562817; 2017.
2. IDC. Worldwide Semiannual Public Cloud Services Spending Guide. http://www.idc.com/getdoc.jsp?containerId=prUS42321417; 2017.
3. Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I. Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. *Future Gener Comput Syst*. 2009;25(6):599-616.
4. KPMG. Journey to the Cloud. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/02/the-creative-cios-agenda-journey-to-cloud.PDF; 2017.
5. Heilig L, Voß S. Decision analytics for cloud computing: a classification and literature review. In: Newman A, Leung J, eds. *Tutorials in Operations Research: Bridging Data and Decisions*. Catonsville, MD: INFORMS; 2014:1-26.
6. Heilig L, Lalla-Ruiz E, Voß S. A cloud brokerage approach for solving the resource management problem in multi-cloud environments. *Comput Ind Eng*. 2016;95:16-26.
7. Heilig L, Buyya R, Voß S. Location-aware brokering for consumers in multi-cloud computing environments. *J Netw Comput Appl*. 2017;95:79-93.
8. Buyya R, Srirama SN, Casale G, et al. A manifesto for future generation cloud computing: research directions for the next decade. arXiv preprint arXiv:1711.09123; 2017.
9. Glover FW, Kochenberger GA, eds. *Handbook of Metaheuristics*. New York, NY: Kluwer; 2003.
10. Blum C, Puchinger J, Raidl GR, Roli A. Hybrid metaheuristics in combinatorial optimization: a survey. *Appl Soft Comput*. 2011;11(6):4135-4151.
11. Caserta M, Voß S. Metaheuristics: intelligent problem solving. In: Maniezzo V, Stützle T, Voß S, eds. *Matheuristics: Hybridizing Metaheuristics and Mathematical Programming*. New York, NY: Springer; 2009:1-38.
12. Maniezzo V, Stützle T, Voß S, eds. *Matheuristics: Hybridizing Metaheuristics and Mathematical Programming*. New York, NY: Springer; 2009.
13. Faragardi HR, Dehnavi S, Nolte T, Kargahi M, Fahringer T. An energy-aware resource provisioning scheme for real-time applications in a cloud data center. *Softw Pract Exper*. 2018;48(10):1734-1757.
14. Yousefipour A, Rahmani AM, Jahanshahi M. Energy and cost-aware virtual machine consolidation in cloud computing. *Softw Pract Exper*. 2018;48(10):1758-1774.

15. Arroba P, Risco-Martin J, Moya J, Ayala J. Heuristics and metaheuristics for dynamic management of computing and cooling energy in cloud data centers. *Softw Pract Exper*. 2018;48(10):1775-1804.
16. Li K. Optimal load distribution for multiple classes of applications on heterogeneous servers with variable speeds. *Softw Pract Exper*. 2018;48(10):1805-1819.
17. Ziafat H, Babamir SM. Optimal selection of VMs for resource task scheduling in geographically distributed clouds using fuzzy c-mean and MOLP. *Softw Pract Exper*. 2018;48(10):1820-1846.
18. Chainbi W, Sassi E. A multi-swarm for composite SaaS placement optimization based on PSO. *Softw Pract Exper*. 2018;48(10):1847-1864.
19. Ghobaei M, Rahmanian A, Souri A, Rahmani AM. A moth-flame optimization algorithm for web service composition in cloud computing: simulation and verification. *Softw Pract Exper*. 2018;48(10):1865-1892.
20. Faragardi HR, Vahabi M, Fotouhi H, Nolte T, Fahringer T. An efficient placement of sinks and SDN controller nodes for optimizing the design cost of industrial IoT systems. *Softw Pract Exper*. 2018;48(10):1893-1919.
21. Calheiros RN, Ranjan R, Beloglazov A, De Rose CAF, Buyya R. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Softw Pract Exper*. 2011;41(1):23-50.