Virtual Machine Placement Methods using Metaheuristic Algorithms in a Cloud Environment – A Comprehensive Review

Deafallah Alsadie
dbsadie@uqu.edu.sa
Department of Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia

Abstract
Cloud Computing offers flexible, on demand, ubiquitous resources for cloud users. Cloud users are provided computing resources in a virtualized environment. In order to meet the growing demands for computing resources, data centres contain a large number of physical machines accommodating multiple virtual machines. However, cloud data centres cannot utilize their computing resources to their total capacity. Several policies have been proposed for improving energy proficiency and computing resource utilization in cloud data centres. Virtual machine placement is an effective method involving efficient mapping of virtual machines to physical machines. However, the availability of many physical machines accommodating multiple virtual machines in a data centre has made the virtual machine placement problem a non deterministic polynomial time hard (NP hard) problem. Metaheuristic algorithms have been widely used to solve the NP hard problems of multiple and conflicting objectives, such as the virtual machine placement problem. In this context, we presented essential concepts regarding virtual machine placement and objective functions for optimizing different parameters. This paper provides a taxonomy of metaheuristic algorithms for the virtual machine placement method. It is followed by a review of prominent research of virtual machine placement methods using meta heuristic algorithms and comparing them. Finally, this paper provides a conclusion and future research directions in virtual machine placement of cloud computing.
Keywords:
Cloud computing, Virtualization, VM Placement, Metaheuristic Algorithms.

1. Introduction

Cloud computing is an emerging field to access a shared pool of configurable resources to the Internet. It prepares the services with an extended customer count by providing many on-demand configurable resources over the Internet [1]. Several cloud computing giants such as Microsoft Azure, Amazon Web Services, Google Drive and Dropbox provide Cloud Computing services. There are mainly three types of cloud services, infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS). IaaS is the topmost layer for providing on-demand infrastructure resources of cloud computing [2, 3]. It offers infrastructure resources such as hardware, processor, storage and power requirements to the virtual machines running over a physical machine [4]. The cloud users using IaaS services maintain control over applications, data, storage, operating system and security. The most common examples include AT & T, GoGrid, Verizon, Amazon web services. Using PaaS services, cloud users can develop, test, and deploy their applications [4]. They have control over the deployment and execution of their applications. PaaS facilitate cloud users by offering a working framework and programming environment. SaaS service provides access to different applications to the cloud users through the Internet without installation or purchase.

Different services such as IaaS, PaaS and SaaS are provided to the cloud users using different deployment models such as private cloud, public cloud, community cloud and hybrid cloud [5]. These deployment models are used as virtual distributed systems. Virtualization plays a significant role in the cloud computing environment that divides hardware resources of physical machines into several executable and isolated environments in the form of virtual machines [1]. The virtual machines mounted on a single physical machine are isolated and can execute applications for different users in a different environment using shared computing resources of the same physical machine. The isolated environment of a virtual machine is maintained by a software called virtual machine monitor or hypervisor. Several real-time realizations of hypervisors have been developed, such as XEN and VMware. In addition, efficient management of computing resources in cloud data centres is also required for maintaining the quality of services per cloud user, optimal utilization of physical machines, and maximum return on investment [1, 6].

Manuscript received April 5, 2022
Manuscript revised April 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.4.19

In the recent past, the tremendous use of cloud computing services has increased the abrupt increase in energy utilization of data centres [2, 7]. A promising direction for solving the energy utilization problem is to minimize the number of active physical machines and shut down the active Cloud servers. Handling unutilized cloud servers is a critical challenge for addressing the energy efficiency problem of cloud data centres [8]. An efficient virtual machine placement in a minimum number of physical machines while ensuring the quality of service to cloud users is an effective way to handle energy efficiency issues. Virtual machine placement helps to improve the utilization of computing resources and enhance return on investment [9].

Several methods have been developed for optimizing energy efficiency and computing resource utilization. The most effective method includes server consolidation and virtual machine placement [10, 11]. Virtual machine placement is the main focus of the research community due to its potential benefits for managing Cloud Computing resources, energy efficiency and return on investment effectively. Several virtual machine placement methods have been developed by considering different constraints and objectives.

This paper provides a comprehensive review and taxonomy of metaheuristic algorithm-based virtual machine placement methods. This paper analyses different aspects of virtual machine placement methods and categorize different virtual machine placement methods according to a taxonomy of metaheuristic methods. The well-known research work in virtual machine placement is analyzed and compared in different aspects to identify different optimization criteria used in virtual machine placement methods. This paper contributes in the following ways.

- It provides background and essential knowledge of virtual machine placement in cloud computing.
- It provides a taxonomy of meta heuristic methods.
- It provides a comparative analysis of metaheuristic algorithm based virtual machine placement method as per proposed taxonomy
- Finally, it provides significant research directions in virtual machine placement of cloud computing.

This paper is structured as follows. Section 2 presents recent work of the existing review studies. Section 3 introduces the concepts of virtual machine placement and taxonomy of objective functions for effective virtual machine placement in cloud computing. Section 4 presents different virtual machine placement methods based on meta heuristic algorithms and compare them in different dimensions. Section 6 discusses the metaheuristic algorithms. Finally, the paper concludes and provides future research directions in Section 5.

2. Related Work

Many researchers presented comprehensive reviews of different virtual machine placement mechanisms in the recent past. Their analysis focused on different perspectives of the virtual machine placement problem. For example, Beloglazov et al. [12] provide a comprehensive review of virtual machine placement methods focusing on energy consumption. The authors mainly focused on intelligent management for minimizing the energy consumption of cloud computing resources. Accordingly, they reviewed energy-efficient design classification for cloud computing resources at different operating system levels, hardware, data centre and virtualization. They highlighted the number of requirements for an effective virtual machine placement solution, including virtual machine migration due to changes in the workload, virtualization, ensuring service level agreement, meeting quality of service of different applications based upon hardware computing resources.

Khan et al. [13] presented a comprehensive review of dynamic virtual machine consolidation. They focused on multiple classifications. Firstly they divided virtual machine consolidation mechanisms into different categories, Static and dynamic, centralized and distributed architecture. They are also considered fresher level as a classification criterion of virtual machine consolidated mechanisms. They proposed that threshold based methods are limited to search space and find optimal solutions for a short span. They have also classified virtual machine consolidation based upon the number of virtual machine migrations. They divided virtual machine consolidation methods based upon a single button machine selection or a group of virtual machines. They have also provided virtual machine selection methods in their review.

Ahmad et al. [14] presented a survey on virtual machine migration methods and server consolidation methods. They proposed a survey of live migration methods and a consolidation environment taxonomy based on architecture, computing resource assignment method, migration triggering method, and their models. They also proposed to divide virtual machine migration methods based on dynamic voltage, bandwidth optimization, and frequency scaling enabled power optimization.

Varasteh and Goudarzi [15] provided a survey of server consolidation methods and classified them in different aspects such as migration point, static, dynamic or predictive. They classified optimization methods into the exact method, heuristic method for metaheuristic method based upon different types of objectives, such as minimizing physical machines, minimizing service level agreement violations, performance overhead, optimizing computing resource utilization and performance evaluation framework.

On similar lines, Usmani and Singh [16] reviewed the virtual machine placement method based on energy efficiency and service parameters. They mainly focused on dynamic energy efficient algorithms and divided them into four categories: constraint programming, stochastic integer programming, bin packing, and genetic algorithm.

Masdari et al. [17] classified virtual machine placement methods based upon parameters and objectives considering energy, computing resources, cost, service level agreement, application, security, and evaluation framework.

Alboaneen et al. [18] focused on nature inspired metaheuristic to methods for virtual machine placement in a cloud computing environment.

Xu et al. [19] conducted a systematic review of virtual machine placement methods by focusing on load balancing methods. They analyzed and classified the existing methods as centralized and distributed virtual machine load balancing methods. They explored the virtual machine load balancing methods in different dimensions, including the dynamicity of virtual machine allocation, virtual machine uniformity, optimization strategy, and virtual machine resource type. However, they have not considered metaheuristic methods in their review. They only focused on load-balancing based solutions for virtual machine placement in the cloud computing environment.

Al-Dulaimy et al. [20] also reviewed power efficiency methods for cloud computing data centres. They presented a taxonomy for power management methods, followed by critical challenges and future directions about power management methods in the cloud computing area. They summarized power efficiency methods in different aspects such as single or multiple data centres, online or offline scheduling, homogeneous or heterogeneous environment and computing resources utilization for improving energy efficiency in a cloud computing environment. They mainly focused on power aware scheduling methods used for energy consumption minimization. But they ignore other aspects such as virtual machine placement, virtual machine migration and server consolidation.

This paper mainly focuses on presenting a comprehensive review of current virtual machine placement methods proposed using metaheuristic algorithms. In this work, we considered multi-objective metaheuristic based methods and analyzed them in different dimensions such as objectives, computing resources, and environmental evaluation. Table 1 presents summary of various review studies focusing virtual machine placement. Table 1 presents summary of various review studies focusing virtual machine placement.

3. Virtual Machine Placement

In the cloud computing environment, the virtualization process plays a significant role that partitions the physical computing resources such as memory, storage, and CPU into multiple isolated execution environments called virtual machines [2]. These virtual machines are located in data centres that are geographically distributed. The virtual machines with computing resources such as memory, CPU, storage, and network bandwidth are available and managed effectively by the users according to their requirements in the cloud computing environment. Virtual machine selection, involving selecting an appropriate physical machine for running virtual machines, is very significant [21]. Virtual machine placement involves allocating an appropriate virtual machine at each physical machine in cloud data centres. It is also called the placement plan of the virtual machine to physical machine mapping. Virtual machine placement mainly aims to save power and deliver the quality of services to cloud users [16].

Virtual machine placement is a complex task due to unpredictable arrival patterns of virtual machine instance requests and large cloud data centres, leading to NP hard problem of finding optimal or near optimal physical machines for a given virtual machine [17, 22]. In a cloud computing environment, there are several physical machines holding a number of virtual machines. Virtual machines are generally migrated from one physical machine to another when some physical machine is overburdened or under burdened. It has been observed that cloud data centres often are not using their physical machines to their maximum capacity, leading to energy inefficiency and hence increasing power consumption cost of the data centre. It has been observed that an idle cloud server consumes up to 70% of energy than of its total capacity [23]. In order to reduce the number of active physical machines, virtual machines must be consolidated to reduce power consumption off the data centre. Underutilized cloud data cloud computing resource utilization and reduce the power consumption cost, leading to improve return on investment of the cloud data centre. To solve this problem, virtual machines must be migrated from an overloaded cloud server to under loaded cloud server. It involves identifying the over loaded and under loaded physical machines, followed by virtual machine migration. The selected virtual machines must be placed to an appropriate physical machine so that the overall objective of migrating virtual machines gets achieved without violating the service level agreement of cloud users.

The virtual machine placement method can be divided into different categories depending upon criteria for solving virtual machine placement problems based upon optimization objectives [24]. The considered period can change over time

Table 1. Summary of virtual machine placement method's reviews

Study	Focus	Classification criteria
[12]	Energy efficiency design	Computing resources at different layers
[13]	Virtual machine consolidation approaches	Centralized vs distributed methods
[14]	Virtual machine migration	Online migration, consolidation environment
[15]	Virtual machine consolidation	Migration point, migration decision strategy, optimization method
[16]	Virtual machine placement	Energy efficiency, QoS based
[17]	Virtual machine placement	Optimization methods
[18]	Virtual machine placement	Optimization methods, objectives

leading to several possible solutions with different objective functions optimized in different scenarios.

Table 2 shows the most important criteria for solving virtual machine placement problems described in the following subsections.

Table 2. Objective functions for virtual machine placement methods

Virtual machine placement methods	Resource utilization maximization
	Performance maximization
	Network traffic minimization
	Quality of Service Maximization
	Energy consumption minimization
	Economic cost optimization

3.1 Resource utilization maximization

Cloud data centres store multiple computing resources such as memory, storage, processing units, GPU and network bandwidth. In the cloud computing environment, it is challenging to make use of all computing resources efficiently. Considered different objective functions are resource utilization maximization as listed below.

1) Resource utilization maximization
2) Resource wastage minimization
3) Maximum average utilization minimization
4) Elasticity maximization

A. Performance maximizaion

Performance maximization has also been considered a significant criterion for defining virtual machine placement policies. The significant objective functions for performance maximization are listed below.

1) Total job completion time minimization
2) Shared last level cache (SLLC) contention minimization
3) Security metrics maximization
4) Resource interference minimization
5) QoS maximization
6) Performance maximization
7) Deployment plan time minimization
8) CPU demand satisfaction maximization
9) Availability maximization

B. Network traffic minimization

Many researchers also focused on minimizing network traffic for solving virtual machine problems at different nodes. The most common objective functions that have been considered for optimization in solving virtual machine problems are listed below.

1) Worst case cut load ratio minimization
2) WAN communication minimization
3) Overall communication cost minimization
4) Node cost minimization
5) Network cost minimization
6) Network performance maximization
7) Network traffic minimization
8) Network utilization minimization
9) Migration number minimization
10) Migration overhead minimization
11) Migration time minimization
12) Link congestion minimization
13) End-to-end delay minimization
14) Data access minimization
15) Data transfer time minimization
16) Cloud QoE maximization (response time minimization)
17) Cloud service response time minimization
18) Average traffic latency minimization

C. Quality of Service Maximization

Many researchers solved the virtual machine placement problem by considering different constraints [2]. One significant constraint is ensuring your service quality to the cloud customers. The most common objective functions related to the quality of service maximization includes the following [26].

1) Reduction of the resource interference
2) High accessibility, Performance Improvement
3) Resource interference minimization
4) Reliability

D. Energy consumption minimization

Most of the research in the field focused on minimizing the energy consumption for developing virtual machine placement methods at different levels. The most common objective functions related to energy minimization are described below.

1) WDM layer Power Consumption
2) Power Consumption Minimization
3) Number of PMs Minimization
4) Network Power Consumption Minimization
5) IP Layer Power Consumption Minimization
6) Energy Consumption Minimization
7) energy efficiency Maximization
8) Datacenter Power Consumption Minimization

E. Economic cost optimization

Few researchers have considered the economic cost of defining virtual machine placement policies. They mainly focused on minimizing the different costs and increasing the return on investment for the data centre. The most significant objective functions that the researchers optimize are listed below.

1) Total infrastructure cost minimization
2) Thermal dissipation costs minimization
3) SLA violations minimization
1) Server cost minimization
2) Reservation cost minimization
3) Operational cost minimization
4) Electricity cost minimization
5) Economic revenue maximization

4. Metaheuristic algorithms for effective virtual machine placement

An efficient virtual machine placement method is one significant method for addressing the issue of energy efficient and resource usage in cloud data centres [27]. A virtual machine placement method attempts to plan virtual machines to appropriate physical machines to optimise the objectives of the cloud data centre such as optimizing power consumption, maintaining throughput while keeping service level agreement and appropriate quality of service [28]. However, considering multiple virtual and physical machines in the cloud computing data centre, the virtual machine placement problem can not be solved in polynomial time. Thus, it is considered as a non deterministic polynomial hard (NP hard) problem in the cloud computing [29].

Many approaches have been proposed for solving NP-hard problems like virtual machine placement problems, such as heuristic and metaheuristic methods. Heuristic methods are designed for solving particular problems. In contrast, meta heuristic methods find near optimal solutions. The meta heuristic methods are generally applied in three forms for solving virtual machine placement problems as depicted in Table 3, single objective methods, multi-objective as single objective and multi objective as multi objective methods [30, 31].

Table 3: Meta heuristic optimization method classification

Meta-heuristic methods	Multi objective as multi objective
	Multi objective as single objective
	Single Objective

Multi objective as multi objective methods consider multiple objectives as different objectives and optimize them at a time. These methods generate many non inferior solutions using Pareto front concept and provides trade off to the cloud suppliers.

Single objective methods mainly used for optimizing a single objective from a set of objectives at a time. Multi objective as single objective methods are used for optimizing multiple objective functions by fusing them into a single function. These methods incorporate domain knowledge to allow correct fusion of the objective functions [32]. Single objective and multi objective as single objective have limitations to solve multiple and conflicting objective problems.

Meta heuristic-based virtual machine placement methods can also be divided into three categories: single solution based, population based, and hybrid methods, as presented in Table 4.

Table 4: Meta heuristic algorithms based virtual placement methods

Meta heuristic algorithms based virtual placement methods	Single solution based methods
	Population based methods
	Hybrid methods

A. Single solution based methods

The single solution based virtual machine placement methods involves initiation with a single solution, and that solution is manipulated and processed during the optimization process. Search methods attempt to find local reasons for efficient solutions. The most common examples include simulated annealing and local search based solutions.

Many virtual machine placement solutions have been proposed using single solution based method. For example, Li et al. [33] proposed a model for partitioning multidimensional space by proposing an algorithm called "EAGLE" for reducing power consumption and balance the multidimensional resource utilization. The proposed algorithm evaluates the resources for each physical machine to accommodate the new virtual machine. Accordingly, it selects the appropriate physical machine to place the new virtual machine and avoids unnecessary resource fragments. The multidimensional partitioning of space allows finding computing resource leakage quantitative. Therefore, their algorithm helps reduce computing resource wastage and reduces the number of active physical machines, reducing the power consumption of cloud data centres. The authors evaluated their algorithm against the greedy first fit algorithm. The comparative results demonstrate a 10% reduction in power consumption using the EAGLE algorithm compared to the first fit algorithm.

Jamali et al. [34] introduced an evolutionary framework for solving multi-objective virtual machine placement problems. They focused on optimizing research utilization and power consumption simultaneously. They proposed their framework based upon the concept of competition among imperialist countries that attempt to control colonies of other countries. The proposed algorithm has been validated for good converge and a balance between exploitation and exploration capability. The authors combine the two objectives into a single objective using different weights.

Xu and Fortes [35] suggested an approach for the dynamic placement of virtual machines by considering a cross-layer control system. They proposed their approach based upon thermal emergency, power consumption and resource contention. Their approach inverse a global controller for receiving data from virtualization layers and platforms. Accordingly, the global controller initiates virtual machine placement. The global controller works in 3 phases. Firstly, it detects threshold based condition for triggering virtual machine migration using sliding window protocol. Secondly, it selects virtual machines based upon thermal emergency, power consumption and resource contention. Finally, it selects a physical machine based upon power, performance, and temperature after accepting virtual machine migration. Experimental results demonstrated an 80% reduction in virtual machine migrations and enhanced up to 30% performance of cloud data centres.

Addya et al. [36] suggested an approach called MVMP to maximize profit and reduce power cost based upon a simulated annealing algorithm. Their approach optimized to active using a real multi objective approach based upon the Pareto concept.

Geronimo et al. [37] suggested a virtual machine organization framework to relocate virtual machines based upon 4 elements, rules, priorities, qualifiers and cost.

Similarly, Baalamurugan and Bhanu [38] introduced virtual machine placement method based upon multi-objective krill herd algorithm for finding non-dominated solutions. They optimized power consumption and resource wastage.

Mollamotalebi and Hajireza [39] introduced dynamic system virtual machine management for optimizing energy consumption and service level agreement violation. They evaluated their approach using DCSim simulation environment.

Song et al. [40] introduced multi-objective virtual machine selection method for improving the impact of CPU temperature, power consumption and computing resource utilization using weighted approach. They applied horror stick approach to get an optimized solution based upon weighted objective. Single solution based virtual machine placement methods can be summarised in Table 5.

B. Population based methods

Population-based virtual placement methods involves finding near optimal solutions using evolution concept in different generations. These methods explore different solutions using diversity of search space and apply the dominance concept for selecting and identifying Pareto optimal solutions in each generation. Several algorithms have been proposed in this category. Significant methods in this category are described in the following sections.

1) Genetic-based approaches

Xu et al. [41] Presented a two level control system to manage workload- virtual machine and virtual machine physical machine mapping for reducing computing resource wastage and power consumption of cloud data centre. They applied an improved genetic algorithm based fuzzy multi objective evaluation for achieving the objective of resource wastage minimization and power consumption minimization of data centre.

Similarly, Sinong Wang [42] also proposed to use improve genetic algorithm with local heuristic and elitism strategy for solving virtual machine placement problem. They focused on maximizing computing resource usage, minimizing communication traffic, and balancing resource utilization. They optimized the objective by combining resource utilization and balancing resource utilization by applying constraint of minimizing communication traffic. Their experimental results demonstrated that their approach can achieve maximum resource utilization and minimum communication traffic.

Table 5. Summary of single-solution based virtual machine placement methods

Study	optimization objectives	Optimization method
Li et al. [33]	• Balancing resource utilization • Number of active PMs	EAGLE algorithm
Jamali et al. [34]	• Power consumption • resource wastage	Imperialist competitive Algorithm
Xu and Fortes [35]	• Energy consumption • Number of migrations • SLA violations	DVFS combined with fuzzy weights method
Addya et al. [36]	• Maximizing profit • Reducing energy consumption	Multiobjectivesimulated annealing
Geronimo et al. [37]	• implementation costs • Increasing placement quality	VM ranking method
Baalamurugan and Bhanu [38]	• Power consumption • Resource wastage	Krill herd algorithm
Mollamotalebi and Hajireza [39]	• Energy consumption • SLA violations	CPU, RAM
Song et al. [40]	• CPU temperature • Power consumption • Resource usage	Multi-objective VM selection algorithm

Xiaoli Wang [43] presented an energy aware and locality aware task scheduling method based on mapreduce framework. They formulated the problem of task scheduling as integer bi level programming and solve it using multi-objective genetic algorithm.

Liu [44] applied NS-GGA method to solve virtual machine placement problem for reducing number of active physical machines, balancing computing resource utilization and communication traffic. They applied NSGA-II algorithm for optimizing these objectives.

Sofia and Kumar [45] suggested approach for minimizing energy consumption and maintaining cloud user satisfaction.

Riahi and Krichen [46] proposed a Framework for virtual machine placement to reduce power consumption and hence operational cost by optimizing computing resource wastage based upon genetic algorithm and Bernoulli distribution.

2) ACO-based approaches

Gao et al. [47] suggested a virtual machine based upon ant colony optimization with the name of reducing computing resource wastage and power consumption. They considered memory and processor as computing resource but ignored disk size. Experimental results demonstrated that the performance of their approach is better than multi objective genetic algorithm and two single objective algorithm called FFD heuristic and Min-Max Ant System.

Similarly Malekloo et al. [48, 49] also applied AC metaheuristic approach with properly stick decision rule and heuristic information formula. They use Pareto front approach for identifying near optimal solutions. In their approach they focused on minimizing power consumption and minimizing computing resource wastage. They implemented their approach using cloud sim simulation environment. They compared their experimental results with multi objective GA and three single objective algorithms, namely FFD, DVFS, and linear regression methods. The comparative results indicate that their approach can lead to minimize the power consumption and computing resource wastage in comparison to other algorithm.

Tan et al. [50] proposed an energy aware multi-objective ACO based approach for virtual machine placement. The main objective of their approach is to optimize power wastage. Computing resource uses and service level agreement violation. They used CPU utilization for computing service level violation. For optimizing the multiple objectives, they used weighted Coefficient method, based upon their expertise. They demonstrated that their approach pleads to minimize energy consumption in comparison to BFD method.

Zhang et al. [51] suggested a multi-objective combinatorial optimization method using ACO method to reduce energy consumption and ideal computing resources in cloud data centre.

3) BBO-based approaches

Zheng et al. [54, 55] has applied BBO and complex system optimization method for solving virtual machine placement problem. They formulated virtual machine placement problem in two complex system having multiple objectives and constraints. Multiple objectives and constraints of the complex system are salt using BBO method. In their approach, the authors attempted to optimize power conversion, storage traffic, computing resource wastage, inter virtual machine traffic and intra virtual machine traffic, migration time, extensibility and adaptability. They compared performance of their approach with MGGA, VMPACS, and a Pareto-based best-fit algorithm. The comparative results demonstrates period performance of their proposed approach.

Li et al. [56] also applied BPO and complex system to optimize load balance of cloud data centre.

4) PSO-based approaches

Ramezani et al. [57] proposed a fruit for solving virtual machine placement problem using fuzzy logic and PSO algorithm to optimize power consumption, virtual machine transfer time and computing resource use. They applied fuzzy logic to attractively control weights of PSO algorithm. They validated better results than conventional Mamdani fuzzy inference technique and conventional PSO algorithm.

Luo et al. [58] suggested a multi-objective PSO algorithm to optimize link loss ratio by considering service reliability and quality of tenant. Their approach can reduce energy consumption and resource wastage. However, in their approach they have not clarified for static or dynamic consolidation and evaluation environment used in the experiment.

Guo et al. [60] presented a multi objective data placement approach in a cloud computing environment and attempted to optimize cost and time using the PSO algorithm and processor interaction graph. They proposed mapping of all data of task to a data centre for solving data placement problem.

5) ABC-based approaches

Li et al. [61] proposed a energy aware and multi resource overload scheme called "EC-VMC". They attempted to improve resource uses of physical machines for reducing energy consumption of data centre. In their approach they used different optimization methods for different phases of dynamic consolidation. The experimental results validate superiority of their proposed method regarding energy consumption and quality of service.

6) Memetic approaches

Pires and Barán [62] propose a memetic multi objective approach for solving virtual machine placement problem by considering service level agreement. They focused on reducing power consumption, increasing return on investment and reducing network traffic. Define the fitness function in their approach using non domination rank, Euclidean distance date comparison.

Study	optimization objectives	optimization method	Study	optimization objectives	optimization method	Study	optimization objectives	optimizatio n method
Xu and Fortes [41]	• Avoiding hotspots • Power consumption • Reducing resource wastage	Control system based on GGA	Tan et al. [50]	• Resource utilization • Energy wastage SLA violation	ACO algorithm and weighted coefficient method	Xu et al. [59]	• Migration times • Resource usage	Improved PSO algorithm
Wang et al. [42]	• Resource utilization • Resource usage balancing • Communication traffic	GA and local heuristic with elitism strategy	Zhang et al. [51]	• Idle resources • Energy consumption	ACO algorithm and vector packing	Guo et al. [60]	• Communication time • Communication cost	PSO algorithm
Wang et al. [43]	• Data locality • Energy consumption • Performance • Resource utilization	MapReduce framework and GA	Pham and Le [52]	• Energy consumption Load balancing	ACO algorithm	Li et al. [61]	• Power consumption • VM migrations • Overload probability	ABC algorithm
Liu et al. [44]	• Communication traffic • Number of hosts • Resource balancing	NSGA-II	Ashraf and Porres [53]	• Reducing the number of migrations • Resource utilization	ACO algorithm	Lopez Pires and Barán [62]	• Energy consumption • Network traffic • Economic revenue	Memetic algorithm

Sofia and Kumar [45]	• Minimizing makespan • Energy consumption	DVFS method NSGA-II	Malekloo and Kara [48]	• Inter VM traffic • Power consumption • Resource wastage • Server unevenness • Migration Storage traffic	BBO algorithm and complex systems	LópezPires and Barán [63]	• Power consumption • Network traffic • Economic revenue • QoS • Network load balancing	Memetic algorithm
Riahi and Krichen [46]	• Power consumption • Resource wastage	GA and Bernoulli distribution	Zheng et al. [54]	• Inter host loads Migration time	BBO algorithm and complex systems	Li et al. [64]	• Resource wastage • Energy consumption	Firefly algorithm
Gao et al.[47]	• Power consumption • Resource wastage	ACO algorithm	Ramezani et al. [57]	• Power consumption • Resource utilization VM migration time	Fuzzy logic and PSO algorithm	Ding et al. [65]	• Security • Resource loss • Energy consumption • Performance	Firefly algorithm
Malekloo et al. [48, 49]	• Power consumption • Resource wastage • Communication cost	ACO algorithm and modified decision rule and heuristic information formula	Luo et al. [58]	• Resource utilization • Data center link lose	Improved PSO algorithm			

Table 6. Summary of population based virtual machine placement

7) Firefly algorithm

Li et al. [64] applied the Firefly algorithm to optimize energy consumption and virtual machine placement resource utilization in a cloud computing environment. They used weighted sum method for optimizing multiple objectives.

Ding et al. [65] proposed a virtual machine placement framework to address the problem of security and performance using the Firefly optimization algorithm.The above cited population based virtual machine placement methods can be summarized in Table 6.

C. Hybrid methods

Many researchers used hybrid approaches for developing virtual machine placement methods. For example, Chang et al. [66] attempted to optimize resource wastage and power consumption by using dynamic programming and a knapsack 0/1 policy. They considered physical machines like knapsacks and virtual machines as their items. They formulated virtual machine placement problem into a Multi stage knapsack sub problems that includes physical machine selection, virtual machine list sorting and dynamic placement. In their approach virtual computing resources are allocated to larger virtual machine resource requests. They proposed to select suitable physical machine based upon their ability evaluation function during dynamic placement sub problem. If a physical machine is unable to accept virtual machine then virtual machine request will be rains in descending order and accordingly virtual machine list is allocated to a physical machine that can afford it. In their experiments using cloudSim simulator they compared the proposed algorithm with openstack maximum remaining strategy method. They demonstrated the superior performance of their method in comparison to openstack maximum remaining statically method.

Zheng et al. [67] proposed a hybrid multi objective problem using BBO and differential evolution method for effective virtual machine placement. Their hybrid approach considered three objectives of optimizing power consumption load balancing in cloud data centre and decreasing the transfer time. In their approach they propose day modified operators for BBO algorithm and applied on migration model. They demonstrated that their modified version of BBO and differential evolution method resulted in acceptable migration time compared to other conventional algorithm, like genetic algorithm, ACO and differential evolution.

Qian et al. [68] introduced a hybrid model of genetic algorithm and simulated annealing algorithm for optimizing network traffic communication and operating cost in cloud data centres. They use genetic algorithm to optimized hardware cost and used it for Global search. In the next phase, simulated annealing algorithm is utilized for changing the position of components in virtual machine and used it for local search. In their experiment, they demonstrated that for small components their approach has little effect compared to the other algorithms regarding solution quality. However, there proposed hybrid approach produce better quality solution in compare two single genetic and single simulated annealing as algorithm.

Ariyan et al. [69] proposed a fuzzy multi objective DVFS-aware approach per optimizing a cloud data centre's energy consumption and carbon footprints. Their main focus was to optimize power consumption, service level agreement violations and virtual machine migrations.

Saber et al. [70] presented a H2D2 approach based upon multilayer design to reallocate virtual machines in different hosting departments of cloud data centres. They focus on different objectives in their approach, such as migration cost of virtual machines, cost of hosting virtual machine and reliability of cloud services. They applied non-dominated method for solving multi objective problem and handling multi-objectives.

Regaieg et al. [71] suggested a hybrid approach by using mixed integer linear programming for minimizing virtual machine rejection ratio, number of physical machines, and resource wastage.

Different approaches used for hybrid virtual machine placement can be summarized in Table 7.

5. Discussion, research challenges and future directions

The main objective of this study is to examine and evaluate different metaheuristic algorithms for solving virtual machine placement problems in the cloud computing environment.

Metaheuristic algorithms are a promising solution in tackling complex problems with conflicting objectives, such as virtual machine placement problems [2]. However, these algorithms have a limitation of returning sub-optimal solutions requiring trade off between exploration and exploitation that are difficult to obtain.

It can be noted from different tables (cited above) summarizing different types of metaheuristic approaches used for virtual machine placement that most office urges focused on optimizing Power consumption, virtual machine migrations, service level agreement violations, performance degradation after migrating virtual machine, and resource usage. These parameters are considered primary factors for designing virtual machine placement policies. In addition, two researchers also focused on other significant parameters such as security, load balancing, and computing resource wastage and network traffic management. Hybrid metaheuristic approaches benefit from different algorithms and have been validated to solve virtual machine placement problems.

In future, hybrid metaheuristic approaches can be explored as a promising research direction to solve multi-objective virtual machine placement problems in the cloud computing environment

6. Conclusion and future directions

This work presents a comprehensive review of the metaheuristic methods used to solve virtual machine placement problems. It analyses virtual machine placement methods based upon metaheuristic algorithms from different perspectives. It can be concluded from this analysis that most researchers focused on optimizing the energy consumption of cloud data centres. However, some researches also focused on other significant parameters such as computing resource utilization, execution time etc.

The primary objective of virtual machine placement methods includes reducing the energy consumption of data centres without any impact on cloud users' service level agreement, leading to an enhanced return on investment for cloud service providers.

Table 7. Summary of hybrid virtual machine placement methods

Study	Optimization objectives	Optimization method
Chang et al. [66]	• Resource wastage • Energy consumption	Dynamic programming and applied knapsack 0/1 method
Zheng et al. [67]	• Power consumption • Load balance • Migration time	BBO and DE
Qian et al. [68]	• Hardware costs • Communication overheads	GA and SA
Arianyan et al. [69]	• Energy consumption • SLA violations • VM migrations	DVFS, fuzzy logic and energy aware approach
Saber et al. [70]	• VM migration cost • reliability of servers • VM hosting cost	GRASP, NSGA, local search
Regaieg et al. [71]	• VM rejection ratio • Resource wastage • Number of used PMs	Mixed integer linear programming

This review provides a classification of meta heuristic algorithms and analyses the recent research in virtual machine placement in the cloud computing environment. The identified meta heuristic algorithm based virtual machine placement methods have been compared differently.

This review concludes to initiate research on multi objective virtual machine placements considering multiple and conflicting objectives of cloud data centres simultaneously, such as optimizing network traffic, energy consumption, load balancing and computing resource utilization. Hybrid metaheuristic algorithms can also be explored to enhance the quality of service by defining effective and efficient virtual machine placement policies. Future research can focus on addressing the following issues.

- Security: it is a significant factor that many researchers in the virtualization area have ignored. However, malicious users can access confidential data from a shared pool of physical resources in a cloud computing environment.
- Workload prediction: accurate prediction of future workload in cloud computing should be explored in future research based upon historical data. Accurate prediction of workload can reduce virtual machine migrations and, hence, lower computational overhead.
- Inter cloud communication: the future search must also consider inter cloud communication while placing and migrating virtual machines by considering conflicting policies of different organizations in different clouds.

References

[1] Masdari, M., Gharehpasha, S., Ghobaei-Arani, M., & Ghasemi, V. (2020). Bio-inspired virtual machine placement schemes in cloud computing environment: taxonomy, review, and future research directions. Cluster Computing, 23(4), 2533-2563.

[2] Gabhane, J. P., Pathak, S., & Thakare, N. M. (2021). Metaheuristics Algorithms for Virtual Machine Placement in Cloud Computing Environments—A Review. Computer Networks, Big Data and IoT, 329-349.

[3] Zheng Q et al (2016) Virtual machine consolidated placement based on multi-objective biogeography-based optimization. Future Gener Computer Syst 54:95–122

[4] Da Cunha Rodrigues G et al (2016) Monitoring of cloud computing environments: concepts, solutions, trends, and future directions. In: Proceedings of the 31st annual ACM symposium on applied computing

[5] Fatima A et al (2019) An enhanced multi-objective gray wolf optimization for virtual machine placement in cloud data centers. Electronics 8(2):218

[6] Ghobaei-Arani, Mostafa, Shamsi, Mahboubeh, Rahmanian, Ali A.: An efficient approach for improving virtual machine placement in cloud computing environment. J. Exp. Theor. Artif. Intell. 29(6), 1149–1171 (2017)

[7] Abreu DP et al (2019) A comparative analysis of simulators for the cloud to fog continuum. Simul Modell Pract Theor:102029

[8] Chowdhury MR, Mahmud MR, Rahman RM(2015)Implementationandperformanceanalysis of various VM placement strategies in CloudSim. J Cloud Comput 4(1):20

[9] Beloglazov A, Buyya R(2012)Optimalonlinedeterministicalgorithmsandadaptiveheuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurr Comput Practice Experience 24(13):1397–1420

[10] Donyagard Vahed, N., Ghobaei-Arani, M., & Souri, A. (2019). Multiobjective virtual machine placement mechanisms using nature-inspired metaheuristic algorithms in cloud environments: A comprehensive review. International Journal of Communication Systems, 32(14), e4068.

[11] Li Z, Li Y, Yuan T, Chen S, Jiang S. Chemical reaction optimization for virtual machine placement in cloud computing. Appl Intell. 2019;49(1):220-232.

[12] Beloglazov A, Buyya R, Lee YC, Zomaya A. A taxonomy and survey of energy-efficient data centers and cloud computing systems. In: Advances in Computers. Vol.82 Elsevier; 2011:47-111.

[13] Khan MA, Paplinski A, Khan AM, Murshed M, Buyya R. Dynamic virtual machine consolidation algorithms for energy-efficient cloud resource management: a review. In: Sustainable Cloud and Energy Services. Cham: Springer; 2018:135-165.

[14] Ahmad RW, Gani A, Hamid SHA, Shiraz M, Yousafzai A, Xia F. A survey on virtual machine migration and server consolidation frameworks for cloud data centers. J Netw Comput Appl. 2015;52:11-25.

[15] Varasteh A, Goudarzi M. Server consolidation techniques in virtualized data centers: a survey. IEEE Syst J. 2017;11(2):772-783.

[16] Usmani Z, Singh S. A survey of virtual machine placement techniques in a cloud data center. Procedia Comput Sci. 2016;78:491-498.

[17] Masdari M, Nabavi SS, Ahmadi V. An overview of virtual machine placement schemes in cloud computing. J Netw Comput Appl. 2016;66:106-127.

[18] Alboaneen, D.A., Tianfield, H. and Zhang, Y., 2016, July. Metaheuristic approaches to virtual machine placement in cloud computing: a review. In Parallel and Distributed Computing (ISPDC), 2016 15th International Symposium on (pp. 214-221). IEEE

[19] Xu M, Tian W, Buyya R. A survey on load balancing algorithms for virtual machines placement in cloud computing. Concurrency and Computation: Practice and Experience. 2017 Jun 25;29(12):e4123.

[20] Al-Dulaimy A, Itani W, Zekri A, Zantout R. Power management in virtualized data centers: state of the art. J Cloud Comput. 2016 Dec;5(1):6.

[21] Shahapure NH, Jayarekha P (2018) Distance and traffic based virtual machine migration for scalability in cloud computing. Proc Comput Sci 132:728–737

[22] Gahlawat, M., & Sharma, P. (2014, February). Survey of virtual machine placement in federated clouds. In 2014 IEEE International Advance Computing Conference (IACC) (pp. 735-738). IEEE.

[23] Buyya, R., A. Beloglazov, and J. Abawajy, Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges. arXiv preprint arXiv:1006.0308, 2010.

[24] Lopez-Pires, F., & Baran, B. (2015). Virtual machine placement literature review. arXiv preprint arXiv:1506.01509.

[25] Alashaikh, A., Alanazi, E., & Al-Fuqaha, A. (2021). A Survey on the Use of Preferences for Virtual Machine Placement in Cloud Data Centers. ACM Computing Surveys (CSUR), 54(5), 1-39.

[26] Attaoui W, Sabir E (2018) Multi-criteria virtual machine placement in cloud computing environments: a literature review. arXiv:1802.05113

[27] Rong, H., Zhang, H., Xiao, S., Li, C., & Hu, C. (2016). Optimizing energy consumption for data centers. Renewable and Sustainable Energy Reviews, 58, 674-691.

[28] Agarwal M, Srivastava GMS (2016) A genetic algorithm inspired task scheduling in cloud computing. Proc IEEE Int Conf Comput Com-mun Autom

[29] Masdari, M., Salehi, F., Jalali, M., & Bidaki, M. (2017). A survey of PSO-based scheduling algorithms in cloud computing. Journal of Network and Systems Management, 25(1), 122-158.

[30] Ebadifard, F., & Babamir, S. M. (2017, April). Optimizing multi objective based workflow scheduling in cloud computing using black hole algorithm. In 2017 3th International Conference on Web Research (ICWR) (pp. 102-108). IEEE.

[31] Pietri, I., Chronis, Y., & Ioannidis, Y. (2017, December). Multi-objective optimization of scheduling dataflows on heterogeneous cloud resources. In 2017 IEEE International Conference on Big Data (Big Data) (pp. 361-368). IEEE.

[32] Beloglazov, A., Abawajy, J., & Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future generation computer systems, 28(5), 755-768.

[33] Li X, Qian Z, Lu S, Wu J. Energy efficient virtual machine placement algorithm with balanced and improved resource utilization in a data center. Math Comput Model. 2013;58(5-6):1222-1235.

[34] Jamali S, Malektaji S, Analoui M. An imperialist competitive algorithm for virtual machine placement in cloud computing. J Exp Theor Artif Intell. 2017;29(3):575-596.

[35] Xu J, Fortes J. A multi-objective approach to virtual machine management in datacenters. In: Proceedings of the 8th ACM International Conference on Autonomic Computing. ACM; 2011, June:225-234.

[36] Addya SK, Turuk AK, Sahoo B, Sarkar M, Biswash SK. Simulated annealing based VM placement strategy to maximize the profit for cloud service providers. Int J Eng Sci Technol. 2017;20(4):1249-1259.

[37] Geronimo, G.A., Uriarte, R.B. and Westphall, C.B., 2016, April. Order@ Cloud: A VM organization framework based on multi-objectives placement ranking. In Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP (pp. 529-535). IEEE.

[38] Baalamurugan KM, Bhanu SV. A multi-objective krill herd algorithm for virtual machine placement in cloud computing. J Supercomput. 2018;1-18.

[39] Mollamotalebi M, Hajireza S. Multi-objective dynamic management of virtual machines in cloud environments. J Cloud Comput. 2017;6(1):16.

[40] Song A, Fan W, Wang W, Luo J, Mo Y. Multi-objective virtual machine selection for migrating in virtualized data centers. In: Joint International Conference on Pervasive Computing and the Networked World. Berlin, Heidelberg: Springer; 2012, November:426-438.

[41] Xu, J. and Fortes, J.A., 2010, December. Multi-objective virtual machine placement in virtualized data center environments. In Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom) (pp. 179-188). IEEE.

[42] Wang, S., Gu, H. and Wu, G., 2013, July. A new approach to multi-objective virtual machine placement in virtualized data center. In Networking, Architecture and Storage (NAS), 2013 IEEE Eighth International Conference on (pp. 331-335). IEEE.

[43] Wang X, Wang Y, Cui Y. A new multi-objective bi-level programming model for energy and locality aware multi-job scheduling in cloud computing. Futur Gener Comput Syst. 2014;36:91-101.

[44] Liu, C., Shen, C., Li, S. and Wang, S., 2014, June. A new evolutionary multi-objective algorithm to virtual machine placement in virtualized data center. In Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on (pp. 272-275). IEEE.

[45] Sofia AS, Kumar G. P. Multi-objective task scheduling to minimize energy consumption and makespan of cloud computing using NSGA-II. J Netw Syst Manag. 2018;26(2):463-485.

[46] Riahi M, Krichen S. A multi-objective decision support framework for virtual machine placement in cloud data centers: a real case study. J Supercomput. 2018;74(7):1-32.

[47] Gao Y, Guan H, Qi Z, Hou Y, Liu L. A multi-objective ant colony system algorithm for virtual machine placement in cloud computing. J Comput Syst Sci. 2013;79(8):1230-1242.

[48] Malekloo M, Kara N. Multi-objective ACO virtual machine placement in cloud computing environments. In: Globecom Workshops (GC Wkshps), 2014. IEEE; 2014, December:112-116.

[49] Malekloo MH, Kara N, El Barachi M. An energy efficient and SLA compliant approach for resource allocation and consolidation in cloud computing environments. Sustainable Computing: Informatics and Systems. 2018;17:9-24.

[50] Tan M, Chi C, Zhang J, Zhao S, Li G, Lü S. An energy-aware virtual machine placement algorithm in cloud data center. In: Proceedings of the 2nd International Conference on Intelligent Information Processing. ACM; 2017, July:1.

[51] Zhang L, Wang Y, Zhu L, Ji W. Towards energy efficient cloud: an optimized ant colony model for virtual machine placement. J Commun Inf Networks. 2016;1(4):116-132.

[52] Pham NMN, Le VS. Applying ant colony system algorithm in multi-objective resource allocation for virtual services. J Inf Telecommun. 2017;1(4):319-333.

[53] Ashraf A, Porres I. Multi-objective dynamic virtual machine consolidation in the cloud using ant colony system. Int J Parallel Emergent Distrib Syst. 2018;33(1):103-120.

[54] Zheng, Q., Li, R., Li, X. and Wu, J., 2015, May. A multi-objective biogeography-based optimization for virtual machine placement. In Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on (pp. 687-696). IEEE.

[55] Zheng Q, Li R, Li X, et al. Virtual machine consolidated placement based on multi-objective biogeography-based optimization. Futur Gener Comput Syst. 2016;54:95-122.

[56] Li R, Zheng Q, Li X, Yan Z. Multi-objective optimization for rebalancing virtual machine placement. Futur Gener Comput Syst. 2017.

[57] Ramezani, F., Naderpour, M. and Lu, J., 2016, July. A multi-objective optimization model for virtual machine mapping in cloud data centres. In Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on (pp. 1259-1265). IEEE.

[58] Luo J, Song W, Yin L. Reliable virtual machine placement based on multi-objective optimization with traffic-aware algorithm in industrial cloud. IEEE Access. 2018;6:23043-23052.

[59] Xu B, Peng Z, Xiao F, Gates AM, Yu JP. Dynamic deployment of virtual machines in cloud computing using multi-objective optimization. Soft Comput. 2015;19(8):2265-2273.

[60] Guo L, He Z, Zhao S, Zhang N, Wang J, Jiang C. Multi-objective optimization for data placement strategy in cloud computing. In: International Conference on Information Computing and Applications. Berlin, Heidelberg: Springer; 2012, September:119-126.

[61] Li Z, Yan C, Yu L, Yu X. Energy-aware and multi-resource overload probability constraint-based virtual machine dynamic consolidation method. Futur Gener Comput Syst. 2018;80:139-156.

[62] López-Pires F, Barán B. Multi-objective virtual machine placement with service level agreement: a memetic algorithm approach. In: Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing. IEEE Computer Society; 2013, December:203-210.

[63] López-Pires F, Barán B. Many-objective optimization for virtual machine placement in cloud computing. In: Research Advances in Cloud Computing. Singapore: Springer; 2017:291-326.

[64] Li, X.K., Gu, C.H., Yang, Z.P. and Chang, Y.H., 2015, December. Virtual machine placement strategy based on discrete firefly algorithm in cloud environments. In Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2015 12th International Computer Conference on (pp. 61-66). IEEE.

[65] Ding W, Gu C, Luo F, et al. DFA-VMP: an efficient and secure virtual machine placement strategy under cloud environment. Peer-to-Peer Network Appl. 2018;11(2):318-333.

[66] Chang, Y., Gu, C. and Luo, F., 2016, October. A novel energy-aware and resource efficient virtual resource allocation strategy in IaaS cloud. In Computer and Communications (ICCC), 2016 2nd IEEE International Conference on (pp. 1283-1288). IEEE.

[67] Zheng, Q., Li, J., Dong, B., Li, R., Shah, N. and Tian, F., 2015, December. Multi-objective optimization algorithm based on BBO for virtual machine consolidation problem. In Parallel and Distributed Systems (ICPADS), 2015 IEEE 21st International Conference on (pp. 414-421). IEEE.

[68] Qian, B., Meng, F. and Chu, D., 2015, December. A cost-driven multi-objective optimization algorithm for SaaS applications placement. In Smart City/SocialCom/SustainCom (SmartCity), 2015 IEEE International Conference on (pp. 1086-1091). IEEE.

[69] Arianyan E, Taheri H, Khoshdel V. Novel fuzzy multi-objective DVFS-aware consolidation heuristics for energy and SLA efficient resource management in cloud data centers. J Netw Comput Appl. 2017;78:43-61.

[70] Saber T, Thorburn J, Murphy L, Ventresque A. VM reassignment in hybrid clouds for large decentralized companies: a multi-objective challenge. Futur Gener Comput Syst. 2018;79:751-764.

Regaieg, R., Koubaa, M., Osei-Opoku, E. and Aguili, T., 2018, July. Multi-objective mixed integer linear programming model for VM placement to minimize resource wastage in a heterogeneous cloud provider data center. In 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 401-406). IEEE.