Optimal Resource Allocation of Cluster using Hybrid Grey Wolf and Cuckoo Search Algorithm in Cloud Computing

Devagnanam J  
Thiruthangal Nadar College,  
Chennai, Tamil Nadu, India  
devagnanam.j@gmail.com

Elango N M  
Vellore Institute of Technology,  
Vellore, Tamil Nadu, India

Abstract: In the progress of the markets and enterprises, cloud computing is considered the novel technology that has a great well-known factor. In cloud computing, the main effort is the allocation of resources. Moreover, for the task, the optimal allocation of resources is that it allocates the optimal appropriate cluster resources to perform in contemplation for various parameters, like scalability, cost, time, availability, resource utilization, throughput, reliability, etc. For the allocation of resources, an optimization algorithm is presented in this paper for the cloud environment through the Hybrid Grey Wolf and Cuckoo Search Algorithm (Hybrid GW-CS). For cloud computing, the hybridization algorithm is on the basis of the resource allocation which accumulates the execution and run time; also it enhances the profits for the cloud provider. Finally, the proposed Hybrid GW-CS based allocation of resource algorithms are evaluated over the conventional PSO, GWO, and CS by exploiting the performance measures like CPU usage rate, profit, and memory usage rate.

Keywords: Cloud Computing; Resource Allocation; Memory Utilization; Profit Evaluation; CPU Utilization

Nomenclature

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>Cuckoo search</td>
</tr>
<tr>
<td>GWO</td>
<td>Grey wolf optimizer</td>
</tr>
<tr>
<td>MSP</td>
<td>Mobile Service Provider</td>
</tr>
<tr>
<td>EC</td>
<td>Extra clouds</td>
</tr>
<tr>
<td>C-RAN</td>
<td>Cloud Radio Access Network</td>
</tr>
<tr>
<td>OBL</td>
<td>Opposition-based learning</td>
</tr>
<tr>
<td>PD-NOMA</td>
<td>Power-Domain Non-Orthogonal Multiple Access</td>
</tr>
<tr>
<td>RA</td>
<td>Resource Allocation</td>
</tr>
<tr>
<td>PaaS</td>
<td>Platform as a Service</td>
</tr>
<tr>
<td>IaaS</td>
<td>Infrastructure as a Service</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>MUC</td>
<td>Modularity-based User-Centric</td>
</tr>
<tr>
<td>MEC</td>
<td>Mobile Edge Cloud computing</td>
</tr>
<tr>
<td>UDNs</td>
<td>Ultra-Dense Networks</td>
</tr>
<tr>
<td>SE</td>
<td>Spectral Efficiency</td>
</tr>
<tr>
<td>APs</td>
<td>Access Point</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>RAS</td>
<td>Resource Allocation Strategy</td>
</tr>
<tr>
<td>RB</td>
<td>Resource Block</td>
</tr>
<tr>
<td>D2DC</td>
<td>Device-to-Device Cluster</td>
</tr>
<tr>
<td>SaaS</td>
<td>Software as a Service</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
</tbody>
</table>

1. Introduction

Nowadays, cloud and distributed computing are advancing technology. In several applications, cloud computing can be employed that consists of IoT applications, data analytics, and storing data [1]. Cloud computing is a technology that involves evolved conventional approaches in that functions are set up by individuals or enterprises. It presents various services types like web services, which are registered to users as a result users do not require establishing in computing framework. Cloud computing offers
services like IaaS, PaaS, and SaaS. Through the medium of the Internet, the users are supposed to propose the service provider requests in each type of service. To satisfy requests proposed by users, the service provider is subject to carrying out the resources. To set up the incoming request (tasks), service providers use the scheduling methods in order to deal with their computing resources adequately. Resource management and task scheduling enable providers to augment revenue and the employment of resources for their restrictions. In a study, regarding cloud computing resources performance, allocation, and resource scheduling, are significant hurdles [2]. Hence, numerous investigators were fascinated by cloud computing applications for scheduling the task.

In cloud computing, RA is to deal with the allocation of available resources to the necessary cloud applications over the Internet [16]. RA starves services if the allocation is not dealt with accurately. For each individual module, to carry out the resources, resource provisioning explains that issue by providing the service providers. RAS is all regarding connecting cloud provider activities for employing and allocating scarce resources within reducing of cloud environment with the intention of finding a requirement for the cloud application [17]. It causes the type and number of resources required by each application so as to develop a user job. For an optimal RAS, the time and order of allocation of resources are further considered as input [4].

As per their requirements for a given period of time, RA can be described as accurately distributing the resources between multiple users [18]. Nevertheless, in cloud computing, RA has turned out to be a bit obscured. Then, to develop the computing capacity, there is a requirement for resource allocation. The primary purpose of smartly resources allocation is to achieve economic profits in the market. This method further advances the purposes of cloud computing that are paid as peruse since the client required not to pay for the resources that he becomes not employed [19]. Dynamic RA hits up the workflow implementation and assigns the users to comprehend among various policies feasible [2].

In this work, a cluster RA model using the exterior clouds in the requirement, while the private cloud resources are not adequate is developed. A novel analytical analysis is devised by exploiting the multiple objective criteria like resource model, fiscal mode, and user-defined model consisting the criteria like memory usage and CPU, deadline and runtime circumstances and deployment and cost of operating the cloud, which leads the optimization issue in RA.

To solve the optimization issue, a Hybrid GW-CS approach on the basis of the RA algorithm is developed. The final aim of the developed approach is to obtain an optimal solution for RA, which depends on the cloud consumer request with minimum computational time for the cloud provider using the resources with the utmost profit.

The main contribution of this work is to propose a novel fitness model is devised for the optimal RA strategy. The proposed fitness model proposes the income and allocation cost increase the cloud provider profit in RA, moreover, rate of the CPU usage and the rate of memory usage are further dealt with the fitness model to create the RA high understandable when controlling the rate of the maximum utilization of the system.

2. Literature Review

In 2019, Mylene Pischella and Didier Le Ruyet, concentrates in multi-carrier uplink networks on RB and clustering allocation employing PD-NOMA. PDNOMA was exploited per frame and RB merely if the attained individual rates of data were higher with PD-NOMA than the TDMA. The clustering approach recursively exploits the utmost weight corresponding to construct clusters per RB. Certainly, clusters, allocation of RB were subsequently optimal. The developed technique gives a better trade-off among minimized generally latency of network and enhancement rate evaluated with the two reference approaches.

In 2018, Jong-ho Kim et al [2], designed a basically competent RA method for multiple D2DC multicast communications. The transmission of the channel, and the power, were assigned to D2DC to utmost the summation of effectual throughput offered which the cellular communications uphold a convinced QoS level. By exploiting the partial information of device positions, they had formulated a channel allocation method. Hence, the outage probability, and an effectual D2DC communications throughput, was derived in estimated models.

In 2018, Yan Lin et al [3] worked on a new MUC clustering that was envisaged for RA in UDNs, with the intention of exploiting the summation-rate of per orthogonal RB. By using the inherent group model of user types of equipment, the Modularity-based User-Centric clustering design was exploited for the decomposing of the Ultra-Dense Networks into numerous sub-networks. Especially, an enhanced Louvain algorithm was presented for Modularity-based User-Centric relying on clustering on well-organized RA heuristics. Finally, the simulation outcomes exhibit the advantage of the proposed MUC model.
In 2018, Yan Lin et al [4], developed a new distributed three-phase solution for user-centric clustering that was attentive to the APs traffic-loads. Subsequently, a novel two-phase graph-based user-centric RA method on the basis of the overlapped clusters constructed was presented for extenuating the ensuing inter-cluster interference, in spite of the inadequate ease of use of orthogonal RBs.

In 2018, Xinhou Wang et al [5], developed a unifying model for MSP in cooperation with resources of scheduling networks in C-RAN and computation resources in MEC to exploit the MSP profit. To attain this purpose, the scheduling of resource problem was devised as a stochastic issue and by exploiting an extended Lyapunov model a new optimization method was designed. Particularly, due to the typical Lyapunov model significantly presumes. Moreover, it was not appropriate for the dynamic circumstances whereas the requests mobile job had lengths for the variable. In order to resolve this issue, the conventional Lyapunov approach and model the Varied Len approach, which was extended in this paper.

3. Cloud Computing Model

The resource allocation in cloud computing is a vital component due to the allocation of demands to persuade the demands of the multiple tasks for the request of the cloud consumer within the goal and to create the additional sensible resources allocation although all the resources are used in the cloud. Moreover, the private cloud uses its individual resources by considering cloud service providers, and the private cloud EC offers the resource in the shortage of circumstances (resource restriction) [6]. In Fig. 1, the cloud model of resource allocation is demonstrated. Fig. 1 indicates the hybrid cloud environment comprising of the EC and the private cloud. At first, from the cloud, the cloud consumer requests for the resources of computing. The user interface module is considered as an interface in a private cloud; hence the request of user application for the cloud consumer is collected using the cloud. In cloud task buffer, the tasks which are requested by the user are saved. In the cloud, the request manager module completes the control of all the approved requests of the user.

The state of non-availability of resources for inward application task is attained while all the private cloud resources are used. The private cloud looks for the EC resource due to the resources cost model offered using the EC. By means of the cloud interface, the EC cost model is composed since on the basis of the price merely the resource of the external is exploited for the allocation of a resource without distressing the cloud provider income in the private cloud.

In the private cloud, the allocator or scheduler is the significant module available. For the task allocation, in the requirement of the EC, to the scheduler, ECs pricing models are transmitted during the interface of the cloud, hence outsourcing the precise EC is chosen.

In RA of cloud computing, the major purpose is the resource allocation for independent bag tasks. In the applications of the enterprises, the independent bag tasks are widespread like customer behavior mining, analysis of the sensor data [6].

Consider cloud providers set as C for diverse application user requests and C = {C₁, C₂, ...., Cₘ}; n number of cloud providers in that presumes CPi as the private cloud and {C₂, C₃, ...., Cₘ} as the EC's. The VM, V = {V₁, V₂, ...., Vₙ} is a set of the VM categories given the provider of the cloud for the implementation of the equivalent application tasks. Consider Λ as the application set, here the user request that is acknowledged using the cloud environment, and it is stated using Λ = {a₁, a₂, ...., aₓ}. Every application aᵢ (i ∈ {1, 2, 3, ...., x}) possesses a runtime RTᵢ and a deadline Dᵢ. Based on the cost CTᵢ, the categorization of the task execution is performed. For the exacting task, the RA by cloud provider must be performed with characteristic consideration to the run time and deadline.

In the cloud, the application is requested by the user which comprises the task set Taskᵢ = {tᵢ₁, tᵢ₂, ...., tᵢₜᵢ}. In the developed analytical analysis, the time is indicated at the slot of time with 1-hour granularity. Considered Maxᵢ as the utmost amount of the time slot, and it is stated using Maxᵢ = maxᵢₑ∈[₁,₂,₃,....ₓ]|Dᵢₑ|. By the cloud provider, the main purpose is the allocation of resources for the application aₓ so that to make the most of the profit for the private cloud. For the applications, the individual task is assigned as the Cloud providers.
3.1. Optimal Resource Allocation for Cluster

In this section, the optimized cluster RA on the basis of the proposed Hybrid GW-CS is shown. Here, in the runtime and the deadline, the task flow with constraint, the fitness calculation for the optimization, the Hybrid GW-CS algorithmic process is described.

**a) Allocation of resource:**
RA is defined as the processes of distributed computing, networking, storage; nodes and VM to the cloud consumer applications that comprise of several tasks in cloud computing [7].

**b) Importance of allocation of resource approach:**
The cloud computing service starves, while RA is not controlled accurately. An optimal RA model must shun the subsequent criterion [8]:
- Resource disputation happens while two applications of the request of the consumer ingress a similar resource at a duration.
- Resource over-provisioning is assigned the resource which is higher than the application requirement.
- Resources under-provisioning is distributed the resources which are lesser than the application demand.

3.2. Flow of Task

In the hybrid cloud environment, the flow of tasks is shown in this section. In Table 1, the flow of task stipulation is indicated. The allocation of the task to the VM V for the application of the user is executed based upon the VM availability and the deadline meeting criterion of the VM. Let task $t_i (i = 1,2,3)$ and the deadline of the task are implicit to be $D_{t_{i=1,2,3}} = \{1,2,3\}$ and the tasks run times are implicit to be $RT_{t_{i=1,2,3}} = \{1,1,1\}$.
At first, the user request the application from the cloud with 3 tasks for the computing of resource, the task \( t_1 \) that is with the minimum price and the cost is initially offered with the resource on the basis of the accessibility from the cloud. In the private cloud, the VM availability is verified, and during the first time in the private cloud resources are not used, and the task allocation is all set to achieve. Mainly the task deadline is verified, and, if \( V \) have the ability to convene the deadline, and it is assigned as the resource for the task \( t_1 \). The \( V_1 \) resources are used with the intention that for promoting inward applications, the private cloud follows the aid of the EC for the allocation of the resource owing to the resource limitations in private regarding the cost form of the EC’s.

### 3.3. Fitness Function

The fitness function for the optimized RA issue consists the decision variable and parameters, like the RA income, EC’s resources cost, the rate of memory usage and the rate of CPU usage to set up the equitable resources allocation obtain majority out of the cloud provider profit [9]. Eq. (1) represents the fitness function for the optimization issue.

**Fitness function,**

\[
F_{\text{fit}} = \alpha \cdot f_{n_1} + \beta \cdot f_{n_2} \tag{1}
\]

The fitness function \( f_{n_1} \) and \( f_{n_2} \) indicates the two objective functions as stated in eq. (1), whereas \( f_{n_1} \) indicates related to the income and variables of the cost decision and \( f_{n_2} \) indicates related to the CPU utilization rate and memory.

**Objective model \( fn_1 \):**

The objective model \( f_{n_1} \) is associated with the RA cost and income of the cloud provider. Moreover, RA is superior if the cloud provider income is large with the intention of that the objective model one in the fitness function is superior for the RA merely if the \( f_{n_1} \) value is high. The objective model one of the fitness function is calculated as below:

\[
fn_1 = (\text{Income} - \text{Cost}) \tag{2}
\]

Income = \[ \sum_{j \in t_i} \sum_{k \in V_m} R_{kj} \cdot P_{kj} \tag{3} \]

Cost = \[ \sum_{i = 1}^{n} \sum_{k \in V_m} R_{ji} \cdot C_{ij} \tag{4} \]

Here, \( PR_{ij} \) indicates the RA price to the task, \( RT_{ij} \) indicates the run time of the task and \( CT_{ij} \) indicates the RA cost to the task.

**Objective model \( fn_2 \):**

In the fitness estimate, objective model 2 is associated with the guarantee that the RA is done within the memory capacity and CPU, for each time slot.

The objective model 2 is calculated as below:

\[
fn_2 = \sum_{j = 1}^{i} \sum_{k = 1}^{n} \left[ \frac{N_{\text{CPU}}_{kj} \times M_{r_{kj}}}{T_{\text{CPU}} \times T_{M}} \right] \tag{5}
\]

In eq. (5), \( M_{r} \) and \( N_{\text{CPU}} \) indicates the amount of the memory size and CPU used for the RA and \( T_{M} \) and \( T_{\text{CPU}} \) indicates the ability of the memory size and CPU in the equivalent cloud provider.

Also, the objective model 2 should be high for the optimal fitness model; maximum indicates maximum Memory and CPU utilization up to the provider’s ability.
4. Proposed Methodology for Resource Allocation

From the user, the resource allocation for the inward application task of the request, an appropriate optimization approach is required for the allocation of the task to the VM with a developed fitness model. For this research, the optimization approach proposed is the second contribution and it is Hybrid GW-CS, which hybridizes the Grey Wolf optimization approach and Cuckoo Search algorithm.

4.1. GWO Algorithm

GWO [10] is a novel adopted a meta-heuristic approach that emulates the hunting wolf’s behavior of swarm. In Grey Wolf Optimization, the optimal individual is known as $\alpha$ wolf, $\beta$ represents the second-best individuals, and $\delta$ represents the third-best individuals, correspondingly, and $\omega$ represents the remaining individuals. The behavior of the swarm wolf encircles the prey and it is formulated as below [10]:

$$Y(t+1)=Y_p(t)-A_{\omega} \cdot Y_{\omega}(t)-Y(t)$$

(6)

In eq. (6), $Y$ indicates the location wolf vector, $t$ indicates the number of iterations, $Y_p$ is the location prey vector,

$$A = 2 \cdot a \cdot \gamma_1 - a$$

(7)

$$C = 2 \cdot \gamma_2$$

(8)

The eq. (7) and (8) represents the coefficient, $\gamma_1 \in [0, 1]$ and $\gamma_2 \in [0, 1]$ represents an arbitrary number and $a$ represents computed as eq. (9), $\max_{itr}$ indicates the maximum number of iterations.

$$a = 2 - \frac{2t}{\max_{itr}}$$

(9)

In the population, $\alpha$, $\beta$ and $\delta$ wolves are used to update the locations of the other individuals which is stated as below:

$$Y_1 = Y_a - A_1 \cdot C_1 \cdot \gamma_a - Y$$

(10)

$$Y_2 = Y_\beta - A_2 \cdot C_2 \cdot \gamma_\beta - Y$$

(11)

$$Y_3 = Y_\delta - A_3 \cdot C_3 \cdot \gamma_\delta - Y$$

(12)

$$Y(t+1)=\frac{Y_1(t)+Y_2(t)+Y_3(t)}{3}$$

(13)

Here, $A_1$, $A_2$, and $A_3$ are approaching $A$, $C_1$, $C_2$, and $C_3$ are close to $C$.

4.2. CS Algorithm

CS [11] is an extensively exploited meta-heuristics approach that imitates the behavior of cuckoos breeding parasitism. In CS, a cuckoo’s egg stands for a solution. In the iterative process, the novel candidate solution is produced using Lévy flight is stated as below [11]:

$$Y_i = Y_i - \gamma \cdot (Y_i - Y_h) \oplus \text{Levy}(\lambda)$$

$$= Y_i + \frac{0.01v}{|u|^{\lambda/2}} (Y_i - Y_h)$$

(14)

In eq. (14), $Y_i$ represents the $i$th solution, $Y_h$ represents the global optimal solution, $\gamma > 0$ represents the step scaling size, $\oplus$ indicates the multiplications of entry-wise, $\lambda$ indicates Levy flight exponent, $v$ and $u$ indicates arbitrary numbers correspondingly and they are fulfilled with normal distribution [11]:

$$v \sim \mathcal{N}(0, \sigma_v^2) u \sim \mathcal{N}(0, \sigma_u^2)$$

(15)

$$\sigma_v = \left[ \sin\left(\frac{\sigma_2^{\lambda}}{2}\right) \frac{\Gamma(1+\lambda)}{\sqrt{\lambda} \cdot \Gamma\left(\frac{1+\lambda}{2}\right)} \right]$$

(16)

In eq. (16), $\Gamma(\cdot)$ indicates the Gamma function. In addition, CS further uses the operator of the detection to reinstate the identified nests with probability $p_0$ are stated as below [11]:

$$Y_i = \begin{cases} Y_i + m \cdot (Y_j - Y_k) & \text{if } p > p_0 \\ Y_i & \text{else} \end{cases}$$

(17)
In eq. (17), \( p \in [0, 1] \) indicates an arbitrary number, \( Y_j \) and \( Y_k \) indicates the candidate solutions from the population, correspondingly.

### 4.3. Opposition Learning Scheme

In numerous engineering fields, the GWO approach has been effectively used in [12]. Nevertheless, in accordance with [13], at local exploitation the fundamental GWO approach is superior; at global exploration, it excluded the underprivileged. Hence, one important research study is to improve the global exploration capability of GWO. This scheme [14] is an effective exploration-improved approach and it has been extensively used meta-heuristic approaches to reinforce their global exploration capability. Moreover, this scheme is described below:

**Description 1:** Opposite number: Taking into consideration of a real number \( y' \), its opposite number \( y' \) is described as below [14]:

\[
y' = lb + up - y
\]

(18)

In eq. (18), \( up \) and \( lb \) indicates the upper and lower boundary of \( y \), correspondingly.

**Description 2:** Opposite solution. The search space of \( D \)-dimensional can be extended using eq. (18). Presumptuous that \( Y = \{y_1, y_2, \ldots, y_D\} \) indicates a solution in the search space of \( D \)-dimensional. The opposite solution for \( Y \) is explained as \( Y' = \{y_1', y_2', \ldots, y_D'\} \) [14]:

\[
y'_i = lb_i + up_i - y_i \quad i = 1, 2, \ldots, D
\]

(19)

Ultimately, the values of fitness function \( f(Y) \) and \( f(Y') \) are computed. If \( f(Y) \)is superior to the \( f(Y') \), \( Y \) is chosen; or else, \( Y' \) is chosen. Further lately, the OBL scheme is used for the GWO approach to enhance the GWO performance [15]. Dissimilar from [15], this work uses the OBL scheme for the decision layer individuals \( \alpha, \beta \) and \( \gamma \) with probability \( p_b \) to more improve the population diversity. Fig 2 shows the flow chart of the proposed model.

### 5. Results and Discussions

#### 5.1. Experimental Procedure

In this section, the proposed algorithm for the RA in cloud computing is explained. In the cloud computing environment, the simulation of the proposed algorithm for the optimized RA is done on Windows 7 operating system. The experimentation of this study was exploited in JAVA. Here, for the simulation, one problem instances were modeled of the proposed hybrid algorithm.

Problem Instance 1 comprises of five applications. By each application, The Virtual Machine instance category requested is arbitrarily chosen from the VMs of the external or private clouds. With the intention to limit the search space, the deadline for every application is limited among arbitrary numeral of 1 and 7 hours.

#### 5.2. Performance Analysis

Fig 3 exhibits the graphical representation of the profit evaluation for problem instance. Here, the proposed method can accomplish the best solution in a rational time. During the 20th iteration, the profit obtained by the proposed method is high; simultaneously the traditional methods like PSO, GWO and CS, attained the less profit of correspondingly. Fig 4 demonstrates the rate of CPU utilization for the problem instance. During all the iteration, the rate of CPU utilization for the proposed technique is superior to the conventional algorithms. Fig 5 shows the memory utilization rate assessment. In the proposed algorithm, the task of application uses the utmost memory size present in the private cloud raising the cloud provider profit for all the iteration.

Fig. 2. Flow chart of the proposed hybrid GW-CS algorithm

Fig. 3. Graphical representation of the proposed technique concerning the profit

![Graphical representation of the proposed technique concerning the CPU utilizations](image1.png)

**Fig. 4.** Graphical representation of the proposed technique concerning the CPU utilizations

![Graphical representation of the proposed technique concerning the memory utilization rate](image2.png)

**Fig. 5.** Graphical representation of the proposed technique concerning the memory utilization rate

### 6. Conclusion

In a hybrid cloud computing environment, an optimal cluster RA algorithm was developed. Here, an optimization approach was proposed by combining the Grey wolf optimization approach and the Cuckoo Search approach known Hybrid GW-CS approach. Moreover, a novel fitness model was developed to reduce the cost of RA and to maximize the profit in addition to the utilization rate in the cloud computing environment for the resources. The hybrid GW-CS algorithm was capable to assign the resource to the tasks effectively on the basis of the proposed fitness model. The simulation of the proposed RA algorithm was examined with three issue cases. The performance of the proposed model was evaluated with the conventional models by exploiting the performance metrics such as rate of CPU usage, profit, and memory usage. The investigational consequences demonstrate that proposed Hybrid GW-CS optimization algorithm on the basis of the RA was effectual and competent for the sensible utilization of resources in cloud computing with a sensible running time.

### References


