Budapest University of Technology and Economics
Department of Networked Systems and Services

# Quantum Optimization Based Resource Distribution Management System

Ph.D. dissertation summary

Sara El Gaily

Supervised by
Prof. Dr. Sándor Imre (cmHAS)

Hungary, 2022

# 1. Introduction & Background

Quantum computing and communication aim to reach a reduction in terms of computational complexity and achieve optimum and efficient results with respect to the requirements of the given problem. It could solve problems that the classical computing counterpart could not deal with, as implementing a secure quantum communication channel, finding the prime number, search an item in an unsorted database using Grover's Method [1]-[2], …etc.

In engineering and science domains, solving optimization problems has a wide range of applications. Optimization algorithms aim to find the best extreme (maximum or minimum) outcome of an objective function or an unsorted database. It is worth noting that these optimization methods can be divided into two categories: unconstrained and constrained optimization methods. The majority of existing problems are computationally challenging. Quantum optimization algorithms, unlike their classical counterparts, use the principles of quantum mechanics to speed up computations.

The resource distribution management (RDM) challenges arise due to the need to maximizing the throughput; reducing the power consumption, the cost, the delay, and the computational complexity; reaching a load balance. The complexity degree of solving the optimization problem depends also on the dependencies nature of the adopted attributes defining the elements of the model, the architecture and the size of the system, and the type of the metric used, etc. For this purpose, many research has been driven to find solutions to the above-mentioned problem paradigms. Furthermore, there always exists a trade-off between all the challenging factors mentioned previously. Most of the existing methods of cloud resource distribution cannot guarantee a reduction in terms of computational complexity. Some metaheuristic algorithms are often used in resource allocation systems such as Ant Colony Optimization [3-5], Particle Swarm Optimization [6-10], Genetic Algorithm [11-14]. These algorithms achieve an efficient and accurate search but they perform a time-consuming computation.

It is important to point out that the RDM must be designed to be highly reacting fast with maximum accuracy performance to any unpredictable task workload. As it is known, any product has its life cycle, classical computing machines will be replaced soon by quantum computers, so there will be an intensive need for developing new adaptive models.

As it is clearly noticeable from the above discussion, selecting an efficient algorithm for the RDM system is an important task. However, finding a robust and powerful classical optimization algorithm that guarantees better results and less computational complexity is a challenging job. For this reason, my research work exploits quantum optimization techniques as a computational infrastructure for optimizing the RDM system. My objective is to highlight the benefits of using quantum methods as an embedded computational tool in the RDM system.

# 2. Research Goals

The starting point of my research work was to incorporate the quantum computing techniques and RDM system, to increase the performance and the QoS. By changing the classical decision-maker module with the quantum extreme value searching algorithm (QEVSA) [15-16]. The following challenge was to set up carefully and appropriately the parameters of this quantum strategy for optimizing an unconstrained metric function. The next goal was to show that the proposed optimization problem cannot be solved by utilizing classical computers and the only solution is to exploit the capabilities of the quantum nature of quantum optimization algorithms. The next step was to extend and develop new quantum optimization methods that can handle at the same time an unconstrained and constrained optimization problem, and exploit them in optimizing the RDM system. My research work also highlights the similarities and differences between the developed quantum optimization algorithms. My final task was to exploit the newly developed algorithms in optimizing the RDM system where queueing scenarios have been considered.

# 3. Methodology

The working methodology of my research work mainly relied on three steps, First, I have exploited the newly developed methods in RDM as a toy example for pointing out the efficiency and the strength of the proposed quantum strategies. Next, I have constructed an analytical mathematical model which fits at the same time the developed quantum method and the proposed RDM model. I have utilized mathematical tools that were mainly based on linear algebra, probability, and statistics. Finally, to test and validate the results obtained analytically, I have relied on extensive simulation environments using Matlab and Python.

# 4. New Results

## 4.1 Quantum optimization algorithms for continuous and non-continuous database structures

***Thesis Group 1: [C7, C8, C9, C10, J3] I developed two new quantum optimization algorithms, the first quantum method finds the extreme optimum result of an unconstrained goal function or an unsorted database, where the type of the database structure can be continuous or non-continuous. The second one is an extension of the previous algorithm which searches the extreme value of a constraint goal function or an unsorted database with respect to a certain constraint. Finally, I showed that these new quantum optimization algorithms can provide less computational complexity and high accuracy compared with the existing classical optimization methods.***

These two quantum optimization algorithms are inspired by the QEVSA, the efficiency of the QEVSA relies on exploiting the quantum existing testing (QET) in enhancing the performance of the binary searching algorithm (BSA) [17]. It is important to point out that the QEVSA can only search for extreme optimum results in a continuous database structure (CDS).

In many optimization problems, it is not expected that the database will always have a CDS. For this sake, I developed a new quantum optimization algorithm that finds the extreme optimum result of an unconstrained goal function or unsorted database, where the database may have a CDS or a non-continuous structure (NCDS), named Quantum Optimization Algorithm for a Non-Continuous Database Structure (QOANCDS). It is worth mentioning that the unsorted databases can be regarded as goal functions with discrete input such that the database entries can be associated with numbers. In order to optimize the constrained objective function, an extension quantum method of the QOANCDS has been developed, this method searches the extreme value of a constraint goal function, labeled Constrained Quantum Optimization Algorithm (CQOA). More details are presented in Chapter 2.

**Thesis 1.1**

*In this thesis [C10], I developed the QOANCDS. Next, I investigated the classical and quantum uncertainty of the QOANCDS as well as its computational complexity. Then, I showed that the computational power provided by the generalized version of QRT called quantum relation testing (QRT) method is responsible for the exponential speed and the efficiency of the QOANCDS.*

The QEVSA can only perform a search in the CDS. To support the search in the NCDS, I extended the QET technique to quantum relation testing $QRT(ref, R)$ which returns only an answer (YES or NO) whether there are at least one or more database entries in a given region of the database with respect to the applied index relation *R* and the reference value *ref*, where the parameter *ref* refers to the updated value which divides the database into two vertical sections, and the index relation *R* refers to the used relation (finding the minimum or maximum value). More details about the multiple choices of the relation *R* according to the given optimization problem are presented in Chapter 2.

The QOANCDS determines the extreme (maximum or minimum) optimal value of an unconstrained objective function or an unsorted database, the type of the relation *R* determines the type of the database structure (CDS or NCDS); the key to its effectiveness stems from merging the QRT with the BSA. The QOANCDS is presented in **Algorithm 1**. More details about the working mechanism of this quantum algorithm are presented in Chapter 2.

The proposed QOANCDS is introduced in detail as follows:

**Algorithm 1: QOANCDS**

1. We start with $S = 0$ : $T_{min1} = T_{min0}$, $T_{max1} = T_{max0}$, and $T = T_{max0} - T_{min0}$
2. $S = S + 1$
3. $T_{med\,S} = T_{minS} + \left[\frac{T_{maxS} - T_{minS}}{2}\right]$
4. $flag = QRT\ (T_{med\,S}, R)$:
    - $flag = Yes,$ then $T_{max\,S+1} = T_{med\,S}$, $T_{\min S+1} = T_{\min S}$
    - Else $T_{max\,S+1} = T_{max\,S}$, $T_{min\,S+1} = T_{med\,S}$
5. If $S < log_2\ (T)$, then go to 2, else stop and $y_{opt} = T_{med\,S}$

Where $T_{med\,S}$ refers to the value *ref*, and the stochastic parameter *T* is the maximum number of steps needed to run the BSA embedded in the QOANCDS.

**Thesis 1.2**

*In this thesis [C7, C8, C9, C10, J3], I developed the CQOA where I extended the QRT to the constrained quantum relation testing (CQRT) method. Next, I investigated the classical and quantum uncertainty of the CQOA as well as its computational complexity. Finally, I highlighted the relationship between the QET, QRT, and CQRT functions.*

Similarly to what has been discussed in Thesis 1.1, I extended the QRT version to the constrained quantum relation testing $CQRT(ref, R, C)$ which gives back only an answer whether there exist at least one or more database entries with respect to the applied relation *R* of the reference value *ref* and the constraint *C* in a certain region of the database.

The computational complexity of the CQRT is strongly connected to the classical and quantum uncertainty of the application. Similar to the QRT function, the CQRT is nothing else than the quantum phase estimation (QPE) in disguise. The CQRT checks whether there exists a certain item in a certain region of the database which fulfills at the same time $\Omega \neq 0$ ($M \neq 0$) and $C = 1$, i.e. if the constraint $C$ is fulfilled then $C = 1$, else $C = 0$. The total number of qubits used for the physical implementation of the CQRT is denoted by $n_{CR}$ and can be described as follows $n_{CR} = c_{CR} + p_{CR}$. It is interesting to note that involving the constraint $C$ in the QRT does not modify the conversion of the QPE. For this sake, the number of qubits required for the classical and quantum certainty of the QRT denoted by $n_R$ equals the one belonging to the CQRT method ($n_{CR} = n_R$). More details are given in Chapter 3.

The CQOA is presented in **Algorithm 2**. More details about the working mechanism of this quantum strategy are presented in Chapter 2.

---

**Algorithm 2: CQOA**

---

1. We start with $S = 0$ : $T_{min1} = T_{min0}$, $T_{max1} = T_{max0}$, and $T = T_{max0} - T_{min0}$
2. $S = S + 1$
3. $T_{med\ S} = T_{minS} + \left[\frac{T_{maxS} - T_{minS}}{2}\right]$
4. $flag = CQRT\ (T_{med\ S}, R, C)$:
   - $flag = Yes,$ then $T_{max\ S+1} = T_{med\ S}$, $T_{\min S+1} = T_{\min S}$
   - Else $T_{max\ S+1} = T_{max\ S}$, $T_{min\ S+1} = T_{med\ S}$
5. If $S < log_2\ (T)$, then go to 2, else stop and $y_{opt} = T_{med\ S}$

---

On the other hand, I showed that the special quantum core functions used in each quantum optimization algorithm (QEVSA, QOANCD, and CQOA) are connected by this formula $QET \subseteq QRT \subseteq CQRT$.

**Thesis 1.3**

*In this thesis [J3, J2, C8, C9, C10], I showed the similarities and differences between the QEVSA, QOANCDS, and CQOA in terms of computational complexity, database structure type, and goal function type. Next, I showed that these quantum optimization algorithms achieve a high computational complexity reduction compared with the existing classical optimization methods.*

I showed that the computational complexity of the QEVSA, QOANCDS, and CQOA are identical, it equals $O\left(log_2(T) log_2{}^3\left(\sqrt{N}\right)\right)$, and it depends on the computational complexity:

- The quantum core function $log_2{}^3\left(\sqrt{N}\right)$, i.e. the QET, QRT, or CQRT depending on the optimization problem and the database structure type, where $N$ refers to the entry size of the database.
- The BSA embedded in the quantum optimization algorithm $log_2(T)$, where $T$ is the maximum number of steps needed to run the BSA.

Table 1 summarizes the similarities and differences between the QEVSA, QOANCD, and CQOA.

Table 1: The similarities and differences between the QEVSA, QOANCD, and CQOA.

| | QEVSA | QOANCD | CQOA |
|---|---|---|---|
| Type of objective function (OF) | Unconstraint OF | Unconstraint OF | Constraint OF |
| Database Structure (sorted/unsorted) | Unsorted database | Unsorted database | Unsorted database |
| Computational complexity | $O\left(log_2(T)log_2{}^3(\sqrt{N})\right)$ | $O\left(log_2(T)log_2{}^3(\sqrt{N})\right)$ | $O\left(log_2(T)log_2{}^3(\sqrt{N})\right)$ |
| The BSA is combined with | QET | QRT | CQRT |
| Database type (CDS/NCDS) | CDS | CDS/NCDS | CDS/NCDS |

I compared the QEVSA with some classical optimization algorithms (the randomized, the exhaustive, and the sequential-searching algorithms). Table 2 presents different classical optimization algorithms and compares them with the QEVSA in terms of computational complexity and accuracy. While Figure 1 illustrates a comparison between the computational complexity of the exhaustive and QEVSA.

Table 2: Comparison between the classical heuristic algorithms and the QEVSA

| **Methods** | **Accuracy** | **Computational complexity** |
|---|---|---|
| Randomized | Very rough | $O(const)$ or $O(1)$ |
| Sequential | Very rough | $O(const)$ or $O(1)$ |
| Exhaustive | Exact | $O(N)$ |
| QEVSA | Close exact | $O\left(log_2(T)log_2{}^3(\sqrt{N})\right)$ |

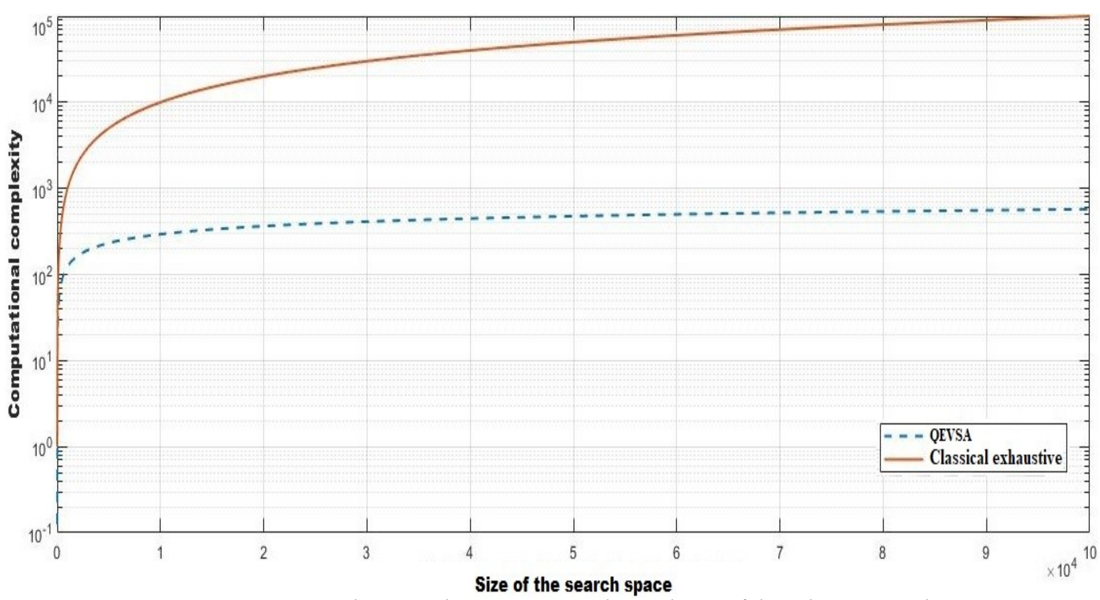


Figure 2: Comparison between the computational complexity of the exhaustive and QEVSA.

## 4.2 Quantum Optimization of Resource Distribution Management in the Multi-Tasks Case

***Thesis Group 2: [J1, C1, C2, C3, C4, C5, C6], I exploited the QEVSA in balancing the load of the RDM model in the multi-tasks case. I build an analytical model that corresponds to the proposed deployment system. Next, I presented and investigated the implementation mechanism of the quantum method. Then, I carefully set up the stochastic parameters of the quantum strategy in accordance with the RDM's requirements. Finally, I validated the theoretical result by extensive simulations. More details are given in Chapter 3.***

**Thesis 2.1**

*In this thesis* ***[J1, C3, C4, C5]****, I developed a general model of RDM with multiple task generators and I defined an appropriate unconstrained metric to measure the efficiency of the quantum method.*

To model the general RDM system, I divided its functionalities into three main blocks. The first one has multiple task generators denoted by *G*, each of them generates tasks that are served by computing units. The memory requirement for each task type is denoted by $m_l$. The computing units may have different theoretical capacities. The total number of computing units is denoted by *K*. The third block is the decision-maker that is responsible for the task deployment. Figure 2 illustrates the proposed RDM model. As an example of application, I have chosen to distribute uniformly the tasks among computing units, so, I used the variance of the relative load (1) as a metric.

$$\sigma^2 = \frac{1}{K} \sum_{k=1}^{K} \left( \bar{b} - \frac{\sum_{l=1}^{G} N_{kl} m_l}{s_k} \right)^2 . \tag{1}$$

where $\bar{b}$ is the average relative load of the system, the parameter $N_{kl}$ refers to the number of task type $l$ in the $k^{th}$ computing unit, and $s_k$ is the theoretical capacity of the $k^{th}$ computing unit. Conserving the load balance among the computing units is an important property in a large resource distribution process. In the case of optimal task distribution, the variance of the relative load tends to zero which means that the tasks are distributed uniformly, else, they are not.

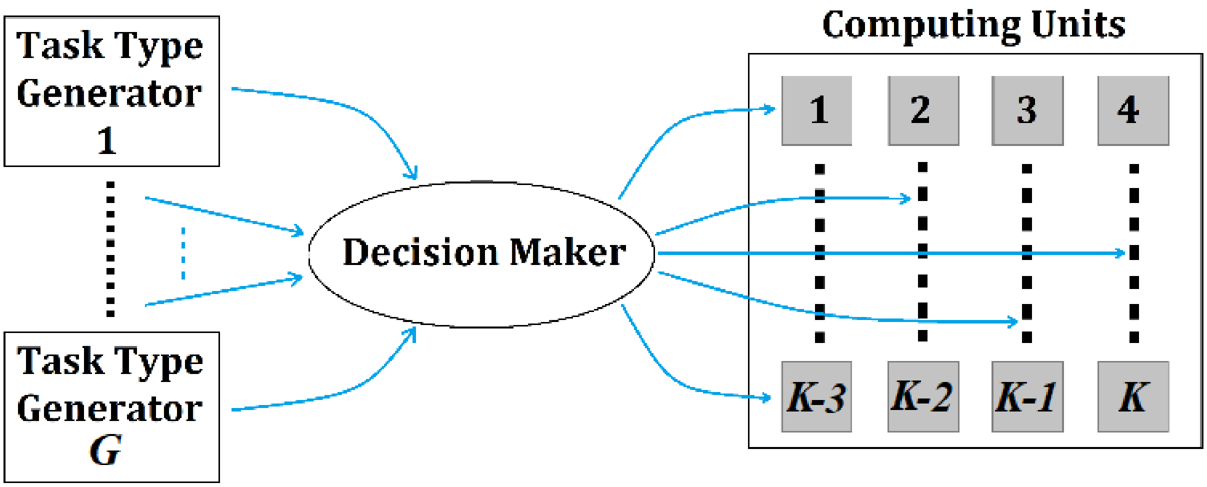


*Figure 3: Resource distribution model architecture.*

**Thesis 2.2**

*In this thesis **[J1, C1, C2, C5],** I showed how can the QEVSA be applied as a computational infrastructure for the RDM. I configured the stochastic parameter T of the quantum algorithm by taking into account the restrictions of the proposed RDM, this quantum optimization strategy can be applied for an unconstrained goal function.*

To distribute uniformly the tasks among computing units, I used the QEVSA as a quantum minimum searching algorithm (QMSA).

The goal is to set up the appropriate value for the maximum number of steps needed to run the logarithm search $T$ (2). This parameter depends on two parameters which are the step size of the search space which is according to the proposed RDM model is the minimum nonzero distance between variances of any two scenarios $\alpha$ as presented in (3) and the size of the region of variance's values $\Delta\sigma^2 = {\sigma^2}_{max} - {\sigma^2}_{min}$,

$$T = \frac{{\sigma^2}_{max} - {\sigma^2}_{min}}{\alpha}, \tag{2}$$

$$\alpha = \min_{\forall\, i,j} \left| ({\sigma^2}_i - {\sigma^2}_j) \right|. \tag{3}$$

The variation's variance that is bounded by 0 and 1, in the unfavorable case, the maximum variance of the relative load will be equal to 0.25. Thus, $\Delta\sigma^2 = 0.25$.

The configuration of $T$ is formulated as finding the minimum distance between variances of two scenarios according to (3). To that end, our key focus is to set up the $\alpha$. Using linear algebraic operations, I showed that the value $\alpha$ is mainly affected by the relative load of the $i^{th}$ and the $j^{th}$ computing units as expressed in (4). To calculate the smallest non-zero difference between variances of two different assignments, it is enough to investigate $\left(\sigma_i^2 - \sigma_j^2\right)^2 = 0$. Furthermore, I showed that the minimum places of the function $\left(\boldsymbol{\sigma_i^2} - \boldsymbol{\sigma_j^2}\right)^{\mathbf{2}}$ are situated in a hyperplane.

$$\sigma_i^2 - \sigma_j^2 = \frac{1}{K}\left[\left(\frac{m_s}{s_i}\right)^2 - \left(\frac{m_s}{s_j}\right)^2 + 2m_s\left(\frac{\sum_{l=1}^{G} N_{il}m_l}{{s_i}^2} - \frac{\sum_{l=1}^{G} N_{jl}m_l}{{s_j}^2}\right)\right]. \quad (4)$$

A simulation example has been built to determine the location of the minimum places of $\alpha$. Figures 3 and 4 show that the solutions of the minimum distance between the two different scenarios are linear solutions, which means that all the appropriate solutions are situated in the middle of the dark blue region, according to our management problem case, the area within the red border is investigated.

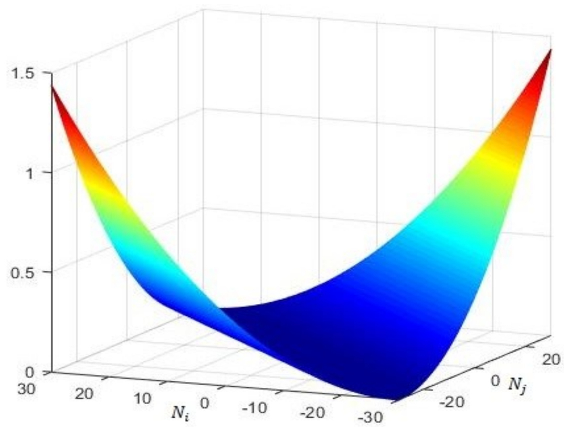


*Figure 3: Shape of the function $\left(\sigma_i^2 - \sigma_j^2\right)^2$.*

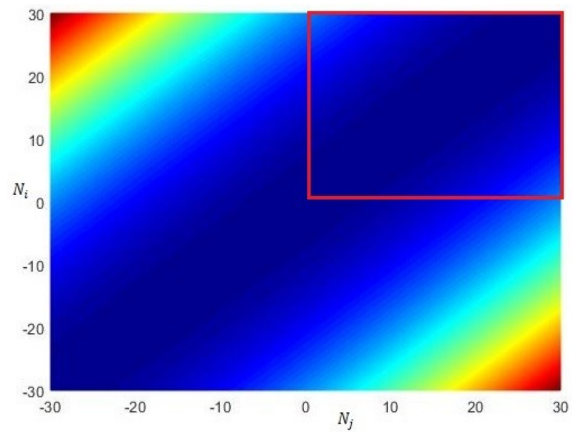


*Figure 4: The Contour of the Shape of the function $\left(\sigma_i^2 - \sigma_j^2\right)^2$.*

**Thesis 2.3**

*In this thesis [C5], I calculated the computational complexity of the QMSA. Next, I showed that the computational complexity of the proposed quantum method can provide a significant reduction in the term of computational complexity and a high level of accuracy.*

The computational complexity of the QMSA is $O\left(log_2(T){log_2}^3(\sqrt{d})\right)$, it depends on the computational complexity of the QET function ${log_2}^3(\sqrt{d})$ and the BSA embedded in the QEVSA $log_2(T)$, where *d* refers to the total number of possible deployment scenarios. The value of *d* according to the proposed RDM equals *K*. On the other hand, setting up the parameter *T* is strongly influenced by the value of $\alpha$.

To compute the optimum $\alpha$, we need to set up the value of $\alpha$ repeatedly for every new incoming task. Therefore, the computational complexity of this operation is $O\left(\frac{K(K-1)}{2}\right)$ which is remarkably a hard computational problem. So, I proposed to compute only one time the global non-zero minimum denoted by $\alpha_{global}$, before starting the operation distribution of the system. Thus, the computational complexity of the entire system will equal $O\left({log_2}^3(\sqrt{d})\right)$.

**Thesis 2.4**

*In this thesis [C3, C5], I demonstrated the gain of the quantum algorithm against the randomized reference, this is done by extensive simulations, where I showed that whatever the intensity type distribution of the incoming tasks, the quantum approach guarantees the best uniformity distribution of the system and low computational complexity.*

I showed that when the intensity distribution of the incoming tasks increases, the randomized strategy diverges completely and fluctuates dramatically, while the optimized strategy keeps and conserves the load uniformity of the system. These results are illustrated in Figure 5.

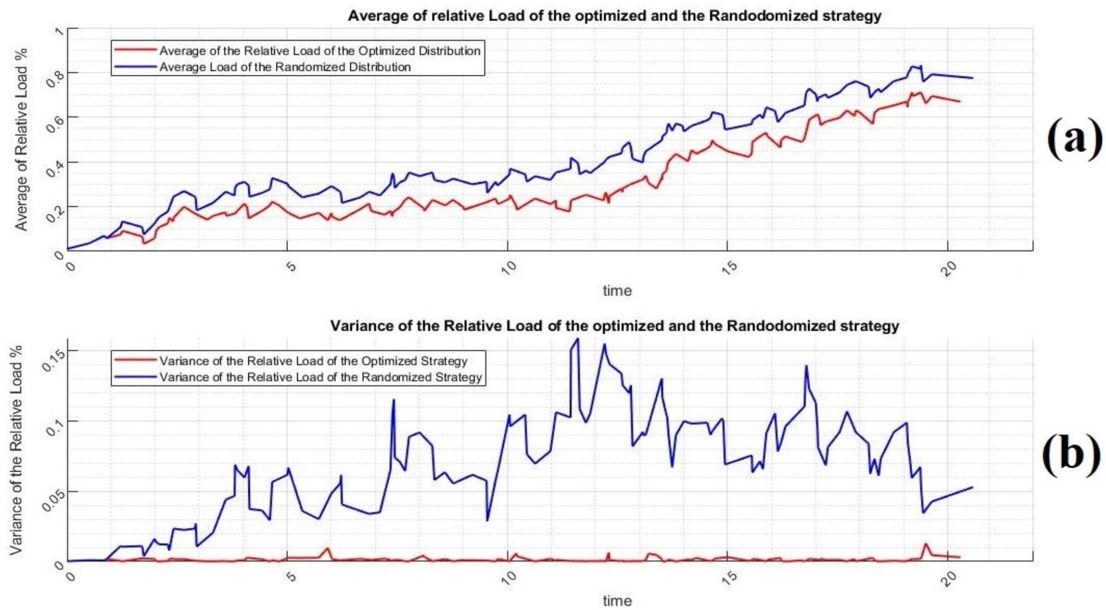


*Figure 5: (a) The average relative load of the optimized strategy (red) and the randomized one (blue). (b) The variance of the relative load of the optimized strategy (red) and the randomized one (blue). In case the mean (intensity) of the exponential distribution of the arrival time of task type 1 is 0.3 and task type 2 is 0.4.*

## 4.3 Quantum Optimization of Resource Distribution Management with Multi-Tasks, Multi-Subtasks

***Thesis group 3: [J2, C6]. I extended the resource assignment model to multi-tasks, multi-subtasks. Next, I implemented QMSA as an embedded computational infrastructure in the decision-maker. Similar to what has been discussed in Thesis group 2, I used the variance of the relative load to test the efficiency of the quantum method. Then, I configured the QMSA according to the new extended model. After, I demonstrated that the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing units but exponential in terms of the number of subtasks. Finally, I validated the obtained result by extensive simulations where I compared the quantum solution with the randomized one. More details are given in Chapter 4.***

**Thesis 3.1**

*In this thesis [J2, C5, C6], I extended the general model of RDM that was proposed in Thesis 2.1, to multi-tasks, multi-subtasks, I used the same unconstrained goal function (the variance of the relative load) to test the efficiency of the quantum method.*

I considered multiple task generators, each of them releases a task type, each task type $l$ is made of several subtasks selected from a set of subtask types, where the total number of different subtask types is $W$. Moreover, every subtask type $v$ has a memory requirement denoted by $\Delta_v$. Let $N_{kv}^{b}$ be the number of task type $v$ under process on the $k^{th}$ computing unit. The expression of the variance of the relative load can be written as

$$\sigma_{V_i}^2 = \frac{1}{K}\sum_{k=1}^{K}\left(\overline{b_{\boldsymbol{V}_i}} - b_k^{\boldsymbol{V}_i}\right)^2. \tag{5}$$

- $\boldsymbol{V_i}$ : The matrix which denotes the possible distribution scenario $i$, its rows represent the computing units and the columns denote the subtask types.
- $\boldsymbol{b_k^{V_i}}$: The relative load of the $k^{th}$ computing unit after receiving subtasks of scenario $\boldsymbol{V_i}$.
- $\overline{\boldsymbol{b_{V_i}}}$: The average relative load of the $k^{th}$ computing unit after receiving subtasks of scenario $\boldsymbol{V_i}$.

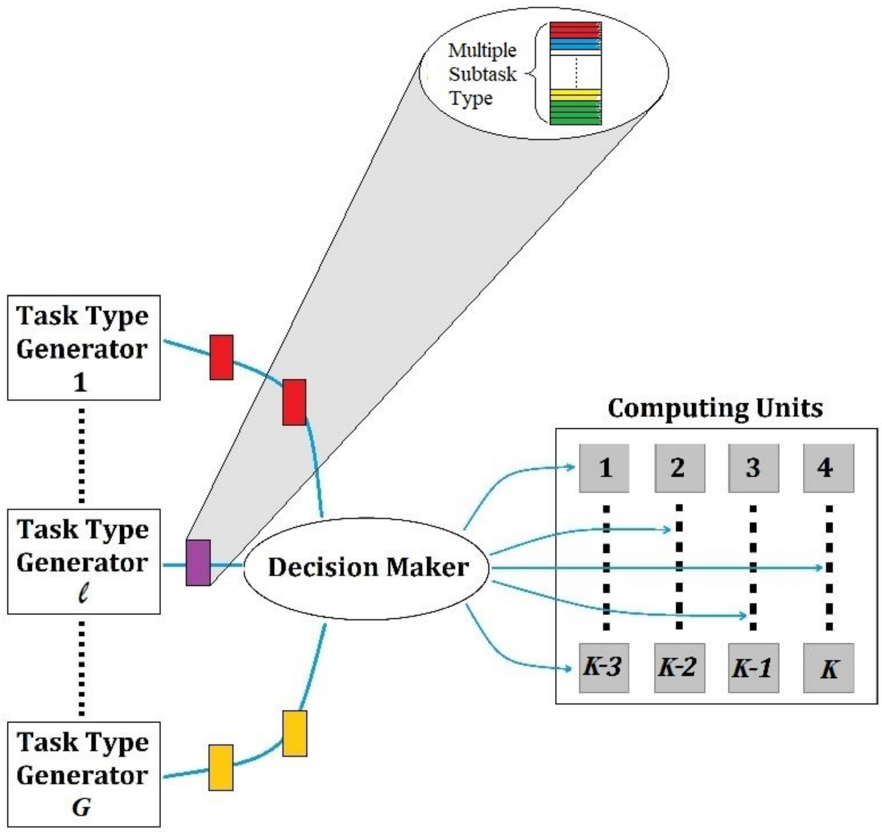

*Figure 6: Resource distribution management model for multi-tasks, multi-subtasks.*

**Thesis 3.2**

*In this thesis [J2, C5, C6], I presented how to set up the stochastic parameter T of the quantum algorithm in the case of the proposed RDM of multi-tasks, multi-subtasks. and I showed that the value of T depends only on the set of computing units $S_{ch}$ that have been assigned a workload, not on all computing units. And I proved that the solutions of T are situated in a hyperplane.*

Similar to what has been discussed in Thesis group 2.2, computing the value of $T$ is traced back to calculate the $\alpha = \left|\sigma_{\boldsymbol{V}_i}^2 - \sigma_{\boldsymbol{V}_j}^2\right|$, where $\boldsymbol{V}_i$ and $\boldsymbol{V}_j$ referred to two different assignment scenarios.

Using linear algebraic operations, I showed that for any distribution, the difference between variances of two different scenarios depends only on the set of computing units $S_{ch}$ that have been assigned a workload, not on all computing units,

$$\sigma^2_{\boldsymbol{V}_i} - \sigma^2_{\boldsymbol{V}_j} = \frac{1}{K}\left[\sum_{k\in S_{ch}^{V^i}}\left(\frac{\sum_v^W N_{kv}^{V^i}}{s_k}\right)^2 - \sum_{k\in S_{ch}^{V^j}}\left(\frac{\sum_v^W N_{kv}^{V^j}}{s_k}\right)^2 \right.$$
$$\left. + 2\left(\sum_{k\in S_{ch}^{V^i}}\left(\sum_{v=1}^{W} N_{kv}^{b}\right)\left(\frac{\sum_v^W N_{kv}^{V^i}}{{s_k}^2}\right) - \sum_{k\in S_{ch}^{V^j}}\left(\sum_{v=1}^{W} N_{kv}^{b}\right)\left(\frac{\sum_v^W N_{kv}^{V^j}}{{s_k}^2}\right)\right)\right]. \tag{6}$$

- $N_{kv}^{V^i}$ and $N_{kv}^{V^j}$ are respectively the number of subtask type $v$ in the k^th computing unit after deploying the new incoming task according to scenarios $\boldsymbol{V}_i$ and $\boldsymbol{V}_j$.
- $N_{kv}^{b}$ is the number of subtask type $v$ in the *k^th* computing unit before deploying the new incoming task.
- $S_{ch}^{V^i}$ and $S_{ch}^{V^j}$ are respectively the set of computing units receiving the new subtasks of scenario $\boldsymbol{V}_i$ and $\boldsymbol{V}_j$.

On the other hand, I showed that to calculate the minimum points of $\alpha$, it is enough to investigate $\left(\sigma^2_{\boldsymbol{V}_i} - \sigma^2_{\boldsymbol{V}_j}\right)^2 = 0$, this expression derives an important property of the minimum points. In compliance with this result, I concluded that the minimum places of the function $\left(\sigma^2_{\boldsymbol{V}_i} - \sigma^2_{\boldsymbol{V}_j}\right)^2$ are situated in a hyperplane.

**Thesis 3.3**

*In this thesis [J2], I calculated the computational complexity of the quantum algorithm in case of considering the proposed RDM model and I proved that the proposed optimization assignment problem cannot be solved classically, because the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing units but exponential in terms of the number of subtasks.*

The computational complexity of the entire system is $O\left(log_2(T){log_2}^3\left(\sqrt{d}\right)\right)$, similarly to what has been discussed in thesis 2.3, we used an alternative solution to set up the value of *T*. Therefore, the computational complexity of the entire system will be $O\left({log_2}^3\left(\sqrt{d}\right)\right)$, The size of the search space *d* (the total number of possible deployment scenarios) can be expressed as

$$d = \prod_{v=1}^{W}\binom{K + M_{gv} - 1}{M_{gv}}. \tag{7}$$

Next, I estimated the lower bound of the search space *d*,

$$d = \prod_{v=1}^{W} \binom{K + M_{gv} - 1}{M_{gv}} \geq \frac{K^{M_g}}{\left(\max_v M_{gv}!\right)^W} = \left(\frac{1}{\max_v M_{gv}!}\right)^W K^{M_g}, \tag{8}$$

where $M_{gv}$ subtasks type $v$ released from generator $g$, and $M_g = \sum_{v=1}^{W} M_{gv}$.

It is noticeable that the computational complexity of the lower bound of the size search space $d$ is polynomial in terms of the numbers of computing units $K$ but exponential in terms of the number of subtasks $M_{gv}$. Therefore, performing a classical computation method to find the optimum result will be time-consuming and hard to solve. Thus, the proposed quantum strategy guarantees an accurate and fast computation of the optimum assignment scenarios.

**Thesis 3.4**

*I validated the efficiency of the quantum method against the randomized one in the case of considering the proposed RDM system (multi-tasks, multi-subtasks). I showed that whatever the intensity type distribution of the incoming tasks, the quantum strategy guarantees a load balance and a low computational complexity [J2].*

I showed that although the changes of the average relative load, the variance of the relative load of the optimized strategy tends to zero while the variance of the relative load in the randomized method fluctuates and diverges from zero. Figure 7 illustrates the obtained results.

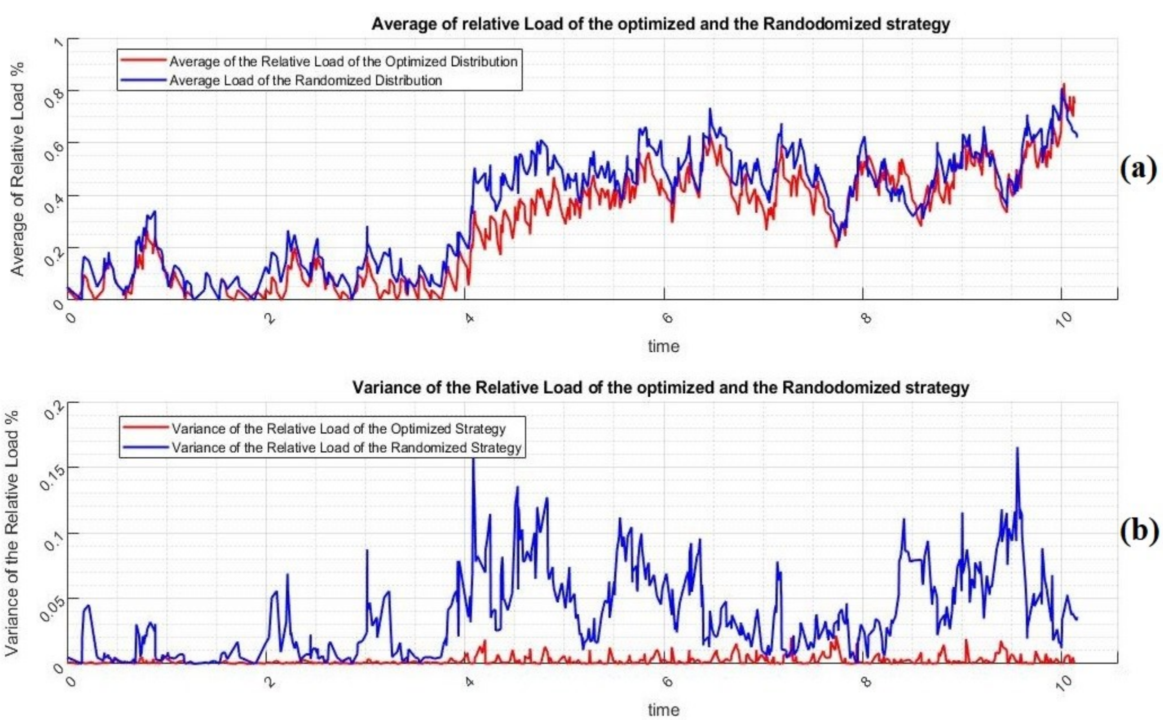


*Figure 7: The average relative load of the optimized strategy (red) and the randomized one (blue). (b) The variance of the relative load of the optimized strategy (red) and the randomized one (blue). In case the mean (intensity) of the exponential distribution of the arrival time of task type 1 is 0.1 and task type 2 is 0.09.*

## 4.4 Constrained Quantum Optimization of Resource Distribution Management with Multi-Tasks, Multi-Subtasks

***Thesis Group 4: [J3]. I exploited the CQOA introduced in Thesis 1.2 in reducing the energy consumption of computing units by taking into consideration the delay constraint of the tasks. Next, I showed that this constrained assignment optimization problem can only be solved by a constrained quantum strategy. Also, I proved that the CQOA provides a significant reduction in terms of computational complexity and energy consumption. More details are given in Chapter 5.***

**Thesis 4.1**

*In this thesis, I extended the functionalities of the RDM model. Next, I proposed to optimize the overall energy consumption of the system by taking into account the delay constraints of tasks as an example of application* ***[J3].***

I defined an extended RDM model for multi-tasks, multi-subtasks. Each generated task type has to obey a specific delay constraint denoted by $\tau_g$. Next, I assume that the subtasks of the incoming tasks are processed sequentially. The delay needed to process the actual load on the $k^{th}$ computing unit is denoted by $\tau_k^{act}$, it has to be always less or equal than the delay constraint of the incoming task type, i.e. $\tau_k^{act} \leq \tau_g$. Also, I assumed that the initial power needed to turn on the computing unit is denoted by $P_k^{init}$. It is important to mention that a subtask is composed of a number of identical pieces called memory units. The processing rate of the $k^{th}$ computing unit is denoted by $\beta_k$ and measured as follows,

$$\beta_k = \frac{number\ of\ memory\ units}{second}. \tag{9}$$

The processing delay of the $k^{th}$ computing unit can be calculated as,

$$\tau_k^{act} = \frac{\sum_{v=1}^{W} N_{kv}\Delta_v}{\beta_k}. \tag{10}$$

Assuming that $\tau_k^{act} \leq \tau_g$ holds, and the energy consumption of one memory unit of the computing unit $k$ is denoted by $\varepsilon_k$. One can write the overall energy consumption of the system as

$$E^{act} = \sum_{k=1}^{K} E_k^{act} = \sum_{k=1}^{K} \left( \varepsilon_k + \frac{P_k^{init}}{\beta_k} \bigg|_{k \in H^{on}} \right) \sum_{v=1}^{W} N_{kv}\Delta_v. \tag{11}$$

**Thesis 4.2**

*In this thesis, I showed how to apply the CQOA in the new model, then, I derived a simplified form of the constrained goal function (Overall energy consumption of the RDM). Next, I configured the implemented constrained optimization method according to the proposed RDM model* ***[J3].***

Similar to what has been discussed previously, to minimize the overall constrained energy consumption of the system, the CQOA is implemented as a minimum constrained searching algorithm (MCSA). Then, I showed that the total constrained energy consumption of the system is influenced by the energy consumption of the incoming task with taking into consideration that it was distributed according to the scenario $Z_i$,

$$E^{inc}(Z_i) = \sum_{k\in Z_i}\left(\varepsilon_k + \frac{P_k^{init}}{\beta_k}\bigg|_{k\in H^{on}}\right)\sum_{v=1}^{W} M_{kvi}\Delta_v, \tag{12}$$

- $M_{kvi}$ denotes the number of subtasks of the incoming task from subtask type *v* deployed onto computing unit *k* in case of the $i^{th}$ deployment scenario.
- $Z_i$ refers to the $i^{th}$ set of those computing units which received one or more subtasks of the incoming task.
- $H^{on}$ refers to the set of the switched-on computing units.

As we already discussed, the value of *T* is written as follows,

$$T = \frac{E_{max}^{act} - E_{min}^{act}}{\alpha}. \tag{13}$$

where $E_{max}^{act}$ and $E_{min}^{act}$ are respectively the maximum and minimum constrained energy consumption. Using (12), one may verify that $\alpha$ can be expressed as

$$\alpha = \min_{\forall\, i\neq j}\left|E^{inc}(Z_i) - E^{inc}(Z_j)\right|. \tag{14}$$

To set up properly the stochastic parameter $\alpha$, it is enough to investigate the non-zero solutions of $\left(E^{inc}(Z_i) - E^{inc}(Z_j)\right)^2 = 0$, the solutions are located on a hyperplane, this result was already discussed in the previous Chapters 3 and 4.

**Thesis 4.3**

*In this thesis, I proved that the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing units K but exponential in terms of $M_{gv}$. Therefore, a classical computational infrastructure cannot solve this type of assignment optimization problem. But, the proposed CQOA can handle this constrained optimization problem exponentially faster* ***[J3].***

Similarly as has been discussed in Thesis 3.3, the computational complexity depends on the BSA $O\left(log_2(T)\right)$ and the CQRT function $log_2{}^3\left(\sqrt{d}\right)$. I exploited the alternative proposed solution for avoiding calculating repeatedly the stochastic parameter *T*. I also showed that the computational complexity of the entire system is $log_2{}^3\left(\sqrt{d}\right)$. Furthermore, I computed the lower bound of all the possible deployment scenarios, and I showed that the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing unit *K* but exponential in terms of $M_{gv}$. Therefore, this constrained problem assignment cannot be solved using the classical existing computing machines.

**Thesis 4.4**

*Using a simulation environment, I proved the effectiveness of the constrained quantum optimization method, by making a comparison with the constrained randomized algorithm. Next, I showed that whatever the distribution intensity is, the constrained optimized approach offers the best performance and can significantly reduce the energy consumption of the computing units* ***[J3].***

I built a simulation environment for the proposed RDM model and I showed that whatever the distribution intensity is, the constrained optimized approach offers a significant reduction in terms of energy consumption and computational complexity. Figures 8 and 9 illustrate the obtained results.

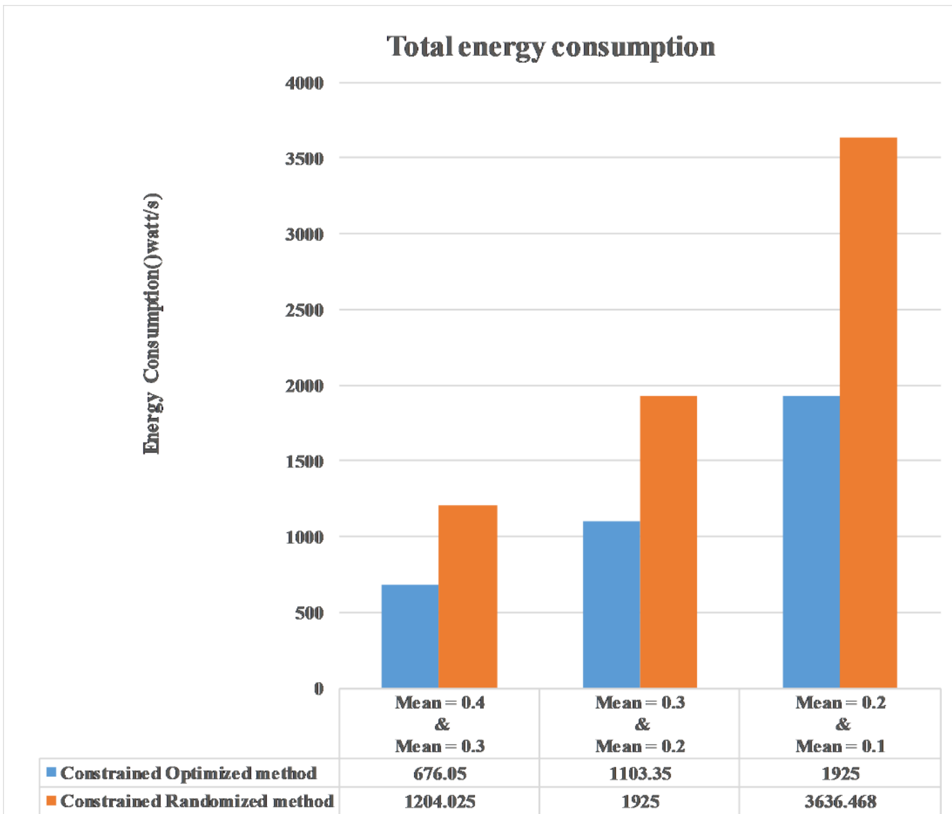


*Figure 8: Energy consumption of the Optimized (Blue bars) and the randomized (Red bars) strategies according to different intensity distributions.*

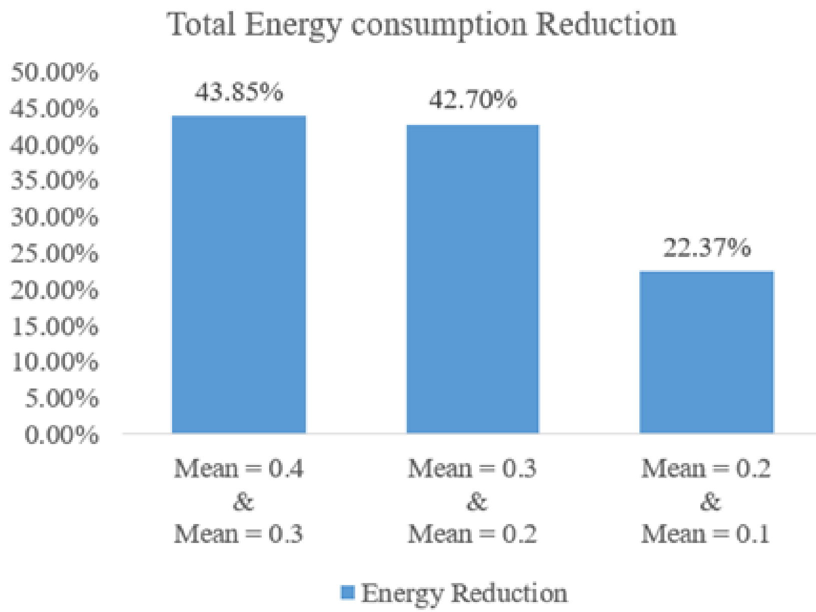


*Figure 9: Energy consumption reduction.*

## 4.5 Constrained Quantum Optimization of Resource Distribution Management with Multi-Tasks, Multi-Subtasks with queueing behavior

***Thesis Group 5: [J4, C7, C8]. I exploited the CQOA introduced in Thesis 1.2 in reducing the energy consumption of computing units by taking into consideration the delay constraint of the tasks as well as considering the queueing scenarios. More details are given in Chapter 6.***

Realistic task deployment systems may face queueing scenarios where jobs might be sent to the waiting queue, and if the intensity of incoming tasks is larger, then the number of tasks in the waiting queue could increase dramatically. As it is known, the length of the waiting queue and the quality of service are two aspects that are linked. For this reason, I focused on optimizing and improving the queueing system process. More details are presented in Chapter 6.

**Thesis 5.1**

*In this thesis, I presented an extended version of the RDM model, and considering queueing scenarios, I chose to share single free access waiting queue among computing units. Next, I defined the rules of task deployment according to capacity and delay constraints* ***[J4].***

We consider the previously described resource distribution model. I proposed to consider the actual processing delay plus the waiting delay as one constrained optimization approach. Also, I assumed that the system contains only one free access waiting queue shared among computing units, where it has no energy consumption. Let $Q$ denote the length of the queue in the memory unit and $q$ the actual number of tasks in the waiting queue including the new task. Figure 10 illustrates the new model. On the other hand, I supposed that,

- If there is free space in the computing units and the delay constraint of the incoming task is met, the subtasks derived from the incoming task are distributed to the optimum places.
- If there is free space in computing units and the delay constraint of the incoming task is not satisfied, the arrived task is rejected.
- If there is no free space in the computing unit, the incoming job is sent back to the free access waiting queue.

I assumed that the total number of generated jobs from different generators is *D*. They are organized according to their arrival time, from 1 to *D*. Let us assume that the $i^{th}$ incoming task released from the generator $g$ requires processing its subtasks within a delay constraint $\tau_g$. If this delay constraint is not met, then the task is rejected. Furthermore, if there is no free capacity space for the incoming task in the system, the new task is sent back to the waiting queue. If there are tasks in the queue, the selection of the next task is based on the one that has the highest priority to be deployed. I proposed giving priority to the tasks which have the largest delay. This delay is the sum of the waiting time of the task in the queue and the time required to serve this task.

I changed the notation of $\tau_k^{act}$ defined in (10) to $\tau_{i,k}^{act}$ which refers to the processing time of the $k^{th}$ computing unit when the subtasks of the $i^{th}$ incoming task is deployed into the system. The waiting delay of the $i^{th}$ incoming task is denoted by $\theta_i$. If the $i^{th}$ task of generator $g$ is coming from the waiting queue and there is free capacity space in the system, then, it is necessary to take into consideration the queueing time,

$$\forall\, k \qquad \theta_i + \tau_{i,k}^{act} \leq \tau_g. \tag{15}$$

On the other hand, if the new $i^{th}$ incoming task released from generator $g$, the value of $\theta_i$ equals zero.

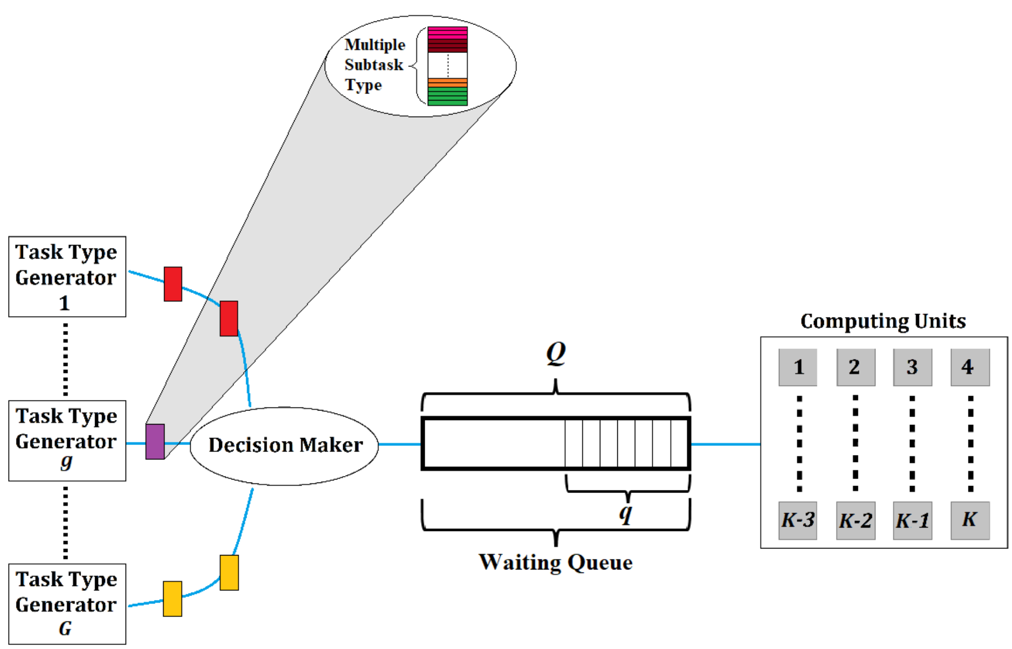


*Figure 10: Resource distribution management system with a shared queue.*

**Thesis 5.2**

*In this thesis, I implemented and configured the CQOA according to the queueing aspect of the RDM model (multi-tasks, multi-subtasks)* ***[J4, C7, C8].***

Similarly to what has been done in Theses 4.2, 3.2, and 2.2. I used the same approach for the implementation and configuration of the CQOA. The only difference remains in considering the new shared free access waiting queue between computing units.

**Thesis 5.3**

*In this thesis, I proved that the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing units K and the number of tasks in the waiting queue q but exponential in terms of* $M_{gv}$. ***[J4].***

The computational complexity of selecting the optimum deployment scenario for the constrained randomized method is $O(d')$ or $O(1)$ where $d'$ is the total number of possible deployment scenarios in case using the constrained randomized strategy, while the computational complexity of selecting the task from the waiting queue according to the priority deployment is $O(q)$. I showed that $d'$ can be written as

$$d' = \prod_{v=1}^{W} \binom{K + M_{gv} - 1}{M_{gv}} \geq \left(\frac{K^{M_{gv}}}{\max\limits_{v} M_{gv}\,!}\right)^{W} \tag{16}$$

On the other hand, the computational complexity of selecting the best optimum deployment among computing units and the selection of the most prior task from the queue of the constrained quantum solution requires only $O\left(log_2(T) log_2{}^3(\sqrt{d})\right)$.

In contrast, the computational complexity of selecting the best optimum deployment scenario among computing units and the selection of the most prior task from the queue using the CQOA strategy requires only $O\left(log_2(T) log_2{}^3(\sqrt{d})\right)$. It is important to note that the configuration of the parameter $T$ is similar to what has been discussed previously. The computational complexity of the CQRT is $O\left(log_2{}^3(\sqrt{d})\right)$, one can easily verify that the value of $d$ can be expressed as

$$d = q!\left(\prod_{v=1}^{W} \binom{K + M_{gv} - 1}{M_{gv}}\right)^{q} \tag{17}$$

Next, I showed that the estimated lower bound of $d$ can be written as

$$q!\left(\prod_{v=1}^{W} \binom{K + M_{gv} - 1}{M_{gv}}\right)^{q} \geq q\left(\frac{K^{M_{gv}}}{\max\limits_{v} M_{gv}\,!}\right)^{W.q} \tag{18}$$

One can observe that the computational complexity of finding the optimal solution within the database is polynomial in terms of the numbers of computing units $K$ and the number of tasks in the waiting queue $q$ but exponential in terms of $M_{gv}$. If the number of subtask type $W$ or the number of tasks in the waiting queue $q$ increase, the computational complexity will augment significantly. Therefore, the computational complexity of the constrained quantum solution requires significantly fewer steps than the constrained randomized classical one, additionally, applying the constrained randomized classical solution will be time-consuming and hard to solve.

**Thesis 5.4**

*I confirmed by a simulation experiment the efficiency of the constrained quantum optimization method in case of considering rejected tasks or queued tasks. I proved that the proposed constrained quantum method guarantees a reduction in terms of energy consumption and computational complexity* ***[J4].***

I built two simulation environments:

- First experiment: The new incoming task is rejected if and only if one of the constraints (capacity constraint or delay constraint) is not met.
- Second experiment: The task is rejected when the delay constraint is not met and it is sent to the waiting queue when there is no free space in the computing units.

*First experiment:*

The outcome of this experiment showed that:

- The total energy consumed by the constrained optimized strategy is lower than that of the energy consumed by the constrained randomized method.
- The number of rejected tasks of the constrained optimized strategy is approximately equal to the one of the constrained randomized case.
- The computational complexity of the constrained randomized method is $O(1)$ or $O(const)$, while in the worst-case, it is $O(d)$. In contrast, the computational complexity of the CQOA is $O\left(log_2(T)log_2{}^3\left(\sqrt{d}\right)\right)$.

*Second experiment:*

The outcome of this experiment showed that:

- The constrained randomized algorithm and the constrained optimized method consume approximately the same quantity of energy.
- The total number of rejected tasks at the end of the simulation of the constrained randomized strategy is slightly higher than the constrained optimized one
- The computational complexity of the constrained randomized is $O(d'q)$, while the computational complexity of CQOA is $O\left(log_2(T)log_2{}^3\left(\sqrt{d}\right)\right)$.

## 5. Applications and Future work

In future work, we would expand the CQOA's functionalities to find the optimum solutions for a constrained multivariable objective function. Following that, the RDM paradigm would be improved with a more challenging task assignment problem in which each computing unit is linked to a waiting queue. Hence, in the latest RDM update, the current expanded version of the CQOA will be used.

Machine-learning, network telecommunications, physical structure simulations, and other problems are concerned with minimizing the constrained multivariate objective function. The contribution of quantum optimization algorithms to these problems may lead to ensuring a high exponential speed and a short execution time.

## Own Publications

### Journal Papers


[J1]: S. El Gaily. EVALUATION OF RESOURCE OPTIMIZATION BASED ON QUANTUM SEARCH. Hungarian Journal of industry and chemistry. 2019. DOI: 10.33927/hjic-2019-03.

[J2]: S. El Gaily, S. Imre, "Quantum Optimization of Resource Distribution Management for Multi-Task, Multi-Subtasks", Infocommunications Journal, Vol. XI, No 4, December 2019, pp. 47-53.

[J3]: S. El Gaily, S. Imre, " Constrained Quantum Optimization for Resource Distribution Management, *Int. J. Adv. Comput. Sci. Appl*, 2021.

[J4]: S. El Gaily, S. Imre, " Implementation of a Constrained Quantum Optimization Method in Resource Distribution Management with Considering Queueing Scenarios ", *Int. J. Communication Networks and Distributed Systems*, 2021.


### Conference Papers


[C1] S. El Gaily:" RESOURCE OPTIMIZATION BASED ON QUANTUM SEARCHING", 15th International Ph.D. Workshop on Systems and Control (2018) pp. 1-3. Hungary.

[C2] S. El Gaily: "Performance evaluation of quantum-based resource management" The International Conference on Quantum Computing (ICoCQ2019), (2018) pp. 1-5. France

[C3] S. El Gaily, S. Imre: Quantum Optimization in Large Resource Management Systems - IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2-5 July 2019, Cannes, France.

[C4] S. El Gaily, S. Imre: Derivation of Parameters of Quantum optimization in Resource Distribution Management - 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), 1-3 July 2019, Budapest, Hungary.

[C5] S. El Gaily, S. Imre: "Quantum Resource Distribution Management in Multi-task environment", 14th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, Serbia.

[C6] S. El Gaily, S. Imre: “Quantum Computing-based Onboard Hardware Management for Spacecrafts ” ITU WORLD 2019. Conference on space communication, space activities and risks of the IT sector for V4 countries”. Budapest, Hungary

[C7] S. El Gaily, Sándor Imre: “Extending the functionalities of the quantum extreme value searching algorithm to a constrained quantum searching algorithm”, The Sixth International Conference for Young Quantum Information Scientists (YQIS 6 or YQIS 2021). 12-16 April 2021, USA.

[C8] S. El Gaily, S. Imre: “Constrained Quantum Optimization Strategy”, The Bristol Quantum Information Technologies Workshop. 26-28 April. England.

[C9] S. El Gaily, S. Imre: “Constrained Quantum Optimization Algorithm”, 20th International Symposium INFOTECH-JAHORINA, 17-19 March 2021, Jahorina, East Sarajevo.

[C10] S. El Gaily, Sándor Imre: “Quantum Computing based Efficient Optimization”, CTP Quantum Information Days, 22-24 February 2021, Poland.

# Bibliography

[1] S. Imre, F. Balázs: Quantum Computing and Communications – An Engineering Approach, Published by John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2005, ISBN 0-470-86902-X, 283 pages. DOI: 10.1002/9780470869048.

[2] M. A. Nielsen and I. L. Chuang. Quantum computation and quantum information. Cambridge University Press. 2000.

[3] Tamene, M., and Nageswara, K. (2018) Intelligent clustering protocol design and optimisation in wireless sensor networks, *Int. J. Communication Networks and Distributed Systems*, Vol. 20, No. 2.

[4] Azad, P., Navimipour N J., and Hosseinzadeh, M. (2019) A fuzzy-based method for task scheduling in the cloud environments using inverted ant colony optimisation algorithm, *Int. J. Bio-Inspired Computation*, Vol. 14, No. 2.

[5] Balusamy, B., Karthikeyan, K., and Sangaiah A K. (2017) Ant colony-based load balancing and fault recovery for cloud computing environment, *Int. J. Advanced Intelligence Paradigms*, Vol. 9. No. 2/3

[6] Tamba, T A., and Joelianto, E. (2019) Network utility maximisation through particle swarm optimisation, *Int. J. Systems, Control and Communications*, Vol. 10, No. 1.

[7] Thaman, J., and Singh, M. (2017) Cost-effective task scheduling using hybrid approach in cloud, *Int. J.Grid and Utility Computing*, Vol. 8.

[8] Meenakshi Priya, A., and Kannigy Devi, R. (2019) Multi-objective optimisation techniques for virtual machine migration-based load balancing in cloud data centre, *Int. J. Cloud Computing*, Vol. 8. No. 3

[9] Rajguru, A A., and Apte, S S. (2018) QoS enhanced distributed load balancing and task scheduling framework for wireless networks using hybrid optimisation algorithm, *Int. J. Communication Networks and Distributed Systems*, Vol. 21, No. 2.

[10] Swagatika, S., and Rath, A K. (2019) SLA-aware task allocation with resource optimisation on cloud environment, *Int. J. Communication Networks and Distributed Systems*, Vol. 22, No. 2.

[11] Hima Bindu G.B., Ramani, K., and Shoba Bindu, C. (2018) Energy aware multi objective genetic algorithm for task scheduling in cloud computing, *Int. J. Internet Protocol Technology*, Vol. 11, No. 4.

[12] Moza, M., and Kumar, S. (2018) Routing in networks using genetic algorithm, *Int. J. Communication Networks and Distributed Systems*, Vol. 20, No. 3.

[13] Jocksam G De Matos; Carla K., Marques D M., and Liberalino, C H.P. (2019) Genetic and static algorithm for task scheduling in cloud computing, *Int. J. Cloud Computing*, Vol. 8.

[14] Kousalya, A., and Radhakrish, R. (2017) Hybrid algorithm based on genetic algorithm and PSO for task scheduling in cloud computing environment, *Int. J. Networking and Virtual Organisations*, Vol. 17.

[15] Imre, S. (2007) Quantum Existence Testing and Its Application for Finding Extreme Values in Unsorted Databases IEEE Transactions on Computers.

[16] Imre, S. and Balázs, F. (2015) Quantum computing and communications: An engineering approach. Wiley, Chichester, UK.

[17] Knuth, D E. (1998) Sorting and searching in the art of computer programming (Addison-Wesley, Reading, MA, USA) 3 (2nd ed)