Multi-Objective Scheduling In Distributed Systems Environment: A Fuzzy TOPSIS Approach

Aidan Ghojoghi
Department of Computer Engineering, Sari Branch, Islamic Azad University, Sari, Iran

aaa.ghojoghi@yahoo.com

***Abstract*—An efficient resource management is the first class concern in distributed systems. As such, scheduler as a main part of resource management framework face multi criteria decision making problem to meet variety criteria having potentially conflicts regard to both service user and service provider viewpoints. In this paper, the fuzzy TOPSIS method is applied to schedule user requests specified in terms of execution time and deadline to be met objectives such as minimizing tardy tasks, makespan in favor of users perspectives and maximizing system throughput and cumulative completion time in favor of service providers perspectives. Before launching the main work by TOPSIS module, the AHP method adopted for data mining on servers data history to weight criteria as their importance. The remarkable result show that the fuzzy TOPSIS module can efficiently opt the best alternative and can make trade-off amongst objectives according to their weights.**

Keywords—Fuzzy TOPSIS, Cloud Computing, Multi-objective Scheduling

I. INTRODUCTION

Distributed systems include variety heterogeneous computing resources which are interconnected via computer networks and seems to be a single node for its users [1-2]. Nowadays, human long-held dreams, utility computing, have been accomplished such as grid and cloud computing with per-per-use basis [3]. Distributed systems infrastructure is embedded on datacenter with intricate architecture [4]. So, one of the most importance concern is efficient computing resource management which needs smart scheduling to control potential conflicts between users and providers viewpoints. For instance, users of google services need quick response time therefore google broker can conduct users' request to prolific cluster servers. Consequently, it incurs power consumption, maintenance costs and total cost of ownership (TCO) as well. On the other hand, broker can convey user workload to consolidated servers in virtualized environment to minimize the number of physical machines in use and maximize resource utilization for cost management. So, this procedure may jeopardize user quality of service (QOS). However, in the clarified complex environment with conflicting perspectives we face with multi criteria decision making (MCDM) problem. The goal of this paper is to schedule user tasks on datacenter high performance computing (HPC) systems to reach several objectives which potentially have conflicts. Several approaches have been presented in literature to figure out MCDM problems, but none of them considers exact weight to criteria to make exact decision. So, the presented works hold a degree of uncertainty. I. Chamodrakas and D. Martakos have developed a utility-based fuzzy TOPSIS method to for energy efficient network selection in wireless sensor network (WSN) environment [5]. The criteria considered were user performance, network conditions, QOS and energy requirement. Their results showed the balance between performance and energy consumption. A fuzzy TOPSIS has been used for group decision making in oil industry to select the best combat responses to oil spills [6]. It has applied Fuzzy TOPSIS to rank between alternatives with different metric weights although it suffered from a degree of uncertainty. To fill the gap, we apply Fuzzy TOPSIS method along with leveraging analytical hierarchical process (AHP) technique which exploit pair comparison between criteria [11]. As such, the near exact weight are assigned to criteria before main fuzzy TOPSIS process launches. The reminder of this paper organized as follows. Section two is dedicated to fuzzy TOPSIS specification. Problem statement is brought in section three. Section four and five present work evaluation and conclusion respectively.

II. FUZZY TOPSIS- Technique for Order Performance by Similarity and Ideal Solution

TOPSIS is broadly applied in literature to handle multi objectives decision making problems in real world. Although cab be easily deployed, this technique is often criticized for the sake of its vagueness and uncertainty in decision process resulting of subjective human comparison. Therefore, fuzzy TOPIS has been developed by hwang et al. to obviate its uncertainty [7]. Moreover, TOPSIS has been improved to deal MCDM with an uncertain decision matrix resulting in fuzzy TOPSIS, which has successfully been utilized to figure out various MCDM problems [8-10]. However, we apply AHP method to determine near to exact weight for criteria (metrics) in the inception of the work [11]. The alternatives are placed as points within a n-dimensional Euclidean space with each dimension corresponding to each criterion and their ranking is produced according to their closeness to the ideal and farness to the anti- ideal points which are modeled as hypothetical alternatives that have respectively the best and the worst utility values for each criterion [5-7]. The TOPSIS method determines the metric of “relative closeness” which is a function of the Euclidean distances of each alternative from the ideal (A^{+}) and the anti-ideal points (A^{-}) in order to represent the simultaneous satisfaction of two objectives: the best alternative should be the closest to the ideal point and the farthest from the anti-ideal point as Fig. 1 illustrates. Also, the relative closeness measure is expressed as $C_i^+ = \frac{S_i^-}{S_i^- + S_i^+}$ where S_i^- and S_i^+ are the distances of alternative i from the anti-ideal and the ideal point, respectively (Fig. 1).

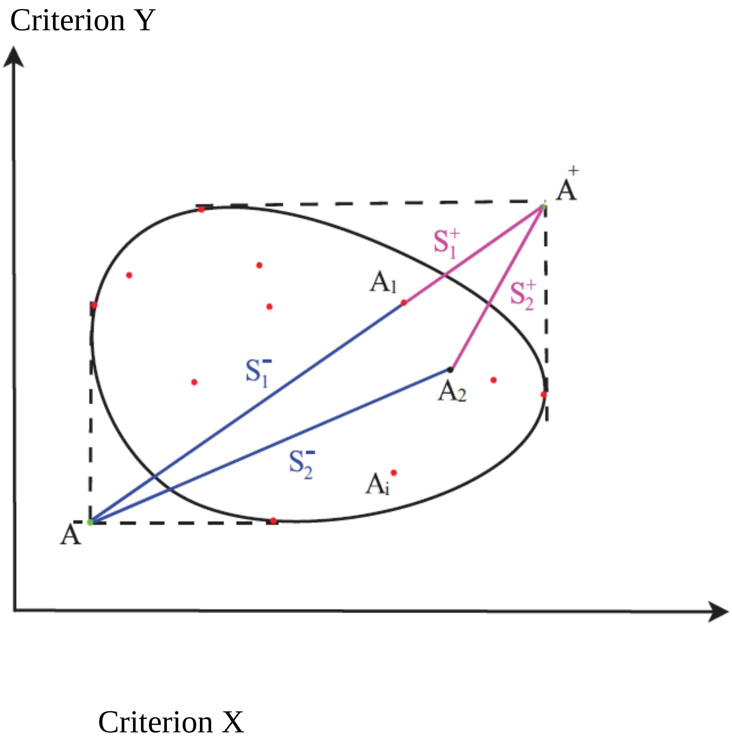

Fig. 1. Alternatives as two points in 2D space corresponding to criteria x and y and their distance from ideal and anti-ideal points

The main method is elaborated as below:
Let us consider the decision matrix A which includes alternatives and metrics (criteria) as following:

$$A=\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \qquad (1)$$

Whereas $A_1, A_2,\ldots, A_m$ are alternatives, and C_1, C_2 , …, C_n are criteria, x_{ij} indicates rate of alternative A_i regarding to criteria C_j. The weight vector W= $(w_1, w_2, \ldots, w_n)$ composed weights w_k (k=1,…,n) for indication the importance of each criteria C_k subject to $\sum_{k=1}^{n} w_k = 1$. Moreover, criteria are divided to benefit and cost types which the first type the higher value is desirable and for the second one lower value is eligible as opposed to the first one. As the data of the decision matrix comes from different sources, it needs to become dimensionless with normalization approach which permits the comparison with various criteria. Then normalized decision matrix R= $[r_{ij}]_{m\,x\,n}$ with i=1,…,m and j=1,…,n is calculated. The normalized value r_{ij} is calculated by following equation:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \qquad (2)$$

Matrix R shows the relative rating of alternatives. Then weighted normalized decision matrix P= $[p_{ij}]_{m\,x\,n}$ with i=1,…,m and j=1,…,n is developed. It is obtained by multiplying the normalized decision matrix and its related weights namely $P_{m\,x\,n} = R_{m\,x\,n}.W_{n\,x\,n}$. Where $W_{n\,x\,n}$ is a diagonal matrix with weights placed on the main diagonal. After that, TOPSIS method is started with four steps as below:

Step 1: To identify positive ideal solution A^+ as benefit and negative ideal solution A^- as cost:

$$A^+ = (p_1^+, p_2^+, \dots, p_m^+) \qquad (3)$$

$$A^- = (p_1^-, p_2^-, \dots, p_m^-) \qquad (4)$$

Whereas $p_j^+ = (max_i\, p_{ij},\ j \in J_1\ ;\ min_i\, p_{ij},\ j \in J_2\ )$ and $p_j^- = (min_i\, p_{ij},\ j \in J_1\ ;\ max_i\, p_{ij},\ j \in J_2\ )$ where J_1 and J_2 are benefit and cost types criteria respectively.

Step 2: To calculate Euclidean distances from positive ideal solution A^+ and the negative ideal solution A^- for each alternative A_i as follows:

$$S_i^+ = \sqrt{\sum_{j=1}^{n} (p_j^+ - p_{ij})^2} \qquad \text{with } i=1,\dots,m$$

(5)

$$S_i^- = \sqrt{\sum_{j=1}^{n} (p_j^- - p_{ij})^2} \qquad \text{with } i=1,\dots,m$$

(6)

Step 3: To calculate the relative closeness C_i for each alternative A_i with respect to positive ideal solution as follow:

$$C_i = \frac{S_i^-}{S_i^+ + S_i^-} \qquad (7)$$

Where 0<C_i<1, A_i is closer that A^+ than to A^- as C_i approaches 1.

Step 4: To rank alternatives according to the relative closeness, so the ranking is done based on parameter C_i in decreasing order which means the higher values C_i is closer to positive ideal solution.

III. Problem Statement

This paper applies Fuzzy TOPSIS method to solve MCDM problem such as in grid HPC and cloud environment. As mentioned, the cloud broker should utilize smart scheduler, optimal virtual machine placement (OVMP), to reach their predefined objectives [12]. The cloud such as other distributed system environment is depicted in Fig. 2.

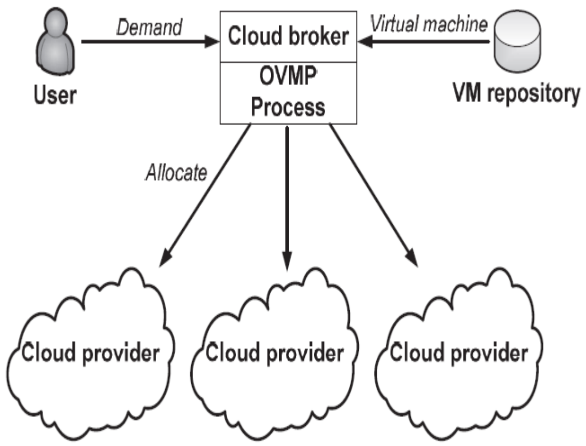

Fig. 2. Cloud Computing Environment [12].

As can be seen in the Fig. 2 when user request arrives, cloud broker searches for virtual machine (VM) needed in VM repository to allocate on datacenter. The scheduler should decide the best alternative between feasible solutions to reach on predetermined metrics. Selecting appropriate metrics is very intricate task so that selecting unsuitable metrics makes misleading positive scheduling [13]. Since the cloud environment is dynamic along with task deadline, we develop Fuzzy TOPSIS module in cloud broker to aid cloud broker not to violate user QOS. Next step is to determine appropriate criteria to cover both user and provider perspectives such as tardy tasks (the number of tasks which violate their due date), makespan (total execution time), throughput (the number of accomplished tasks in time unit) and Cumulative completion time (as indirectly insight resource utilization) the reason why we adopt aforementioned metric used in [13-14]. Moreover, the two first are related to the user perspective whereas the rest pertain to provider viewpoint. Assume that user request arrives with the set of tasks along with their specifications in terms of execution time and deadline. Then the broker should decide based on tasks requirement and system free resources to reach compromising the criteria which have potentially conflicts.

IV. EVALUATION

After To evaluate the effectiveness of fuzzy TOPSIS on broker performance, we take a broker snapshot which received user workload. The workload includes set of tasks possibly independent such as in current case study. Also, tasks can be both interactive and batch processing. Here, tasks are independent and batch processing in nature such as in grid environment. Take specifications in terms of execution time and due date are brought in table 1.

Table 1. Task Specification

Task No.	Task Execution Time (t_i)	Due Date of t_i (d_i)
t_1	1	2
t_2	1	2
t_3	2	3
t_4	3	4

All feasible scheduling are depicted in Fig. 3 though Fig. 5. In addition, the tardy tasks are hatched.

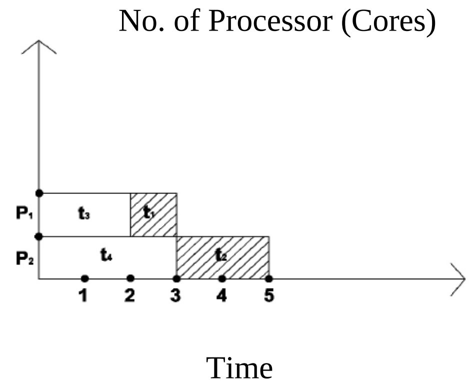

Fig. 3. Feasible Solution 1

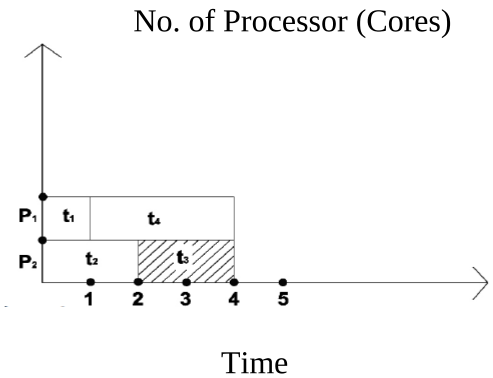

Fig. 4. Feasible Solution 2

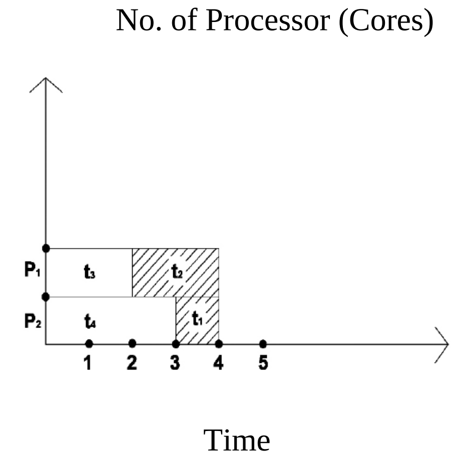

Fig. 5. Feasible Solution 3

Criteria value determination is brought in table 2. Also, metric (criteria) Cumulative Completion for schedule S is calculated by equation (8) as below:

Cumulative Completion for scheduling (S) = $\sum_{i=1}^{n}(1 + C_{max} - t_{finish\_time})^{*}t_j$ (8)

This metric gives partial insight regarding to resource utilization. The only shortcoming of this metric is to use in static environment such as grid static workload in which all of task execution time are determined before execution starts. However, in dynamic environment such as cloud, the broker can apply mining technique to estimate same workload behavior.

Table 2. Determination of Criteria for Feasible Solutions

Solutions No.	Tardy Tasks	Makespan	Throughput	Cumulative Completion
Feasible Solution 1	2	5	0.8	19
Feasible Solution 2	1	4	0.8	24
Feasible Solution 3	2	5	1	23

The broker of distributed system faces with MCDM problem at the moment user request receives. Before it launches the work, it defines criteria weights as their importance. At least two approaches are applied such as mining in data history of datacenter machines along with applying AHP method using pair comparison of administrators' interview [11, 15]. We consider weight vector as W= (0.2, 0.4, 0.2, 0.2). In this example, we have three feasible solutions as alternatives (m=3) along with four criteria (n=4). Moreover, the metrics tardy task (TT), makespan (MS) are known as cost type criteria whereas the metrics throughput (TP) and Cumulative Completion (CC) are known as benefit type criteria. Thus, the dimension of decision matrix A is 3X4. So, initial decision matrix is as below:

$$A_{3\,x\,4}= \begin{bmatrix} 2 & 5 & 0.8 & 19 \\ 1 & 4 & 0.8 & 24 \\ 2 & 5 & 1 & 23 \end{bmatrix} \qquad (9)$$

As nature and source of data are different, it is normalized according to equation (2).

$$R_{3\,x\,4}= \begin{bmatrix} 0.67 & 0.62 & 0.53 & 0.50 \\ 0.33 & 0.49 & 0.53 & 0.63 \\ 0.67 & 0.62 & 0.66 & 0.60 \end{bmatrix} \qquad (10)$$

Then weighted normalized decision matrix P is calculated as following:

$$P_{3\,x\,4} = R_{3\,x\,4}.W_{4\,x\,4} =$$

$$\begin{bmatrix} 0.67 & 0.62 & 0.53 & 0.50 \\ 0.33 & 0.49 & 0.53 & 0.63 \\ 0.67 & 0.62 & 0.66 & 0.60 \end{bmatrix} \cdot \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.134 & 0.248 & 0.106 & 0.100 \\ 0.660 & 0.196 & 0.106 & 0.126 \\ 0.134 & 0.248 & 0.132 & 0.120 \end{bmatrix} \qquad (11)$$

Step 1: The vector of positive ideal solution A^{+} and negative ideal solution A^{-} are calculated according to equation (3) and (4) as following:

$$A^{+}= (0.134, 0.196, 0.132, 0.126)\,, \quad A^{-} =(0.660, 0.248, 0.106, 0.100) \qquad (12)$$

Step 2: Euclidean distances from positive/negative ideal solution are calculated by equation (5) and (6) and illustrated in table 3:

Table 3. Euclidean distances from positive/negative ideal solution

i	S_i^-	S_i^+
1	S_1^-=0.0637	S_1^+=0.5260
2	S_2^-=0.5273	S_2^+=0.0520
3	S_3^-=0.1421	S_3^+=0.5270

Step 3: Relative closeness C_i for each alternative A_i with respect to positive ideal solution are calculated by equation (7) and illustrated in table 4:

Table 4. Relative Closeness for all Alternatives

i	C_i
1	C_1 = 0.108
2	C_2 = 0.910
3	C_3= 0.212

Step 4: In this step ranking between alternatives are done in decreasing order regarding to relative closeness coefficient. So, the second feasible solution is the best, third and first feasible solution are in the next order. It makes trade-of amongst contending objectives.

V. Conclusion and Future Work

In general form, Multi criteria decision making (MCDM) problems belong to NP-Complete category. Therefore, the need for efficient approach is tangible. Specifically, in distributed systems broker machine containing scheduler and dispatcher modules has limited time to decide for scheduling and dispatching tasks over resources since some tasks may have interdependencies and deadlines. The reason why fuzzy TOPSIS module is suitable as it has low overhead. Also, the strong hypothesis behind it, closeness to ideal solution and farness from anti-ideal solution, makes for the results robustness and compromises amongst objectives. However, this module works well in the static and batch processing environment such as grid computing. The only shortcoming is that it is rather appropriate for cloud dynamic environment as the workload is varying in nature. As for future work, we envisage to develop fuzzy TOPSIS method as strong tools for MCDM problems with utilizing data mining technique to excavate data history for discerning workload behavior then launching improved fuzzy TOPSPS to make best decision.

References

[1] A.S. Tanenbaum & M.V Steen. (2007). distributed systems: principles and paradigms, second edition, prentice hall.

[2] P. Mokarippor and Mirsaeid Hosseini Shirvani, A State Of The Art Survey On DVFS Techniques In Cloud Computing Environment, Journal of Multidisciplinary Engineering Science and Technology (JMEST) 3(50): 2016 .

[3] B.C. Tak, B. Urgaonkar, A. Sivasubramaniam, To move or not to move: The economics of cloud computing. HOTCLOUD 2011, Third USENIX Workshop on Hot Topics in Cloud Computing. 2011.

[4] J. Altmann, M. Kashef, Cost model based service placement in federated hybrid clouds, Future Generation Computer Systems 41 (2014) 79–90.

[5] I. Chamodrakas, D. Martakos, A utility-based fuzzy TOPSIS method for energy efficient network selection in heterogeneous wireless networks, Applied Soft Computing 11 (2011) 3734–3743.

[6] R. A. Krohling, V. C. Campanharo, Fuzzy TOPSIS for group decision making: A case study for accidents with oil spill in the sea, Expert Systems with Applications 38 (2011) 4190–4197.

[7] Hwang, C.L.; Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. New York: Springer-Verlag.

[8] Chen, C. T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. Fuzzy Sets and Systems, 114, 1–9.

[9] Chen, T. Y., & Tsao, C. Y. (2008). The interval valued fuzzy TOPSIS methods and experimental analysis. Fuzzy Sets and Systems, 11, 1410–1428.

[10] Dag˘deviren, M., Yavuz, S., & Kılınc, N. (2009). Weapon selection using the AHP and TOPSIS methods under fuzzy environment. Expert Systems with Applications, 36, 8143–8151.

[11] T.L. Saaty, The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill, New York, NY, USA, 1980.

[12] S. Chaisiri et al., Optimal Virtual Machine Placement across Multiple Cloud Providers, Services Computing Conference, 2009. APSCC 2009. IEEE Asia-Pacific.

[13] Mirsaeid Hosseini Shirvani, Evaluating of Feasible Solutions on Parallel Scheduling Tasks with DEA Decision Maker , Journal of Advances in Computer Research, 6(2),2015: 109-115.

[14] A. Burkimsher, I. Bate, L. S. Indrusiak, A survey of scheduling metrics and an improved ordering policy for list schedulers operating on workloads with dependencies and a wide variation in execution times. Future Generation Computer Systems 29(2013) 2009-2025.

[15] Suh, B., & Han, I. (2003). The IS risk analysis based on a business model. Information Management, 41(2), 149–158.