

## Evaluation of scheduling solutions in parallel processing using DEA FDH model

**Gholam R. Amin\***

*Assistant Professor, Dep. of Computer Engineering, Postgraduate Engineering Center, Islamic Azad University, South Tehran Branch, Tehran, Iran*

**M. S. Hosseini Shirvani**

*Research Scholar, Dep. of Computer Engineering, Islamic Azad University, Sari Branch, Sari, Iran*

### Abstract

This paper gives a new application of DEA to evaluate the scheduling solutions of parallel processing. It evaluates the scheduling solutions of parallel processing using the non-convex DEA model, FDH model. By introducing each solution of parallel processing scheduling as a DMU with some relevant inputs and outputs this paper shows that how the most efficient schedule(s) can be identified.

**Keywords:** Data Envelopment Analysis (DEA); Most efficient schedule; Parallel processing; Free disposal hull (FDH)

### 1. Introduction

Data Envelopment Analysis (DEA) is a linear programming method for measuring the efficiency of Decision Making Units (DMUs) such as firms or public sector agencies, first introduced into the *Operations Research* (OR) literature by Charnes, Cooper, and Rhodes (CCR) [5]. The original CCR model was applicable only to technologies characterized by constant return to scale globally. In what turned out to be major breakthrough, Banker, Charnes, and Cooper (BCC) [4], extended the CCR model to accommodate technologies that exhibit variable return to scale. In order to appear a single efficient unit in the reference set of inefficient DMUs, Deprins *et al.* [7], proposed a new type of DEA model called Free Disposal Hull (FDH) non-convex model. Mathematically the non-convex FDH model compares any DMU with an observed unit. This model has received a considerable amount of research attention. Tulkens [17] developed the methodological issues and applications of FDH model in retail banking, courts, and urban transit. Agrell and Tind [1] generalized a dual approach to non-convex DEA models. Kuosmanen and Post [11] have introduced quadratic DEA fron-

tier and models. Also those authors [12] measured the economic efficiency of incomplete price information. Recently Leleu [13] proposed a linear programming framework for free disposal hull technologies and cost functions.

Because of many successful applications and case studies appeared in the DEA literature, it has been growing rapidly [6, 9, 15]. As applications of DEA, Amin *et al.* [3] proposed an improved integrated minimax DEA model for technology selection. Furthermore Amin and Emrouznejad [2] proposed an efficient compact mathematical form to find the optimal value for the discriminating power used in technology selection. Recently Emrouznejad and Amin [8] presented a new convexity consideration for ratio data. The DEA methodology has some successful applications in computer engineering. For example, Pendharkar [14] proposed a DEA based approach for data preprocessing. In addition, software productivity measurement is discussed by Kitchenham and Mendes [10], and Shafer and Bradford [16] proposed a DEA method for measurement of alternative machine component grouping solutions.

This paper gives a new application of DEA in computer engineering where the evaluation of sche-

\* Corresponding author. E-mail: g\_amin@azad.ac.ir

duling solutions of parallel processing is requested. The paper shows how DEA FDH formulation can be used for evaluating the performance of scheduling solutions in parallel processing. We define each scheduling solution of parallel processing as a DMU with some relevant inputs and outputs.

The remaining of this paper is organized as follows: Section 2 gives a brief explanation of DEA FDH formulation. The section also shows the reason of using the non-convex FDH model instead of the other DEA models.

In Section 3 the problem of evaluation of scheduling solutions of parallel processing is formulated as a DEA FDH model. Also the section gives a numerical illustration. Finally the conclusion remarks are given in Section 4.

### 2. The DEA FDH formulation

Assume that there are  $n$  decision making units (DMUs) each of them producing  $s$  outputs by consuming  $m$  inputs. More formally,  $DMU_j$  is denoted by its inputs and outputs, *i.e.*  $(\mathbf{x}_j, \mathbf{y}_j)$ , where  $\mathbf{x}_j = (x_{1j}, \dots, x_{mj})$  and  $\mathbf{y}_j = (y_{1j}, \dots, y_{sj})$ . DEA is a linear programming technique useful to assess relative efficiency among similar entities of DMUs. Consider the input-oriented CCR model for  $DMU_k$  shown below [5].

$$\theta_k^* = \min \theta$$

Subject to:

$$\begin{aligned} -\sum_{j=1}^n x_{ij} \lambda_j + x_{ik} \theta &\geq 0 \quad i = 1, \dots, m \\ \sum_{j=1}^n y_{rj} \lambda_j &\geq y_{rk} \quad r = 1, \dots, s \\ \lambda_j &\geq 0 \quad j = 1, \dots, n \end{aligned} \tag{1}$$

In the above model (1), CCR model, the projected DMU,  $(\theta_k^* x_k, y_k)$ , may be unobserved. The DEA FDH model is motivated to ensure that efficiency evaluations are effected from only actually observed performances. This idea is shown in the following figure, a one input and one output example.

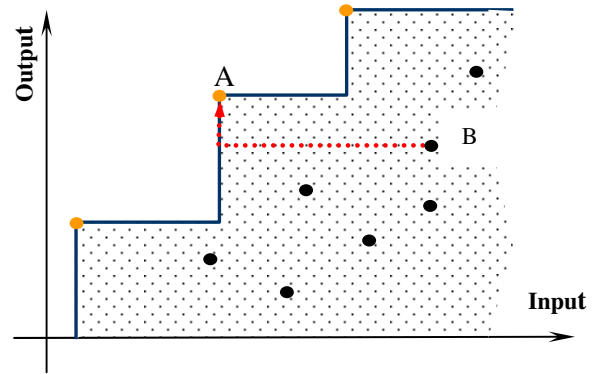


Figure 1. The FDH frontier.

As shown the inefficient  $DMU_B$  is projected to the frontier and its reference set is efficient  $DMU_A$ . Generally, in the FDH formulation each inefficient DMU has only one efficient actual DMU as reference set. Mathematically this assumption converts the convex DEA model (1) into the following non-convex DEA model.

$$\theta_k^* = \min \theta$$

Subject to:

$$\begin{aligned} -\sum_{j=1}^n x_{ij} \lambda_j + x_{ik} \theta &\geq 0 \quad i = 1, \dots, m \\ \sum_{j=1}^n y_{rj} \lambda_j &\geq y_{rk} \quad r = 1, \dots, s \\ \lambda_j &\in \{0, 1\} \quad j = 1, \dots, n \end{aligned} \tag{2}$$

Model (2) gives the input-based FDH technical efficiency score for each  $DMU_k, k = 1, \dots, n$ . The next section introduces a new application of the non-convex FDH model for evaluation of scheduling solutions in parallel processing. The aim of choosing the FDH model for scheduling solutions in parallel processing is also discussed in the following section.

### 3. Evaluation of scheduling solutions

We consider  $n$  activities  $N = \{1, \dots, n\}$  that have to be scheduled in  $p$  parallel identical processors  $P = \{1, \dots, p\}$  under the following assumptions:

- (i) Each activity  $j$  has a due date  $d_j$  and a processing time  $p_j$  that is independent of the processor that processes the activity;
- (ii) There is no preemption or division of activity;
- (iii) All activities are available at time 0.

The completion time of an activity according to schedule  $S$  is denoted by  $c_j^S$ . The maximum completion time for a specific schedule  $S$  (also called the makespan) can be computed as

$$c_{\max}^S = \max \{ c_j^S : j = 1, \dots, n \}$$

The number of tardy activities of a particular schedule solution  $S$  is:

$$N_t^S = \sum_{j=1}^n u_j^S$$

where,

$$u_j^S = \begin{cases} 1 & \text{if } c_j^S > d_j \\ 0 & \text{otherwise} \end{cases}$$

To illustrate the problem consider the following simple example shown in Figure 1. Three activities, A, B and C, to be scheduled on two parallel processors,  $P_1$  and  $P_2$ , and assume that the current time is 0. Activity A and B have both a production time of 1 and both are due at time  $t = 1$ , while activity C has a production time of 2 and is due at time  $t = 3$ .

Figure 2 illustrates three possible schedule solutions where tardy activities appear in gray. The first two schedules are non-dominated, while the last one is dominated by the other two, Solutions 1 and 2. In this paper we assume each scheduling solution is a DMU. So we can define some suitable inputs and outputs for each solution. For instance, consider three solutions shown in Figure 2.

The number of tardy activities and the maximum completion time of a specified schedule  $S$  are both inputs. The number of activities done without tardy is an output of a specified schedule. So the solutions shown in Figure 2 can be summarized as DEA observations in Table 1:

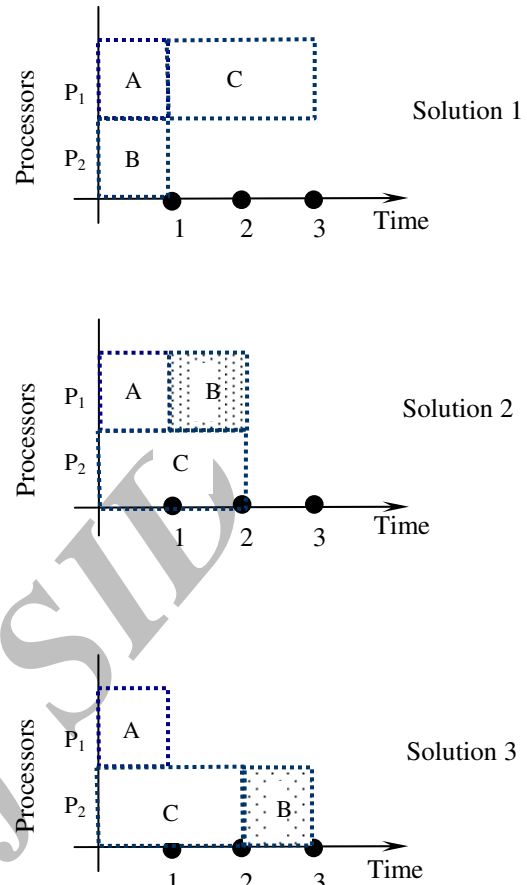


Figure 2. An illustrated example.

Table 1. The corresponding DEA data.

DMU No.	Tardy activities (I <sub>1</sub> )	Maximum completion (I <sub>2</sub> )	No. of activities without tardiness (O)
Solution 1	0	3	3
Solution 2	1	2	2
Solution 3	1	3	2

Now we apply the DEA FDH model for Table 1. The following model evaluates the efficiency score of Solution 3.

$$\theta_3^* = \min \theta$$

Subject to:

$$-\lambda_2 - \lambda_3 + \theta \geq 0$$

$$-3\lambda_1 - 2\lambda_2 - 3\lambda_3 + 3\theta \geq 0$$

$$3\lambda_1 + 2\lambda_2 + 2\lambda_3 \geq 2$$

$$\lambda_1, \lambda_2, \lambda_3 \in \{0, 1\}$$

Using the DEA-solver software provided by Cooper et al. [6] the above FDH model has the following unique optimal solution:

$$(\lambda_1^*, \lambda_2^*, \lambda_3^*, \theta_3^*) = (0, 1, 0, 1)$$

So the third solution is inefficient and the single reference set is Solution 2. Applying the DEA-solver, [6], for Table 1 also shows that the other DMUs are efficient. If we use a ranking DEA procedure, [3], for the data shown in Table 1 it can identify the most efficient DMU is Solution 1. Decision Maker (DM) can use any DEA ranking method if he (she) encounters more than one FDH efficient unit. In general for the given scheduling solutions of parallel processing the methodology proposed in this paper can be applied for selecting the most efficient schedule. Note that the solutions can be obtained by any scheduling algorithm. In fact the DEA FDH model can be applied for the evaluation of the existing scheduling procedures. Using the non-convex FDH model ensures us the projected solution is a given observed one.

#### 4. Conclusion remarks

This paper started with the motivation of the evaluation of scheduling solutions in parallel processing, formulated it as a FDH DEA mathematical model and evaluates the most efficient schedule(s). It is shown that how a given group of scheduling solutions in parallel processing can be evaluated efficiently. Therefore the contribution of this paper is that it introduced a new application of DEA in the parallel processing field of computer engineering.

#### References

- [1] Agrell, P. J. and Tind, J., 2001, A Dual Approach to Nonconvex Frontier Models. *Journal of Productivity Analysis*, 16(2), 129-147.
- [2] Amin, Gholam R. and Emrouznejad, A., 2007, A note on DEA models in technology selection: an improvement of Karsak and Ahiska's approach. *International Journal of Production Research*, 45(10), 2313-2316.
- [3] Amin, Gholam R., Toloo, M. and Sohrabi, B., 2006, An improved MCDM DEA model for technology selection. *International Journal of Production Research*, 44 (13), 2681–2686.
- [4] Banker, R. D., Charnes, A., Cooper, W.W., 1984, Some models measuring the efficiency of decision-making units. *Management Science*, 30, 1078–1092.
- [5] Charnes, A., Cooper, W.W., Rhodes, E. 1978, Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2, 429-444.
- [6] Cooper, W. W., Seiford, L. M. and Tone, K., 2006, Introduction to Data Envelopment Analysis and its Uses with DEA-Solver Software and References, *Springer*.
- [7] Deprins, D., Simar, L., and Tulkens, H., 1984, Measuring labor-efficiency in post offices, the performance of public enterprises: Concepts and measurements. *Amsterdam: North-Holland Publishing Company*.
- [8] Emrouznejad, A. and Amin, Gholam R., 2009, DEA models for ratio data: Convexity consideration. *Applied Mathematical Modelling*, 33(1), 486-498.
- [9] Emrouznejad, A. Parker, B. and Tavares, G., 2008, Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences*, 42(3), 151-157.
- [10] Kitchenham, B., and Mendes, E., 2004, Software Productivity Measurement Using Multiple Size Measures. *IEEE Transactions on Software Engineering*, 30(12), 1023-1035.
- [11] Kuosmanen, T., Post, T., 2002, Quadratic data envelopment analysis. *Journal of the Operational Research Society*, 53(11), 1204-1214.
- [12] Kuosmanen, T., Post, T., 2003, Measuring economic efficiency with incomplete price information. *European Journal of Operational Research*, 144 (2), 454-457.
- [13] Leleu, H., (2006, A linear programming framework for free disposal hull technologies and cost functions: Primal and dual models. *European Journal of Operational Research*, 168, 340-344.
- [14] Pendharkar, P. C., 2005, A Data Envelopment Analysis-Based Approach for Data Preprocess-

ing. *IEEE Transactions on Knowledge and Data Engineering*, 17(10), 1379-1388.

- [15] Ray, S. C., 2004, *Data Envelopment Analysis, Theory and Applications for Economics and Operations Research*, Cambridge University Press, USA.
- [16] Shafer, S. M., and Bradford, J. W., 1995, Efficiency measurement of alternative machine component grouping solutions via data envelopment analysis. *IEEE Transactions on Engineering Management* 42(2), 159-165.
- [17]Tulkens, H., 1993, On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. *Journal of Productivity Analysis*, 4(1-2), 183-210.

Archive of SID