FUZZIFICATION OF THE RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEM

A Fight against Nature

Anikó Csébfalvi, György Csébfalvi and Sándor Danka
University of Pécs, Pécs, Hungary

Keywords: Project scheduling, Stochastic scheduling, Fuzzy scheduling, Resource-constrained project, Heuristic algorithm, Simulation.

Abstract: In a recent article (Bhaskar et al., 2011) the authors proposed a heuristic method for the resource-constrained project scheduling problem (RCPSP) with fuzzy activity times. The apropos of this state-of-the-art work, we try identify and illuminate a popular misconception about fuzzification of RCPSP. The main statement of their approach, similarly to the other fuzzy approaches, is simple: the project completion time can be represented by a "good" fuzzy number. This statement is naturally true: in a practically axiomatic fuzzy thinking and model building environment, using only fuzzy operators and rules, we get a fuzzy output from the fuzzy inputs. But the real problem is deeper. The possibilistic (fuzzy) approach, traditionally, defines itself against the probabilistic approach, so in the "orthodox" fuzzy community everything is prohibited which is connected to somehow to the probability theory. For example, the Central Limit Theorem (CLT) is in the taboo list of this community. We have to emphasize, CLT is a humanized description of a miracle of nature. When we fight against CLT, we fight against nature. The situation in the "neologist" fuzzy community is not better, because they try to redefine somehow the probability theory within the fuzzy approach without using "forbidden" statistical terms. In this paper, we will show that the nature is working totally independently from our "magic" abstractions. According to the robustness of CLT, the distribution function of the completion time of real-size projects remains nearly normal, which is a manager friendly, natural and usable result. An abstraction and its "natural" operators are unable to modify the order of nature. When we want to add a practical scheduling method to the project managers we have to destroy the borders between the probabilistic and possibilistic approaches and have to define a "unified" approach to decrease the gap between scientific beliefs and reality. In this paper we present a unified (probabilistic/possibilistic) model for RCPSP with uncertain activity durations and a concept of a heuristic approach connected to the theoretical model. It will be shown, that the uncertainty management can be built into any heuristic algorithm developed to solve RCPSP with deterministic activity durations. The essence and viability of our unified model will be illustrated by a fuzzy example presented in the recent fuzzy RCPSP literature.

1 INTRODUCTION

In a recent article (Bhaskar et al., 2011) the authors proposed a heuristic method for the resource-constrained project scheduling problem (RCPSP) with fuzzy activity durations. The apropos of this state-of-the-art work, we try identify and illuminate a popular misconception about fuzzification of RCPSP. The main statement of their approach, similarly to the other fuzzy approaches, is simple: the project completion time can be represented by a "good" fuzzy number. This statement is naturally true: in a practically axiomatic fuzzy thinking and model building environment, using only fuzzy operators and rules, we get a fuzzy output from the fuzzy inputs. But the real problem is deeper. The possibilistic (fuzzy) approach, traditionally, defines itself against the probabilistic approach, so in the "orthodox" fuzzy community everything is prohibited which is connected to somehow to the probability theory. For example, the Central Limit Theorem (CLT) is in the taboo list of this community. We have to emphasize, CLT is a humanized description of a miracle of nature. When

Csébfalvi A., Csébfalvi G. and Danka S..
FUZZIFICATION OF THE RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEM - A Fight against Nature.
DOI: 10.5220/0003644402860291
In *Proceedings of the International Conference on Evolutionary Computation Theory and Applications* (ECTA-2011), pages 286-291
ISBN: 978-989-8425-83-6

we fight against CLT, we fight against nature. The situation in the "neologist" fuzzy community is not better, because they try to redefine somehow the probability theory within the fuzzy approach without using "forbidden" statistical terms. In this paper, we will show that the nature is working totally independently from our "magic" abstractions. According to the robustness of CLT, the distribution function of the completion time of real-size projects remains nearly normal, which is a manager friendly, natural and usable result. An abstraction and its "natural" operators are unable to modify the order of nature. When we want to add a practical scheduling method to the project managers we have to destroy the borders between the probabilistic and possibilistic approaches and have to define a "unified" approach to decrease the gap between scientific beliefs and reality. In this paper we present a new unified (probabilistic/possibilistic) model and a conception of a heuristic connected to the unified model for RCPSP with uncertain activity durations. In Section 2 we present a unified theoretical model. In Section 3 we describe the conception of the uncertainty management according to the theoretical model. The essence and viability of our unified model will be illustrated by a fuzzy example in Section 4. Finally, Section 5 draws conclusions from this study.

2 THEORETICAL MODEL

In this section we describe the theoretical model for RCPSP with uncertain activity durations. The approach produces "robust" schedules which are immune against uncertainties in the activity durations. The optimality criterion is defined as a linear combination (weighted sum) of resource-feasible makespans connected to the key terms of the applied uncertainty formulation. Theoretically the optimal robust schedule searching process is formulated as a multi-objective mixed integer linear programming problem (MOMILP) where the number of objectives corresponds to the number of key terms (parameters) of uncertainty formulation.

In this paper, we replaced the MOMILP with a MILP by scalarization. The resulting MILP can be solved directly in the case of small-scale projects within reasonable time. The proposed model is based on the so-called "forbidden set" concept. The output of the model is the set of the optimal conflict repairing relations. Obviously, the solution of the problem depends on the choice of the weights for the objective functions.

In order to model uncertain activity durations in projects, we consider the following resource-constrained project-scheduling problem: A single project consists of N real activities $i \in \{1, 2,..., N\}$.

In this paper, without loss of generality, we assume that each activity duration can be described by three parameters: $\{D_{i1}, D_{i2}, D_{i2}\}$, $i \in \{1, 2,..., N\}$, where triplet $\{D_{i1}, D_{i2}, D_{i2}\}$ may define a triangular membership function in the possibilistic approach, or a density function from beta distribution in the probabilistic approach. We have to note, that in the probabilistic approach the triplet is estimated from a sample using standard statistical tools, assuming that the future can be described from the past, but in the possibilistic approach it is only an abstraction which describe the future according to knowledge of the project managers.

The fuzzy community, under the spell of the challenging but manageable nature of the membership function (it is non-smooth composite of linear segments) tries to recreate everything from the beginning. For example, "normalization" is a "coded message" that the triangle is not a density function, and the horizontal line corresponding to "α-cut" is a theoretically questionable replacement of the two vertical lines, which define the confidence interval in the probabilistic approach. Changing the position of α we change our risk-taking habit, but, at the same time, we omit/add duration segments with totally different left/right tail probability (Figure 1).

Our opinion about the uncertainty management in project scheduling is very simple: we have to replace the triangular membership function with the equivalent triangular density function, have to let the CLP to work. Formalisms which in the uncertainty dimension, try to redefine statistical terms without statistical terms, are meaningless and misleading.

The activities are interrelated by precedence constraints: Precedence constraints force an activity not to be started before all its predecessors are finished. These are given by network-relations $i \to j$, where $i \to j$ means that activity j cannot start before activity i is completed. Furthermore, activity $i = 0\,(i = N+1)$ is defined to be the unique dummy source (sink). Let NR be the set of the network relations.

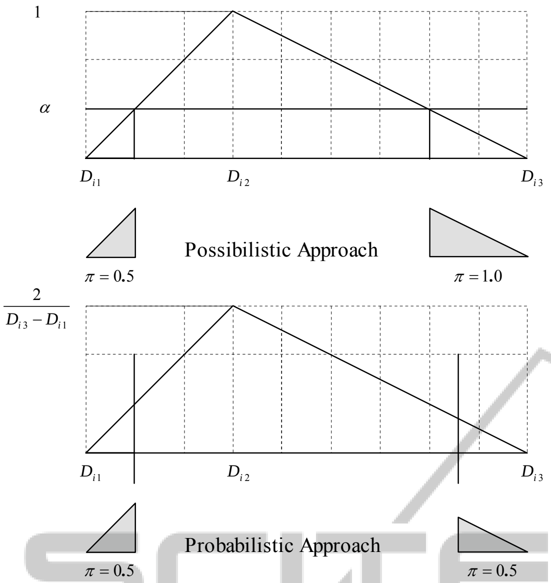

Figure 1: Possibilistic and probabilistic approaches.

Let R denote the number of renewable resources required for carrying out the project. Each resource $r \in \{1,...,R\}$ has a constant per period availability R_r. In order to be processed, each real activity $i \in \{1, 2,..., N\}$ requires $R_{ir} \geq 0$ units of resource $r \in \{1,...,R\}$ over its duration.

A schedule is network-feasible if satisfies the predecessor-successor relations:

$$S_i + D_i \leq S_j \text{, for each } i \to j \in NR \tag{1}$$

Let $\Re$ denote the set of network-feasible schedules. For a network feasible schedule $S \subset \Re$, let $A_t = \{i \mid S_i \leq t < S_i + D_i\}, t \in \{1,...,T\}$ denote the set of active (working) activities in period t and let

$$U_{tr} = \sum_{i \in A_t} r_{ir} \text{ , } t \in \{1,...,T\},\ r \in \{1,...,R\} \tag{2}$$

be the amount of resource r used in period t, where T is an upper bound of the resource-feasible makespan.

A network-feasible schedule $S \subset \Re$ is resource-feasible if satisfies the resource constraints:

$$U_{tr} \leq R_r \text{ , } t \in \{1,...,T\},\ r \in \{1,...,R\} \tag{3}$$

Let $\overline{\Re} \subseteq \Re$ denote the set of resource-feasible schedules. The presented unified MILP formulation is based on the forbidden set concept.

In MILP model the total number of zero-one variables is $|RR|$, and the formulation is based on well-known "big-M" constraints. The presented MILP model is a modified and simplified version of the original forbidden set oriented model developed by Alvarez-Valdés and Tamarit for the deterministic case.

A forbidden activity set is identified such that: (1) all activities in the set may be executed concurrently, (2) the usage of some resource by these activities exceeds the resource availability, and (3) the set does not contain another forbidden set as a proper subset. A resource conflict can be repaired explicitly by inserting a network feasible precedence relation between two forbidden set members, which will guarantee that not all members of the forbidden set can be executed concurrently. We note, that an inserted explicit conflict repairing relation (as its side effect) may be able to repair one or more other conflicts implicitly, at the same time.

Let $T = \sum_{i=1}^{N} D_{i3}$, which is an "extremely weak" resource-feasible upper bound and fix the position of the unique dummy sink in period $T+1$. Naturally, this "weak" upper bound can be replaced by any "stronger" one.

Let $D_i = D_{i1}, i \in \{1, 2,..., N\}$, and let F denote the number of forbidden sets and let RR_f denote the set of explicit repairing relations for forbidden set F_f F_f, $f \in \{1,2,..., F\}$ according to the "optimistic" durations and resource-feasible upper-bound T.

Let

$$RR = \left\{ \bigcup_f RR_f \,\middle|\, f \in \{1,2,..., F\} \right\} \tag{4}$$

denote the set of all the possible repairing relations. In the forbidden set oriented model, a resource-feasible schedule is represented by the set of the inserted resource conflict repairing relations (Alvarez-Valdés and Tamarit, 1993). According to the implicit resource constraint handling, in this model the resource-feasibility is not affected by the feasible activity shifts (movements).

In the time oriented model, a resource-feasible schedule is represented by the activity starting times. In this model, according to the explicit resource constraint handling, an activity movement may be able to destroy the resource-feasibility.

It is very important to note, that after inserting an appropriate conflict repairing set, the "immunised" schedule will invariant to the duration change. In other words, the schedule will be resource-feasible on the set of the possible (and allowed) activity durations because we immunised it according to the optimistic (shorter) durations:

$$D_i = \left[D_{i1}, D_{i3}\right], i \in \{1, 2, ..., N\} \tag{5}$$

Let S_{pi} denote the starting time of activity i, where $i \in \{0, 1, 2, ..., N+1\}$ and $p \in \{1,2,3\}$. By definition, in the optimistic, most likely, and pessimistic schedules the durations are optimistic, most likely, and pessimistic durations:

$$D_i = D_{pi},\ i \in \{1, 2, ..., N\},\ p \in \{1,2,3\} \tag{6}$$

Defining the binary decision variables:

$$Y_{ij} = \begin{cases} 1 & \text{if } i \to j \text{ inserted} \\ 0 & \text{otherwise} \end{cases},\ i \to j \in RR \tag{7}$$

the following MILP model arises:

$$\sum_{p=1}^{3} W_p * S_{N+1\,p} \to \boldsymbol{min} \tag{8}$$

$$\sum_{i \to j \in RR_f} Y_{ij} \geq 1, \quad f \in \{1, ..., F\} \tag{9}$$

$$S_{ip} + D_{ip} \leq S_{jp} + \left(\overline{S}_i - \underline{S}_j + D_{ip}\right) * \left(1 - Y_{ij}\right) \quad i \to j \in RR,\ p \in \{1,2,3\} \tag{10}$$

$$S_{ip} + D_{ip} \leq S_{jp},\ i \to j \in NR,\ p \in \{1,2,3\} \tag{11}$$

The objective function (8) minimizes the linear combination of the resource-constrained makespans, where the weights characterize risk-taking habit of the project manager (for example: "best pessimistic" may be a good scheduling policy, when the project manager is a risk-avoider).

Constraint set (9) assures the resource feasibility (we have to repair each resource conflict explicitly or implicitly, therefore from each conflict repairing set we have to choose at least one element).

Constraint sets (10) take into consideration the precedence relations between activities in the function of the inserted repairing relations.

Constraint sets (11) take into consideration the original precedence (network) relations between activities.

In the "big-M" formulation $\underline{S}_i \left(\overline{S}_i\right)$ define the earliest (latest) starting time of activity i, in the optimistic schedule according to upper-bound T.

We have to note again, that the optimal solution is a function of W_p, $p \in \{1,2,3\}$ weights. According to the model construction in the optimal schedule every possible activity movement is resource feasible, and schedule is "robust" because it is invariant to the variability of activity durations. In other words, a non-critical activity movement (a non-critical delay) or longer (but possible) activity duration is unable to destroy the resource feasibility of the schedule.

3 HEURISTIC ALGORITHM

In this section, describe the conception of a heuristic algorithm connected to the presented theoretical model, Without loss of generality, we assume that we have a deterministic list scheduling algorithm width forward-backward improvement (FBI) to produce resource-feasible schedules in a arbitrary metaheuristic frame. According to the essence of the algorithm, we generate the resource-feasible schedules by taking the selected activities one by one in the given activity order and scheduling them at the earliest (latest) feasible start time using the optimistic activity durations. After that, using FBI we try to improve the quality of the generated schedule.

When the algorithm, in the forward-backward list scheduling process, inserts a precedence relation between an already scheduled activity and the currently scheduled activity whenever they are connected without lag, than we get a schedule without "visible" resource-conflicts in which, according to applied "thumb rule", the number of "hidden" conflicts is drastically decreased.

The importance of the "thumb rule" may be explained by the fact, that in this way we are able to resolve resource conflicts, without explicit forbidden set computation. After that, the algorithm is able (in exactly one step) to repair all of the hidden (invisible) conflicts, inserting always the "best" conflict repairing relation for each forbidden set.

In this context "best" means a relation between two forbidden set members for which the lag is maximal. Naturally, the algorithm memorizes the best schedule found so far by computing the durations of the schedule according to the key point durations. In the search process, according to the "from optimistic to pessimistic" strategy, the algorithm resolves the visible (hidden) resource conflicts using the optimistic durations, after that replaces the optimistic durations with the most likely, and pessimistic ones. The algorithm exploits the fact, that we can not destroy the resource-feasibility, replacing the optimistic durations in a conflict free optimistic solution with longer durations.

After the "best conflict repairing combination" searching phase, the makespan distribution function is generated by simulation. In the simulation phase we have to replace the membership functions with the appropriate density functions (for example: we replace a triangular membership function with a triangular density function and use a triangular random number generator to get duration instances).

4 EXAMPLE

The algorithm of the proposed approach has been programmed in Compaq Visual Fortran 6.5. To solve the presented problem to optimality the callable version of Cplex 12.2 was used. The computational results were obtained by running the algorithm on a 1.8 GHz Pentium IV IBM PC under Microsoft Windows XP® operating system. The conception of Section 3 was inserted to the "Sounds of Silence" harmony search metaheuristic frame (Csébfalvi et al., 2008).

The fuzzy example was borrowed from Bhaskar, Pal, and Pal. The project is shown in Table 1 and Figure 2. The project has only one renewable resource type. The "weak" upper-bound is $T = 397$.

Table 1: A fuzzy RCPSP.

Activity	Duration	Resource requirement
1	{42, 50, 61}	8
2	{36, 40, 42}	17
3	{35, 50, 79}	12
4	{39, 50, 59}	3
5	{16, 25, 30}	13
6	{43, 51, 57}	17
7	{52, 58, 69}	16
Resource availability		30

According to the presented fuzzy RCPSP algorithm, which is based on a "distance base ranking of fuzzy numbers" method, the "good" schedule obtained by the heuristic is: $\{212, 249, 278\}$.

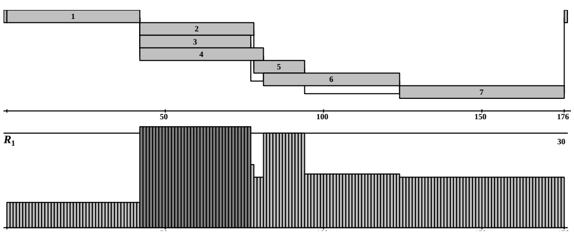

Figure 2: A fuzzy RCPSP with optimistic durations.

Because the project is extremely small, we can prove by explicit enumeration, that this result is wrong. According to $T = 397$ setting and using the optimistic duration estimations, the problem has only two forbidden sets (see Table 2). The implicit enumeration tree is presented in Figure 3.

Table 2: Forbidden sets and repairing relations.

Forbidden sets		Explicit repairs	Implicit repairs
1	{2,6}	2→6 6→2	2→3 2→4
2	{2,3,4}	2→3 3→2 2→4 4→2 3→4 4→3	6→2

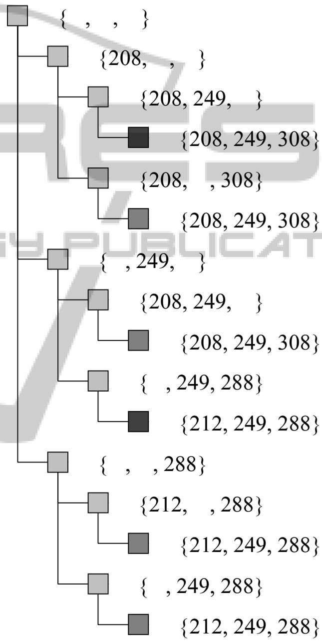

Figure 3: Explicit enumeration tree.

The problem has two non-dominated solutions: $\{208, 249, 308\}$ and $\{212, 249, 288\}$, which illustrate the fact, that a good optimistic schedule not necessarily will be a good pessimistic one and vice versa. The presented "good" solution from (Bhaskar et al., 2011) is better then a non-dominated solution, which is impossible.

When we apply the model of Section 2 to the presented fuzzy problem with unit weights, we get $\{212, 249, 288\}$ as optimal solution within 0.05 sec. In this case, the optimal resource conflict repairing relations are: $2 \to 6$ and $4 \to 2$. The optimistic optimal solution is presented in Figure 4.

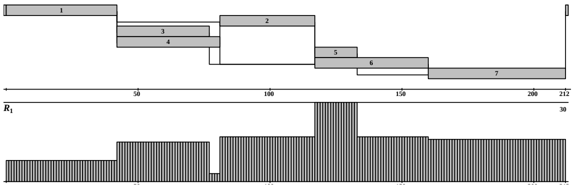

Figure 4: Optimal solution with optimistic durations.

The problem is really simple. The applied harmony search metaheuristic reached the optimal solution in the random repertoire uploading phase setting the repertoire size to ten.

After the "best conflict repairing combination" searching phase, the makespan distribution function is generated by simulation. In the simulation phase we replaced the membership functions with density functions (in this case we replaced the triangular membership functions with a triangular density functions and used a triangular random number generator to get duration instances). We have to mention it, that simulation is a cheap operation, so the sample size may be large enough. In the presented example we set the sample size to ten thousand. Using the Kolmogorov-Smirnov test, we can not reject a null hypothesis that the sample comes from a normal distribution with the following parameters:

$$\pi = 0.158,\ \mu = 256.4,\ \sigma = 10.0 \qquad (12)$$

where $\pi = 0.158$ is the probability of the largest difference (in absolute value) between the observed and theoretical distribution functions when the null hypothesis is true with mean $\mu = 256.4$ and standard deviation $\sigma = 10.0$.

The histogram in Figure 5 reviles the fact, that the nature knows nothing about the fuzzification and does its best according to the CLP.

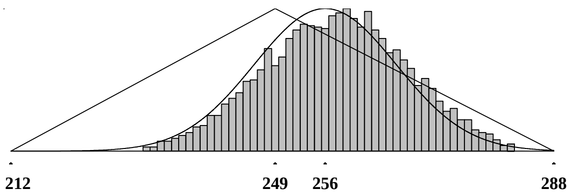

Figure 5: Makespan estimation by simulation.

5 CONCLUSIONS

In this paper, a new unified theoretical model and a concept of the corresponding heuristic approach to solve RCPSP with uncertain activity durations were presented. In the proposed heuristic approach, the uncertainty management is invariant to the applied heuristic frame; therefore it can be built into any other heuristic developed to solve RCPSPs. The essence and viability of our unified approach was illustrated by a fuzzy example presented in the recent fuzzy RCPSP literature. A fast and effective metaheuristic algorithm for large problems is under development and will be presented in a forthcoming paper.

REFERENCES

Alvarez-Valdés, R., Tamarit, J. M., 1993. The project scheduling polyhedron: Dimension, facets and lifting theorems, *Journal of Operational Research*, 96, 204-220.

Bhaskar, T., Pal, M. N., Pal, A. K., 2011. A heuristic method for RCPSP with fuzzy activity times, *European Journal of Operational Research*, 208, 57-66.

Csébfalvi, G., Csébfalvi, A., Szendrői, E., 2008. A harmony search metaheuristic for the resource-constrained project scheduling problem and its multi-mode version, In *Proceedings of the Eleventh International Workshop on Project Management and Scheduling*, Istanbul, 56–59.