Fuzzy resource-constrained project scheduling for GIS software development

L.Bershtein¹ M.Knyazeva¹ I.Rozenberg²

¹Southern Federal University, 44 Nekrasovsky Street, Taganrog, 347922, Russia
²Public Corporation “Research and Development Institute of Railway Engineers”, 27/1, Nizhegorodskaya Street, Moscow, 109029, Russia

Abstract

Resource-constrained project scheduling (RCPS) problem can be defined as a combinatorial problem of constructing a special plan for performing a number of precedence related tasks that have to be executed with the use of limited uncertain resources. In this paper we study the framework of a heuristic approach to deal with uncertainty while modelling RCPS problem by considering a fuzzy set theory. In addition PCRS problems are generally known to be NP-hard because of their combinatorial nature [2], and the goal is to construct a heuristic-based approach to obtain a near optimal solution in a reasonable computational time.

Keywords: Project scheduling, fuzzy graph, fuzzy plan, heuristic priority rules, fuzzy resource allocation, project management software, GIS.

1. Introduction

A resource-constrained project scheduling problem can be defined as a set of precedence tasks (activity network) that have to be sequenced in order to determine optimal starting and finishing times of activities as to minimize project makespan under limited uncertain resources. Thus a fuzzy activity network is a partially ordered set of activities with fuzzy activity durations and other types of fuzzy limited resources (labour man-hour expenditures) that form a directed acyclic fuzzy graph. Fuzzy set theory [1] has been proven to be effective way to handle with such vague information [1].

In literature there are several main approaches to deal with uncertainty, between them authors Herroelen and Leus distinguish such as stochastic project scheduling, reactive scheduling, fuzzy project scheduling, proactive/robust scheduling, etc [4]. Besides fuzzy set theory has been successfully applied to flow-shop and job-shop scheduling problems by Ishibuchi et al., Dubois et al. and Fortemps [5,6,7,8]. Application of fuzzy inference rules to estimate activity durations in software project management was proposed in [24] and suppose to use 14 fuzzy inference rules for semi-automatic estimation to reduce the potential negative aspects of the expert judgment-based estimation for activity duration. Also fuzzy approach to scheduling problem for software development was introduced by Wang X. and Huang W. in [25] by solving fuzzy programming models by means of fuzzy simulation and genetic algorithm as well as hybrid intelligent algorithm. Some questions of crisp project scheduling problem for software development library PSPSWDLIB were discussed in [26].

Usually a fuzzy project scheduling approach is used when activity durations rely on human estimations and due to some unique stages of GIS software development and lack of historical data there is no opportunity to predict durations or times accurately. Thus probability theory is not compatible with specific software development and decision-making projects which are unique to their own sector. However the availability of some resources for an activity performance are usually limited so an efficient project scheduling involves the optimal allocation of such resources to activities. In this paper two ways of handling limited resources are considered: resource-constrained scheduling (RCS) and resource levelling (RL) problems [9].

This paper is organized as follows. Section 2 provides fuzzy arithmetic approach for uncertainty modelling and their configuration overlaps as far as mathematical formulation of the fuzzy project scheduling problem. Section 3 presents fuzzy and qualitative possibility theory to handle with uncertain and preference information, as well as fuzzy priority rules to deal with RCS [10] and finally section 4 provides fuzzy priority heuristic scheduling procedures and resource levelling methods. Section 5 represents numerical example and practical application in fuzzy projects scheduling for GIS ObjectLand software development for cadastral purposes.

2. Basic fuzzy concepts and fuzzy operations

It is supposed that in software development activities the time required to performing certain activities including designing, coding and debugging/testing is not certain and should be estimated by programmer for coding task, for example, according to few cycles it can pass or redebugg to satisfy the target. Thus fuzzy logic approach requires the empirical acquisition of membership functions related to the degree that an activity duration belongs to a fuzzy set durations [11]. Besides experienced project team managers are usually able to specify most and least possible/designated values for ready-time and deadline that can be also flexible [4]. Rommelfanger in [12] suggested a common representation of fuzzy numbers by 6-point piecewise linear membership function to ease and facilitate the acquisition of expert information, where project manager has to provide 3 α-cuts and their corresponding activity duration intervals. Dubois and Prade [13] have shown that a convenient representation of fuzzy numbers is an L-R type flat fuzzy number denoted as $\tilde{M} = (m, \bar{m}, \alpha, \beta)_{LR}$
where \( [\tilde{m}, \tilde{M}] \) is the core of \( \tilde{M} \), and \( \tilde{m}, \tilde{m} \) are the lower and upper modal values of \( \tilde{M} \). \( \alpha, \beta \) are the left-hand and right-hand spreads. A membership function of \( M \) can be expressed by means of symmetric, bell-shaped reference functions \( L \) and \( R \), such that \( L(0) = R(0) = 1 \) and \( L(1) = R(1) = 0 \):

\[
\mu_x(x) = \begin{cases} 
\frac{L(x - \tilde{m})}{\alpha}, & x < \tilde{m}, \\
1, & x \in [\tilde{m}, \tilde{M}], \\
\frac{R(x - \tilde{m})}{\beta}, & x > \tilde{m}, \alpha, \beta \in X.
\end{cases}
\]

(1)

Flat fuzzy number representation with five linear pieces for modelling uncertainty duration in RCPS problems was first presented by Hapke and Slowinski [15] and is shown on Fig.1, symbolic definition can be presented as follows: \( \tilde{D} = (d^e, d^d, d, \tilde{d}, \tilde{d}^e, \tilde{d}^d) \) (2)

Let's consider 3 intuitively well-defined \( \alpha \)-level set:

\( \mu_x(x) = 1 \) means that decision makers consider that the values of \( x \) are certainly belongs to the set of possible values,

\( \mu_x(x) > \lambda \) means that decision makers consider that the values of \( x \) have a good chance of belonging to set of possible values,

\( \mu_x(x) < \varepsilon \) means that decision makers consider that the values of \( x \) have a very poor chance of belonging to set of possible values.

Similarly we can define project ready time and deadline time as follows [16]:

\( \tilde{e} = (\tilde{e}, \tilde{e}, \tilde{e}, \tilde{e}, \tilde{e}, \tilde{e}) \) is shown in Fig.2. (3)

Thus in every \( \alpha \)-level of a fuzzy number we can define optimistic and pessimistic values for each activity duration as well as project makespan. So if the fuzzy set is \( (-\infty, \tilde{e}] \) it means that the level of satisfaction when the deadline is \( x \), and a completion time of a project is not greater than \( \tilde{e} \) it means that the project manager is completely satisfied with a membership function:

\[
\mu_{\langle e, x \rangle}(x) = \sup_{y \geq x} \mu_y(y) = (-\infty, -\infty, -\infty, \tilde{e}, \tilde{e}, \tilde{e})
\]

(4)

Similarly if ready-time is \( \tilde{b} = (\tilde{b}, \tilde{b}, \tilde{b}, \tilde{b}, \tilde{b}, \tilde{b}) \) and fuzzy set \( [\tilde{b}, +\infty) \) shows the level of satisfaction when the ready-time of a project is \( x \), then membership function can be settled as follows:

\[
\mu_{\langle b, x \rangle}(x) = \sup_{y \leq x} \mu_y(y) = (\tilde{b}, \tilde{b}, +\infty, +\infty, +\infty, +\infty)
\]

(5)

In this paper we combine proposed model L-R flat fuzzy numbers with up to five linear pieces with adaptation of fuzzy number representation to RCPS problem by specifying some types of activity duration overlaps for solving RCS problem. Usually such kind of overlapping happens while ranking fuzzy temporal parameters and when it is required to compare fuzzy start and finish times for activities. For example, the start time for an activity should satisfy the precedence relations and value for the start time of activity should be greater than or equal to the finish time of all its predecessors. There were several approaches in literature to deal with ranking order between two fuzzy numbers. Some of them apply compensation areas determined by their membership functions [15,17] and considering types of overlapping showed on Fig.3:

Compensation areas are shown on Fig.4:

In addition Liou and Wang considered approach based on ranking fuzzy numbers with integral value [18]. Let a fuzzy number \( M \) has a left membership function \( f^L_M \) and right membership function \( f^R_M \) and the inverse functions \( g^L_M \) and \( g^R_M \). Then the left and
right integral values of a fuzzy number $\tilde{M}$ can be formulated as follows:

$$I_r(\tilde{M}) = \int_0^1 g^+_{\tilde{M}}(\alpha)d\alpha$$

(6)

$$I_r(\tilde{M}) = \int_0^1 g^-_{\tilde{M}}(\alpha)d\alpha$$

(7)

And a final integral value of fuzzy number $\tilde{M}$ can be defined as a weighted sum of $I_r(\tilde{M})$ and $I_s(\tilde{M})$:

$$I_s(\tilde{M}, \beta) = \beta \times I_r(\tilde{M}) + (1-\beta) \times I_s(\tilde{M}),$$

where $\beta$ denotes the index of optimism, that is determined by project manager, where $0 \leq \beta \leq 1$. The total integral value of $\tilde{M}$ is determined according to Hurwicz’s criterion [19] weighted average of secure and optimism levels [16]. Then we need to define the relationship between two fuzzy numbers $\tilde{M}$ and $\tilde{N}$ as follows:

If $I_s(\tilde{M}, \beta) \geq I_s(\tilde{N}, \beta)$ then $\tilde{M} \geq \tilde{N}$, else $I_s(\tilde{M}, \beta) < I_s(\tilde{N}, \beta)$ then $\tilde{M} < \tilde{N}$.

If $\beta = 0.5$, then the ranking order of fuzzy numbers $\tilde{M}$ and $\tilde{N}$ determined by approach in [18] will be the same as the one that can be determined by Compensation areas are shown on Fig.5 [15,17].

Let’s consider an example with the formulas for left and right integral value, written for 6-point fuzzy number as follows:

$$I_s(\tilde{M}) = 0.5 \times [\mu_l + m + (m^r - m)\lambda]$$

(8)

$$I_s(\tilde{N}) = 0.5 \times [\mu_l + m + (m^r - m)\lambda]$$

(9)

Let’s consider an example on Fig.5:

![Fig. 5: Example of comparing fuzzy numbers.](image)

$\tilde{M} = (5.8,12,14,15,17)$ and $\tilde{N} = (3.9,13,15,16,20)$

Assuming that $\varepsilon = 0$ and $\lambda = \beta = 0.5$, the total integral values for $\tilde{M}$ and $\tilde{N}$ are:

$I_s(\tilde{M}.05) = 0.5 \times 8.25 + 0.5 \times 15.25 = 11.750$

$I_s(\tilde{N}.05) = 0.5 \times 8.5 + 0.5 \times 16.75 = 12.625$

There we have that $\tilde{N}$ is greater than $\tilde{M}$.

3. Fuzzy and possibility approach for project scheduling problem

A software development project $p$ can be represented by a ready-time $\bar{b}$ and deadline $\bar{r}$, where all activities $i$ have to be performed and scheduled according to their start/finish times and precedence relations. In order to schedule all activities, we assume that each activity has specific duration $d_i$ and its execution requires use of some kind of vector resources $N_i = (n_{i1}, n_{i2},...,n_{iq})$, where elements of vector present usage of resource types $k, k=1,\ldots,q$.

Resource availability can be presented by vector $R_i = (m_{i1}, m_{i2},...,m_{iq})$, where $m_k$ represents the availability of resource type $k$. For any period of time we should consider that resources required to perform activity can’t exceed its availability for any type of resource $k$. Precedence relations should be regulated by ranking fuzzy temporal parameters such as fuzzy start time, its duration and fuzzy finish time discussed in section 2. The goal is to determine such schedule that allows to consider some preference and possibility nature of decision making procedure and optimize makespan of project under limited resources.

In order to handle both uncertain and preference information while scheduling we suppose to use qualitative possibility theory introduced by Dubois and Prade [20], and use by Wang [16] to evaluate utility and plausibility for measuring schedule robustness. Thus let’s assume that there is set of U-states and X-possible sequences of his decision. Then a possibility distribution $\pi$ will represent incomplete knowledge on state U and $\mu$ will be possibility distribution presented by decision maker. Finally the utility of decision $d$ will be $x = d(u)$, can be evaluated by combining plausibilities $\pi(u)$ and the utilities $u(x)$. A criterion below was proposed by Dubois and Prade to estimate worth of decision $d$ under uncertain nature of information:

Pessimistic criterion:

$$U.(d) = \inf_{u \in U} \max_{u \in D} (1-\pi(u), \mu(d(u)))$$

(10)

Optimistic criterion:

$$U'(d) = \sup_{u \in U} \min_{u \in D} (\pi(u), \mu(d(u)))$$

(11)

Thus Wang suggested to use pessimistic criterion to measure the robustness of a schedule [16].

For a project let $s$ be a project schedule that shows an activity precedence according to certain resource availability and let $\bar{D}(s)$ be the project duration with $\mu_{\bar{D}(s)}$ possibility distribution of a schedule $s$. A preferred project duration $\bar{P}$ can be calculated as $\bar{P} - \bar{b}$, and measure that a chance of the uncertain project duration by a schedule will be greater than the preferred project duration in the worst situation:

$$\text{Measure} = \inf_s \max_x (1 - \mu_{\bar{D}(s)}(x), \mu_{\bar{P} - \bar{b}}(x))$$

(12)
To illustrate this measure of uncertainty lets observe Fig.6, where measure is calculated as intersection of $\mu_{\hat{D}(x)}(x)$ and $\mu_{\hat{D}^{-}\hat{P}}(x)$, and thus we consider right edge of $\mu_{\hat{D}(x)}(x)$ that shows the worst case of project duration. Thus the plausible duration $\Phi(s) = \max\{\hat{D}(s), \alpha = \text{measure}(s)\}$ (13) from the pessimistic point of view can be defined as maximum value of $\alpha$-level set of $\hat{D}(s)$.

![Fig.6: Illustration of measuring technique.](image)

4. Fuzzy priority heuristic scheduling procedures for RCPS problem and resource leveling

Basically, in literature heuristic approaches for solving PCPS problems have several ways to construct algorithms: single and multiple pass priority rule-based scheduling [21,22], truncated branch-and-bound procedures [3], disjunctive arc concepts, local search techniques and others.

Heuristics using priority rules for constructing list of activities to perform using crisp or fuzzy duration parameters is found to be efficient by many researchers [15], and it is generally known two priority-based approach for scheduling activities: serial and parallel scheduling scheme. In both procedures, tasks are ranked in some order and scheduled according to resource availabilities.

Originally Hapke and Slowinski [15] have proposed a parallel scheduling procedure for fuzzy projects based on fuzzy priority rules and fuzzy time incrementation. Let’s present some rules for constructing priority-based list of activities to be scheduled as discussed in Table 1. Where $\bar{p}_j$ - duration, $r_j$ - is the type requirements of resource $R_k$, $E_{\hat{j}}$ - early finishing times of activity $j$, $F_{\hat{j}}$ - earliest starting times of activity $j$. Let’s $S_j = \{1,...,S\}$ be the set of activities to be scheduled, then we can calculate distribution parameters of each activity $j$ ($H^{\hat{j}}_o$ and $H^{\hat{j}}_p$) according to optimistic and pessimistic criterion, and by using parallel scheduling scheme based on fuzzy ranking technique provided in section 2 we can construct feasible schedule profile for each activity in the project.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Name</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>EST</td>
<td>Early start time</td>
<td>min($\bar{E}_{i}$)</td>
</tr>
<tr>
<td>EFT</td>
<td>Early finish time</td>
<td>min($\bar{F}_{j}$)</td>
</tr>
<tr>
<td>MINSLK</td>
<td>Minimum slack</td>
<td>min($\bar{P}_{j}$)</td>
</tr>
<tr>
<td>SPT</td>
<td>Shortest processing time</td>
<td>max($S_{j}$)</td>
</tr>
<tr>
<td>MIS</td>
<td>Most immediate successors</td>
<td>max($S_{j}$)</td>
</tr>
<tr>
<td>MTS</td>
<td>Most total successors</td>
<td>max($S_{j}$)</td>
</tr>
<tr>
<td>GRD</td>
<td>Greatest resource demand</td>
<td>$\bar{p}<em>{j} + \sum</em>{k \in P_{j}} \bar{p}_{k}$</td>
</tr>
<tr>
<td>GRPW</td>
<td>Greatest rank positional weight</td>
<td>max($\bar{p}_{j}$)</td>
</tr>
</tbody>
</table>

Table 1: Priority rules for makespan optimization.

Let’s introduce few sets to denote activities in algorithm as follows:

- $A_{es}(\bar{t})$ is a set of activities whose fuzzy earliest starting times $E_{s_{j}}$ are less or equal to a fuzzy moment $\bar{t}$.
- $E_{s_{j}} \leq \bar{t}, \forall j \in A_{es}(\bar{t})$.
- $\bar{t}_{j}$ is the least value among the earliest starting times of activities from $A_{es}(\bar{t})$ and the finishing times of activities from $S(\bar{t})$.
- $S(\bar{t})$ is the set of activities in present fuzzy moment $\bar{t}$ if equation $S_{j} \leq \bar{t} \leq F_{j}$ is performed and $S_{j}, F_{j}$ denotes fuzzy starting and finishing times of activity $j$.
- Fuzzy parallel scheme structure is similar to algorithms provided in [15], considering possibilities to schedule each activity in conditions to resource availability, demand and precedence relations in priority list. A fuzzy priority-based heuristic algorithm is showed in Fig.7:

Choose priority rule:
- Initialize $E_{s_{j}}, (\forall j)$, earliest starting time of activity $j$ using critical path method calculations;
- Initialize $\bar{t}$ = $\underline{t}$;
- Initialize the total resources availability at each scheduling time $\bar{t}$;

repeat

Compose the set $A_{es}(\bar{t})$ of all available activities at scheduling moment $\bar{t}$ for each $j$ from $A_{es}(\bar{t})$ according to the priority rule do:
- calculate the corresponding symmetric probabilistic distribution $p_{j}$;
- if symmetric probabilistic distribution $P_{j}$ does not fit the resource availability then calculate a new $P_{j}$ with asymmetric shape considering the min possible value for the left optimistic pa-
parameter $H^{(0)}_t$, if $P_t$ fits the resource availability then
Schedule $j$ with corresponding starting and finishing times,
Integrate the distribution $P_j$ into project profile and update the resource availabilities,
Update the earliest starting times of all successors of $j$,
end if
end if
end if
if all activities from $\lambda_{ES}(\tilde{t})$ are scheduled then
$\tilde{t} = \max(\tilde{t}, \tilde{t}(\tilde{t}))$
Else
$\tilde{t} = \max(\tilde{t}, a_i + 1)$
end if
until all tasks are scheduled.

Fig.7: A fuzzy priority-based heuristic algorithm.

To analyze performance and evaluate efficiency of a fuzzy priority-based heuristic algorithm we used 20 sets of benchmarks from project scheduling problem library (PSPLIB) [27] and randomly made them fuzzy according to 6-point representation with $\lambda = 0.8$. For comparison benchmark we considered beam search algorithm with beam width $(i)$ [28] as sets of benchmarks in PSPLIB.

The set of problems contains problems with 32 and 62 numbers of activities. Computational experiment for 20 benchmarks was done with 20 random runs and it summarized the average and standard deviation from optimal solution for both approaches. Finally the performance a fuzzy priority-based heuristic algorithm for most of problem sets showed the following results:
average deviation from optimum less than 0.070 for 10 problem sets with 32 numbers of activities, average deviation from optimum less than 0.120 for 10 problem sets with 62 numbers of activities, comparing with beam search with beam width:

\begin{align*}
i = 1 & \text{average deviation from optimum 0.130} \\
i = 2 & \text{average deviation from optimum 0.082} \\
i = 4 & \text{average deviation from optimum 0.061}
\end{align*}

According to this results beam search solution quality can be improved if beam width is increased, the solution quality of developed fuzzy priority-based heuristic algorithm showed good computational results for EST-fuzzy priority rule. Although the average deviation from optimum using fuzzy priority-based heuristic is among 0.070-0.120 it can be changed by applying different priority rules for large-scaled projects.

5. Case study for a GIS Objectland software development.

The case study for software development of GEO information system Objectland can be represented as a large project consisting few aggregative stages for constructing product, showed in Table 2.

![Fig.8: Interface of GEO information system Objectland](image)

To illustrate approaches introduced in this paper we divide the entire software development project into several large stages, which include system requirements, analysis, system definition and developing architecture of a system etc. We also simplify and combine stages into 13 activities, concerning several types of staff member as a types of resource to perform activities above, thus we will consider system engineers, software engineers, hardware engineers and finally supporting engineers as 4 types of resources: $r_1, r_2, r_3, r_4$.

<table>
<thead>
<tr>
<th>Definition</th>
<th>Staff members</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Customer definition design and Layout</td>
<td>General professional users</td>
</tr>
<tr>
<td>2 Problem definition</td>
<td>GIS Specialist-1</td>
</tr>
<tr>
<td>3 Feasibility study</td>
<td>GIS Specialist-2</td>
</tr>
<tr>
<td>4 Requirements specification</td>
<td>GIS developers</td>
</tr>
<tr>
<td>5 Analysis of GIS structure</td>
<td>GIS developers</td>
</tr>
<tr>
<td>6 System design</td>
<td>GIS developers, System managers</td>
</tr>
<tr>
<td>7 Data Acquisitions Constructing database architecture</td>
<td>Database programmers</td>
</tr>
<tr>
<td>8 GIS modelling</td>
<td>GIS developers, System managers, engineers</td>
</tr>
<tr>
<td>9 Cartographic modelling</td>
<td>Data providers</td>
</tr>
<tr>
<td>10 Spatial modelling/interface specification</td>
<td>GIS developers, supporting engineers</td>
</tr>
<tr>
<td>11 Programming/Coding</td>
<td>General programmers, software engineers</td>
</tr>
<tr>
<td>12 Integration testing</td>
<td>Supporting engineers</td>
</tr>
<tr>
<td>13 GIS Fundaments Implementation</td>
<td>Supporting engineers</td>
</tr>
</tbody>
</table>

Table 2: Description of activities.

We assume that $\xi = 0$ and $\lambda = 0.5$ for a fuzzy temporal parameters and the index of optimism is set to 0.5 for this example. Table 3 presents numerical description of project.


Table 3: Numerical description of project.

The total availability of resources assigned to this project is 5/5/4/5 human resources per planning horizon in every moment. Some of activities can be performed in parallel in case of resource availability. In this paper we assume that project managers may need some crisp value for starting project (activity) within a project, so we will consider the smallest value of the α -cut as a start time for the project. For other activities fuzzy starting times can overlap fuzzy finishing times of its predecessors, and thus the preceding activities may not be fully completed because of resource availability. The proposed possibility scheduling approach should be used to evaluate the resource allocation decision to avoid the risk of late pessimistic makespan of a project. For this purposes we recommend to use measuring technique and plausible project duration calculations as showed in formulas (12) and (13) respectively.

After performing calculations we got following results, represented in Table 4 for obtained fuzzy schedule:

Table 4: The obtained fuzzy schedule for project.

The final optimal schedule with respect to limited types of resources and some activities that can be performed simultaneously, showed following activity sequence:

1→(2+3)→4→5→6→(7+9)→8→10→11→12→13 with the following fuzzy project makespan $\bar{d}^* = (72,83,90,92,103,110)$ with resource conflict happened in several cases. In case if few activities can be performed simultaneously without resource conflict, the start time of next activity should be calculated as sum of finish time of previous activity plus maximum between durations of activities in process which can be performed simultaneously.

6. Conclusions

This paper presents a generalized fuzzy heuristic approach to resource-constrained project scheduling (RCPS) problem. Practical solution of this problem is often complex due to its NP-hardness and imprecise temporal parameters involved in the project. Combinatorial nature of such projects suppose application of fuzzy heuristic approach to obtain reasonable solutions under fuzzy conditions. In this paper we present a fuzzy priority heuristic approach with evaluated efficiency for medium-scaled projects within average deviation from optimum solutions less than 12 percent for EST-priority rule. Further investigations will be done to discover an influence of each priority rule to algorithm performance.

7. Acknowledgements

This work has been supported by the Russian Foundation for Basic Research, Project № 14-01-00032a.

References


