Proactive Project Scheduling with a Bi-Objective Genetic Algorithm in an R&D Department

Canan Capa[1], Gunduz Ulusoy[2]

[1]Concordia University, Canada
e-mail: c_capa@encs.concordia.ca

[2]Sabanci University, Turkey
e-mail: gunduz@sabanciuniv.edu

Keywords: proactive project scheduling, multi-objective genetic algorithm, R&D

1. Introduction

During project execution, especially in a multi-project environment unforeseen events arise that disrupt project plans resulting in deviations of project plans and budgets due to missed due dates and deadlines, resource idleness, higher work-in-process inventory and increased system nervousness. Therefore, project schedules should also include solution robustness to cope with the uncertainties such that actually realized activity start times during project execution will not differ much from the baseline schedule. Constructing solution robust schedules requires proactive scheduling techniques.

The literature on proactive project scheduling is relatively scarce. Leus (2003) consider the objective of minimizing the total weighted instability of the schedules from a given deadline. Herroelen and Leus (2004) develop mathematical models for the generation of stable baseline schedules. Van de Vonder et. al. (2006) propose resource flow dependent float factor heuristic as a time buffering technique relying completely on the activity weights. Lambrechts et. al. (2008) focus on disruptions caused by stochastic resource availabilities to generate stable baseline schedules. Van de Vonder et. al. (2008) introduce multiple algorithms to include time buffers in a given schedule while a predefined project due date remains respected. In a recent study, Lambrechts et. al. (2011) analytically determine the impact of unexpected resource breakdowns on activity durations and develop an approach for inserting explicit idle time into project schedules in order to protect them from possible resource unavailability. In addition to these proactive strategies, there are some risk integrated procedures. Shatteman et. al. (2008) develop a computer supported risk management system that allows to identify, analyze and quantify the major risk factors and derive the probability of their occurrence and their impact on the duration of the project activities. Creemers et. al. (2011) propose a quantitative approach that allows to address the risk response process in a scientifically-sound manner and shows that a risk-driven approach is more efficient than an activity-based approach when it comes to analyzing risks. Herroelen (2014) propose a methodology that integrates quantitative risk analysis with reliable proactive/reactive project scheduling procedures.

We consider the preemptive resource constrained multi-project scheduling problem (RCMPSP) with generalized precedence relations in a stochastic and dynamic environment and develop a three-phase model incorporating data mining and project scheduling techniques to schedule the projects in the R&D Department of a leading home appliances company in Turkey. Phase I of the model, uncertainty assessment phase, provides a systematic approach to assess uncertainty by identifying the most important sources of uncertainty, measuring the impacts of these factors to resource usage deviation levels of projects and their activities and generating activity deviation distributions by using the most important data mining techniques: feature subset selection, clustering and classification. Phase II, proactive project scheduling phase, proposes two scheduling approaches using a bi-objective genetic algorithm (GA). Phase III, reactive project scheduling phase, aims at rescheduling the disrupted project activities.

In this paper, our focus is limited to Phase II of the three-phase approach. In Section 2, the problem and the problem environment are explained. In Section 3, we present the solution methodology and in Section 4 we present the main results obtained by the implementation of the proposed proactive project scheduling approach with real data. Finally, in Section 5 we conclude and provide suggestions for future work.

2. Problem definition and environment

The problem on hand is proactive scheduling of the R&D projects with a priori assigned resources in a stochastic and dynamic environment present in the R&D Department of a leading home appliances company in Turkey. A project consists of a number of events and activities that have to be performed in accordance with a set of precedence and resource constraints. Activities require two types of renewable resources: human resource and equipment. Equipment includes machines, mechanisms and laboratories. Non-renewable resources are not considered. The problem environment under consideration contains multiple projects using multi-skilled renewable resources. The resource requirement of activities and hence, the durations of activities are uncertain. The project network is of activity-on-node (AON) type with Finish-to-Start (FS) and Start-to-Start (SS) precedence relations with zero or positive time lags. No precedence relation is assumed between projects. The problem on hand can be considered an extension of the RCMPSP with generalized precedence relations and multi-skilled resources to include preemption, stochastic activity duration and resource availabilities and dynamic arrival of projects. The objective is generating solution robust baseline project schedules and minimizing the completion time for the overall project makespan. Solution robustness is a measure of the difference between realized schedule and baseline schedule. In our case, we use total sum of absolute deviations *(TSAD)* for solution robustness. *TSAD* is sum of the absolute deviations between actual starting times and starting times realized in a set of K simulations over all activities. The problem environment differentiates from those in the literature in that a resource is required for the duration of its usage within an activity rather than for the whole deterministic or stochastic duration of the activity requiring that particular resource. Resources can work on more than one activity in a time period (say, a week) and the duration of the usage of the resources can differ over the periods that the activity is executed. Additionally, the concept of preemption of a resource employed by an activity is introduced.

3. Solution methodology

In this section, we present a bi-objective GA that uses the output of the Phase I, uncertainty assessment phase, and two scheduling approaches each using the bi-objective GA. The aim of these approaches is to generate non-dominated solution robust project schedules with the minimum makespan for the completion of all projects scheduled. Solution robustness is measured with *TSAD* of the schedule through K number of possible schedule realizations in both approaches. The single and multi-project scheduling approaches differ in that the single project scheduling approach considers the remaining part of the schedules of the already active projects as fixed and schedules only the newly arrived project using the currently available resources, whereas the multi-project scheduling approach, schedules all the active projects in the system anew together with the newly arrived project. Since the two scheduling approaches differ in the way they adopt for the scope of scheduling, the definitions of *TSAD* and makespan, thus, the objectives considered in the bi-objective GA also differ although they both try to minimize *TSAD* and makespan. Note that in the proactive project scheduling approaches, a set of non-dominated robust project schedules are generated. From these non-dominated robust schedules, the decision maker can choose the schedule that best fits the current project management environment in the system. Proposed bi-objective GA is an adopted version of NSGA-II suggested by Deb et al. (2002), which uses an explicit diversity generation procedure along with an elite-preservation procedure. An individual is represented by a precedence feasible activity list. We make use of one-point crossover and swap mutation operators. Population management is the same as of NSGA-II. However, our bi-objective GA differs in the schedule generation scheme and chromosome evaluation procedures.

3.1. Schedule generation

Since the work of resources on activities are preemptive, a schedule is represented with the lists of resource, activity, week and amount (r,a,t,k) quadruple. Each (r,a,t,k) quadruple shows that resource r works on activity a at time instant t for k working hours. Our resource schedule generation scheme starts with scheduling the resources of the first activity in the chromosome. Note that, resource order for scheduling is not important since all orders give the same work schedule for that activity. Considering the earliest precedence feasible starting time of activities and starting at the first available time instant, resources are scheduled until they reach their required usage hours. After all the resources of the first activity in the chromosome are scheduled, starting and ending time of that activity is determined by simply checking the work schedules of the resources that activity requires. Then, the earliest starting time of the successor activities are updated. This procedure is repeated until all the activities in the chromosome are scheduled.

3.2. Chromosome evaluation

For a given order of activities both the overall makespan and solution robustness are assessed through a set of *K* realizations mimicking the implementation phase, where a realization corresponds to a sample instance obtained by a simulation run using the activities' percentage resource requirement deviation distributions, which is determined calling Phase I of the three-phase approach. For this purpose, two alternative chromosome evaluation heuristics with the objective of quality robustness represented with makespan and solution robustness expressed in terms of *TSAD* value of the robust activity starting times from their counterparts in all *K* realizations, are considered: chromosome evaluation heuristic I (*CEH-I*), and chromosome evaluation heuristic II (*CEH-II*). *CEH-I* solves a *TSAD* minimization model by LP. Using the activity starting times realized in simulations, this *TSAD* minimization model aims at finding robust start times that minimizes the *TSAD* value of the scheduled activities. Note that resulting activity start times might be completely different than the activity starting times in *K* realizations and they might be resource-infeasible. Thus, using the resulting robust starting times, first, feasibility of these starting times is checked and if infeasible, the schedule is fixed with deferring the infeasible activities. On the other hand, in *CEH-II, K* realizations are sorted in their non-domination levels using the corresponding makespan and *TSAD* values and among the schedules that have a rank of 1, the schedule having minimum *TSAD* is selected as the robust schedule of the chromosome. The makespan and the *TSAD* values of the resulting schedule are used as performance measures of the chromosome. Note that the *TSAD* in the multi-project scheduling approach includes the deviations of the starting times of the existing activities as well.

4. Implementation with real data

For the implementation, 37 completed R&D projects are used as test instances to compare the performances of the two proactive project scheduling approaches developed. All codes are written in Microsoft Visual Studio C# and CPLEX 12.5 is used as the MILP solver. All tests are performed on a computer with a 3.20 GHz Intel(R) Core(TM) i7 CPU 960 processor and 8 GB of RAM. The best combination of the parameters to be used in the bi-objective GA is determined through extensive experimentation. In the following sections, the results of the scheduling approaches obtained by using the crossover rate of 0.95, mutation rate of 0.05, population size of 50 and the number of generations and the number of schedule realizations for a chromosome are taken as 50, and 100, respectively.

4.1. Data

All 37 projects are the projects initiated between 2007 and 2011. Project networks are of AON type FS and SS precedence relations with zero and positive time lags. There is no precedence relation between projects. The two types of renewable resources are: Human resource and equipment. Activities require from one human resource to a total of more than 11 human resources and equipment. While the weekly capacity of human resources is 45 working hours, these capacity values differ from nine working hours to 672 working hours for the resources in the equipment category.

4.2. Results

The results obtained by the use of *CEH-I* and *CEH-II* in the bi-objective GA are compared with respect to CPU time, diversity of the solutions and solution quality. Because of space limitations, we do not present the tables showing results but we provide the main results obtained with the analysis of the results. It is seen that the CPU time required to schedule the projects is less for almost all projects when *CEH-II* is used instead of *CEH-I* since fitness of a chromosome is calculated using an already generated schedule in *CEH-II.* Thus, it seems sorting the schedules generated in the simulation with respect to their non-domination level requires less computational time than solving the *TSAD* minimization model and generating a new schedule using the output of the *TSAD* minimization model. It is also seen that when *CEH-I* is used, less number of non-dominated schedules are obtained for each project and it tends to find schedules with less *TSAD* while *CEH-II* tends to find schedules with smaller makespan values.

When we compared the results of the single and multi-project scheduling approaches, we saw that for most of the projects, single project scheduling approach gives better completion times. On the other hand, if we think all the projects as a composite project, the completion time of this composite project obtained with multi-project scheduling approaches approximately 5 months earlier using *CEH-I* and approximately 6 months earlier using *CEH-II*. Hence, if completing the composite project is more important than completing the projects individually, multi-project scheduling approach is better. On the other hand, a disadvantage of multi-project scheduling is that it re-schedules all the active activities with a new project initiation, so an activity is scheduled more than once even if there is no disruption affecting that activity. This re-scheduling increases system nervousness and demotivates the resources that work on the activities. An additional disadvantage is that the multi-project scheduling approach needs more CPU time than the single project scheduling approach.

5. Conclusion and future work

In this paper, we presented the proactive project scheduling phase of the three-phase approach developed for robust project scheduling. To the extent of our knowledge, this study is the first study considering multiple objectives on proactive project scheduling literature for the problem of the preemptive version of the RCMPSP with generalized precedence relations. To obtain robust baseline schedules, in the proactive project scheduling phase, we suggested two scheduling approaches each using a bi-objective GA with two different chromosome evaluation heuristics. Solution robustness is assured with *TSAD* minimization after a pre-specified number of schedule realizations are obtained for a chromosome. The other objective is the minimization of the makespan over all projects. The proactive project scheduling approaches are implemented on the real data from the R&D Department of a leading home appliances company in Turkey. Although we have used these two objectives, some other objectives could be used or added to the model as well. A further extension of our work could be considering the concepts of activity flexibility, project flexibility, activity priority and project priority while scheduling the projects.

References

Creemers, S., Demeulemeester, E., Van de Vonder, S. (2011, December). Project risk management: A new approach. *In Industrial Engineering and Engineering Management (IEEM), 2011 IEEE International Conference on* (pp. 839-843). IEEE.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. A. M. T. (2002). A fast and elitist multi-objective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, *6*(2), 182-197.

Herroelen, W. (2014). A Risk Integrated Methodology for Project Planning Under Uncertainty. In *Essays in Production, Project Planning and Scheduling* (pp. 203-217). Springer US.

Herroelen, W., Leus, R. (2004). Robust and reactive project scheduling: a review and classification of procedures. *International Journal of Production Research, 42(8),* 1599-1620.

Lambrechts, O., Demeulemeester, E., Herroelen, W. (2008). A tabu search procedure for developing robust predictive project schedules. *International Journal of Production Economics, 111(2*), 493-508.

Lambrechts, O., Demeulemeester, E., Herroelen, W. (2011). Time slack-based techniques for robust project scheduling subject to resource uncertainty. *Annals of Operations Research, 186(1)*, 443-464.

Leus, R. (2003). The generation of stable project plans. PhD thesis. Department of applied economics, Katholieke Universiteit Leuven, Belgium.

Schatteman, D., Herroelen, W., Van de Vonder, S., Boone, A. (2008). Methodology for integrated risk management and proactive scheduling of construction projects. *Journal of Construction Engineering and Management, 134(11)*, 885-893.

Van de Vonder, S., Demeulemeester, E., Herroelen, W. (2008). Proactive heuristic procedures for robust project scheduling: An experimental analysis. *European Journal of Operational Research*, *189*(3), 723-733.

Van de Vonder, S., Demeulemeester, E., Herroelen, W., Leus, R. (2006). The trade-off between stability and makespan in resource-constrained project scheduling. *International Journal of Production Research*, *44(2),* 215-236.