7th AIMMS–MOPTA Optimization Modeling Competition
Submit your solution by May 24th, 2015, 23:55 EDT

Project Portfolio Management

Large companies address grand challenges by completing a multitude of projects. For instance, developing an innovative new product may require completing dozens of projects, possibly spanning several phases. Large companies are likely to develop several new products at the same time, given the risk of delays or failures on each project, and the need to adapt products to an ever changing marketplace.

In the present case, the goal is to assist management with the selection of projects to pursue, over a three-year horizon, subject to risk constraints.

The projects under consideration have multiple dependencies:

- Some projects must be completed before others can start (e.g. a project has been decomposed into phases)
- Some projects cannot be carried out simultaneously (e.g. there are resource constraints on teams or equipment);
- Some projects can start conditionally to the completion of any project in a set (e.g. the set of projects represents independent research efforts on different alternatives to overcome a technical hurdle faced by another project).

1 Goal

The management is concerned with the selection and scheduling of projects, the expected rate of return of the portfolio of projects, and the risk embedded in the portfolio. The management needs a decision support system that allows the user to interact with the solution in order to perform what-if types of analysis, and compare different criteria.

In practice, the portfolio of projects will be rebalanced periodically; a three-year horizon is used to be able to schedule long-term projects and be able to feel the long-term interactions of some projects. Time will be discretized into monthly periods for the purpose of scheduling the projects.

A monthly contingency plan should exist in order to adapt or repair the schedule in case of project delays and failures. In this competition, we assume that a delay on a project postpones the start of all projects dependent on the completion of that project, and that failure triggers the abortion of all projects whose completion is made impossible by the failure[1].

2 Project Data

For each project (or phase, or task) j in a pool J of potential projects, data analysts have assigned a predicted cost of undertaking the project c_j (in millions of dollars), a predicted duration for the project d_j (in months), a predicted income b_j (in millions of dollars) if the project is completed successfully, a

[1]Typically, a better reaction will be computed by rebalancing the portfolio. The default contingency plan neglects this possibility and thus will lead to conservative estimates of the expected value and risk of the project portfolio.

probability of failure $p_{f,j}$, a probability that completion is delayed $p_{\delta,j}$ (given that the project did not fail yet), the estimated duration of the delay d_j^+ (in months) and the additional cost c_j^+ of this delay (in millions of dollars), and the updated probability of failure when the project is delayed $p_{f|\delta,j}$.

If project j starts at time $t_j \geq 0$, we will know whether j succeeds without delay, or fails without delay, or is delayed, at time $t_j + d_j$. If the project is delayed, we will know whether j ultimately succeeds or fails at time $t_j + d_j + d_j^+$.

The predicted income typically corresponds to the predicted net present value of a discounted stream of cash flows from the sales of a new product in a certain market during the estimated life of the product. For phased projects, often, only the project representing the final phase will be assigned a nonzero b_j.

In summary, each project $j \in J$ is represented by a tuple $(c_j, d_j, b_j, c_j^+, d_j^+, p_{f,j}, p_{\delta,j}, p_{f|\delta,j})$.

3 Project Interaction Data

There are 3 types of relations among projects that have been identified by the analysts and that should be taken into account: strict ordering, nonsimultaneity, and redundancy.

- Sometimes, some project j cannot start before project i is successfully completed. We write this as $i \succ j$. If project i starts at time t_i and is completed in d_i months, project j can start at time $t_j \geq t_i + d_i$. If the duration of project i is extended of d_i^+ months and then completes, project j can start at time $t_j \geq t_i + d_i + d_i^+$.

  Project ordering relations can be represented by a directed graph (V, E), where vertices $j \in V$ correspond to projects, and where an arc $e \in E$ from i to j exists if $i \succ j$.

  The problem data may contain directed cycles. A directed cycle in the graph (V, E) prevents the execution of all the projects in the cycle.

- Sometimes, some project i would use the same resources as some project j, and therefore the work on projects i and j cannot overlap. We write this as $i \not\sim j$.

  Nonsimultaneity constraints are represented by an undirected graph (V, E'), where an undirected edge $e' \in E'$ exists if $i \not\sim j$. That is, if both projects are selected, either the completion date of project i should be before the start date of project j, or the completion date of project j should be before the start date of project i. To summarize, $i \not\sim j$ is equivalent to ($i \succ j$ or $j \succ i$).

- Sometimes, project j cannot start before any of some projects $i_1, \ldots, i_N$ completes successfully. For instance, $i_1, \ldots, i_N$ can represent independent research efforts on different alternatives to overcome a technical challenge faced by project j. We write this as $\{i_1, \ldots, i_N\} \vdash j$.

  This information is represented as follows. Let $\mathscr{P}(V)$ represent the power set of V, that is, the collection of possible subsets of V. We define E'' as a collection of pairs $(I, j) \in \mathscr{P}(V) \times V$, where I is a set of projects of cardinality at least 2 and not containing j. As soon as one project $i \in I$ completes, j is allowed to start.

  Note that if singletons are allowed for I, then $\{i\} \vdash j$ is equivalent to $i \succ j$.

In summary, project dependencies are encoded by (V, E, E', E'').

4 Hurdle Rate and Discounted Cash Flow Valuation

We assume that investments are done to produce a rate of return of at least 10% yearly. This required rate of return is referred to as the minimal attractive rate of return (MARR) or the hurdle rate. Ensuring the hurdle rate is met can be done by valuing cash flows from investments using a monthly discount factor

$$\gamma = 1/(1 + 0.10)^{1/12} \simeq 0.99208894. \tag{1}$$

Projects producing discounted cash flows with a net present value equal or greater than 0 meet the hurdle rate.

For example, a single project with cost $c_j = 1$, income $b_j = 2$, duration $d_j = 3$, started at time $t_j = 6$ (in months after time 0), and successfully completed at $t_j + d_j$, is valued by the discounted cash flow approach at $v_j = -(1)\gamma^6 + (2)\gamma^9 \simeq 0.90856$, in millions of dollars at time 0. Since $v_j \geq 0$, the investment in this project satisfies the hurdle rate. Since $v_j > 0$, the investment in fact exceeds the hurdle rate.

As you can see, we suppose the project cost c_j is accrued at the time the project starts, while the project income b_j is accrued at the time the project successfully completes.

Additional costs due to delays are accrued at the time the delay occurs. Another example of valuation of a single project is given in Appendix B, that leads to a general formula for the net present value of a project $(c_j, d_j, b_j, c_j^+, d_j^+, p_{f,j}, p_{\delta,j}, p_{f|\delta,j})$ started in t_j months from now, valued using the discounted expected cash flow approach:

$$\begin{aligned} &v_j = \gamma^{t_j}(-\theta_1 c_j - \theta_2 c_j^+ + \theta_3 b_j), \qquad (2)\\ &\theta_1 := 1, \qquad \theta_2 := \gamma^{d_j}(1 - p_{f,j})p_{\delta,j}, \qquad \theta_3 := \gamma^{d_j}(1 - p_{f,j})[(1 - p_{\delta,j}) + \gamma^{d_j^+} p_{\delta,j}(1 - p_{f|\delta,j})]. \end{aligned}$$

You will still want to separate costs and incomes for the purpose of formulating optimization problems (Appendix B provides the details) and make them event-contingent, because formula (2) ignores dependencies among projects. To see this, observe for instance that if a project i fails, the projects j depending on the completion of i will be canceled. This type of cascading risk is *not* embedded in the choice of the hurdle rate. It is the role of project selection and scheduling to mitigate this type of risk, e.g. by selecting income-producing projects that can be completed through several combinations of successful cost-incurring projects.

5 A First Approach

A first approach to the problem is to "ignore" the random aspects, for the sake of simplicity. The project data is reduced to tuples $(\tilde{c}_j, \tilde{d}_j, \tilde{b}_j)$ for each $j \in J$, derived from the original tuples by some method the user chooses, among a set of methods you are asked to propose and motivate.

5.1 Deterministic Model

You are asked to formulate the maximization of the net present value of the project portfolio using the monthly discount factor γ, by selecting and scheduling the projects over the 3-year horizon, subject to the project dependency constraints. If $S \subset J$ denotes the selected projects and T is the associated schedule specifying the start-dates and end-dates of selected projects, the optimization problem can be stated as

$$\begin{array}{ll} \text{maximize} & NPV_\gamma(S,T) \\ \text{subject to} & i \succ j \quad \forall (i,j) \in E, \quad j \in S, \\ & i \not\sim j \quad \forall \{i,j\} \in E', \quad i,j \in S, \\ & I \vdash j \quad \forall (I,j) \in E'', \quad j \in S. \end{array} \tag{3}$$

The net present value of the portfolio is computed by summing up the costs and benefits of the selected projects, discounted appropriately at the time they are incurred, cf. Appendix B and (2). Thus we have

$$NPV_\gamma(S,T) = \sum_{\substack{j \in S \\ 0 \le t_j/12 < 3}} \gamma^{t_j} \tilde{c}_j + \sum_{\substack{j \in S \\ 0 \le (t_j + \tilde{d}_j)/12 < 3}} \gamma^{t_j + \tilde{d}_j} \tilde{b}_j \ . \tag{4}$$

5.2 Solution Analysis

You are asked to develop the AIMMS model and interface to let the user change the input data and explore what-if scenarios. Examples of typical questions a user might ask are: What is the value of the portfolio if the projects are delayed? What would have been the optimal portfolio had we knew the projects would be delayed? What is the value of the portfolio optimized for delayed projects if actually the projects are not delayed? What happens if I increase the hurdle rate? Etc.

In general, you should give the user several alternatives to create reduced project parameters $(\tilde{c}_j, \tilde{d}_j, \tilde{b}_j)$ from the project data, to be used as input data for the optimization, and then allow the user to post-evaluate the optimal solution, using the project parameters from any of the other alternatives that create reduced project parameters.

In the post-evaluation, the delay data may change. In this case, the schedule may have to be repaired or improved to remain compliant with the constraints $i \prec j$, $i \not\sim j$, $I \vdash j$. Changes of project duration should modify the start-date of the successor projects, but should not trigger deeper changes to the plan. It is good practice to try to formulate the post-evaluation as an auxiliary optimization problem. If project j is delayed, this causes an additional cost c_j^+ that should be discounted as appropriate.

6 Second Approach

A second approach is suggested to bring resilience to failures and delays to the project portfolio.

Let $W \ge 0$ be a parameter we call the Bad Luck Budget. Given any fixed schedule, Bad Luck will affect projects by making some of them fail or being delayed.

Let I_f be the set of projects that will fail by the effect of bad luck. Let I_δ the set of projects that will be delayed by the effect of bad luck but will ultimately be completed. Let $I_{f|\delta}$ be the set of projects that will fail by the effect of bad luck after having been delayed. Note that if $i \succ j$ and i fails by the effect of bad luck, i is in I_f, but j need not be in I_f, because j automatically fails when i fails.

We quantify[2] bad luck by

$$w(I_f, I_\delta, I_{f|\delta}) := \sum_{j \in I_f} [-\log_2(p_{f,j})] + \sum_{j \in I_\delta} [-\log_2((1-p_{f,j})p_{\delta,j})] \\ + \sum_{j \in I_{f|\delta}} [-\log_2((1-p_{f,j})p_{\delta,j}p_{f|\delta,j})]. \tag{5}$$

The net present value of the portfolio NPV'_ℓ is then defined as the net present value of the projects that have been selected and scheduled, and where subsequently projects in I_f fail, projects in I_δ and $I_{f|\delta}$ are delayed, and delayed projects in $I_{f|\delta}$ ultimately fail, all of this implying that dependent projects in the portfolio will in turn be be canceled or have their start-date delayed as appropriate. To clarify this, we define $I'_f(I)$ for $I \subset (I_f \cup I_{f|\delta})$ as the set of projects that fail as a consequence of the joint failure of the projects $j \in I$, and $I'_\delta(I)$ for $I \subset I_\delta$ as the set of projects that have their start-date delayed as a consequence of the joint delay of the projects $j \in I$. We write t'_j and d'_j as the updated start date and duration of the surviving projects $S' = S \setminus [(I_f \cup I_{f|\delta}) \cup \bigcup_{I \subset (I_f \cup I_{f|\delta})} I'_f(I)]$. We write T' for the collection of variables defining the updated schedule.

You are asked to formulate and develop a solution approach to the maximin problem

$$\begin{array}{lll} \max\limits_{S,T} & \min\limits_{I_f, I_\delta, I_{f|\delta}} & NPV'_\gamma(S, T, I_f, I_\delta, I_{f|\delta}) \\ \text{subject to} & i \succ j \quad \forall (i,j) \in E, & \\ & i \not\sim j \quad \forall \{i,j\} \in E', & \\ & I \vdash j \quad \forall (I,j) \in E'', & \\ & w(I_f, I_\delta, I_{f|\delta}) \leq W. & \end{array} \tag{6}$$

You are asked to add to the AIMMS interface the ability to let the user change the bad luck budget W, compare various solutions, and analyze the results. Ideally, the user should be able to visualize the optimal post-bad-luck net present value NVP'_γ as a function of the bad luck budget $W \geq 0$. The user should be able to display and dig into particular optimal solutions (S, T) and (S', T').

As a first step, you may consider developing a solution method for the case where there is no random addition of delays. The inner minimization would thus be over I_f only.

7 Deliverables

You are free to browse and use the literature for inspiration. Please cite all sources and carefully distinguish your ideas from those obtained in the literature.

Your team needs to deliver a solution to the problem described in this case study, including

[2] Our definition of bad luck is related to Claude Shannon's concept of information: the failure of a "safe" project is more "surprising" than the failure of a "risky" project. In particular, it takes much more bad luck to make a project fail if its failure probability is small. Bad luck is infinite if the project's failure probability is 0. Bad luck is zero if the project's failure probability is 1. The initial delays and failures due to bad luck are assumed to be independent. Cascading delays and failures due to project dependencies are not counted against the bad luck budget.

- implementation of your models in AIMMS, including a user interface, providing the user graphical and textual output;
- solutions of the models for the given data sets, as well as those for your data sets;
- a 10–15 page report (max 15 pages) that discusses your models, the mathematical background of your techniques, the solutions that you obtained, and further recommendations.

Teams that only develop partial solutions to the case are still encouraged to submit their solution. Teams are also encouraged to address additional or alternative "realistic" factors, variants, or modifications of the underlying problem considered in the case. If data beyond the one provided with the case is necessary to consider these factors, variants, or modifications, you are asked to generate additional corresponding data either from relevant literature or sensible assumptions. Be sure to reference the literature or assumptions you used to generate the additional data.

The deadline for submission is May 24th, 2015, 23:55 EDT. If you have questions about the problem or the competition in general, please contact Boris Defourny at `defourny@lehigh.edu`. The subject of your email should start with "MOPTA:" (without quotes).

Appendix A Data Sets

We are providing 3 data sets. The two first data set are designed to test your formulations and are given below. The third data set is larger. It will be released on Feb 15, 2015.
In all data sets, costs and revenues are in millions of dollars, and durations are in months.

A.1 Data set 1

Project data:

j	1	2	3	4	5	6	7	8	9	10
c_j	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
d_j	12	12	10	9	12	12	12	12	12	12
b_j	0	0	0	0	0	0	0	0	10	20
$p_{f,j}$	0.1	0.1	0.3	0.3	0.3	0.1	0.1	0.01	0.01	0.1

Project interactions data:

$1 \succ 10,\quad 2 \succ 10,\quad 8 \succ 9,$
$1 \not\succ 3,\quad 1 \not\succ 8,\quad 3 \not\succ 4,\quad 10 \not\succ 8,$
$\{3,4,5\} \vdash 10,\quad \{6,7\} \vdash 10.$

A.2 Data set 2

Same as data set 1 except that now, the projects may also be affected by delays.

Additional project data:

j	1	2	3	4	5	6	7	8	9	10
c_j^+	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2
d_j^+	2	2	4	6	1	0	0	0	0	0
$p_{\delta,j}$	0.1	0.1	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1
$p_{f\|\delta,j}$	0.15	0.15	0.3	0.3	0.3	0.2	0.1	0.01	0.01	0.1

Appendix B Isolated Single Project Valuation

Suppose $\gamma = 1/(1+0.10)^{1/12} \simeq 0.9921$ to meet a yearly 10% rate of return. A project that has a γ-discounted nonegative value (benefits minus costs, discounted at the moment they occur) meets or exceeds this target.

Example. A 6-month project scheduled to be started in 3 months from now that costs 1 and has benefit 1.5 is worth $\gamma^6(1.5) - 1 = 0.4302$ in 3 months, and $\gamma^3 0.4302 = 0.4201$ now. If a 3-month delay occurs with additional cost 0.25, the project is worth $\gamma^9(1.5) - \gamma^6(0.25) - 1 = 0.1582$ in 3 months, and $\gamma^3 0.1582 = 0.1544$ now. If the project fails before it incurs the delay, it is worth $\gamma^3(-1) = -0.9765$ now, that is, a net loss. If the project fails after it incurs the delay, it is worth $\gamma^3(-1 - \gamma^6(0.25)) = -1.2092$ now.

Now, if the probability of failure is 0.1, the probability of incurring a delay (assuming no failure yet) is 0.3, and the updated probability of failure given the delay is 0.15, the expected value of this project, ignoring all interactions with other projects, is

$$v = 0.1(-0.9765) + 0.9[0.7(0.4201) + 0.3[0.85(0.1544) + 0.15(-1.2092)]] = 0.1535.$$

The condition $v \geq 0$ indicates that this project meets the 10% rate of return condition in expectation. Note that instead of a start-date in 3 months, the start-date is pushed back 2 months, the expected value of the project becomes $v' = \gamma^2 0.1535 = 0.1510$.

In the worst-case scenario, undertaking this project leads to a loss of 1.2092, without counting the loss of value from dependent projects that should be canceled. A delayed success could scale down the values of all dependent projects by a factor γ^3 due to the 3-month delay of their start-dates.

General case. If t denotes the scheduled start-date of the project (in number of months from now), the expected value formula in terms of the project parameters $(c, d, b, c^+, d^+, p_f, p_\delta, p_{f|\delta})$ is

$$v = p_f v_f + (1 - p_f)[(1 - p_\delta)v_{ok} + p_\delta[(1 - p_{f|\delta})v_{ok|\delta} + p_{f|\delta} v_{f|\delta}]],$$

where

$$v_{ok} = -\gamma^t c + \gamma^{t+d} b, \quad v_f = -\gamma^t c, \quad v_{ok|\delta} = -\gamma^t c - \gamma^{t+d} c^+ + \gamma^{t+d+d^+} b, \quad v_{f|\delta} = -\gamma^t c - \gamma^{t+d} c^+.$$

Equivalently, $v = \gamma^t(-\theta_1 c - \theta_2 c^+ + \theta_3 b)$, where

$$\theta_1 = 1, \qquad \theta_2 = \gamma^d (1 - p_f) p_\delta, \qquad \theta_3 = \gamma^d (1 - p_f)[(1 - p_\delta) + \gamma^{d^+} p_\delta (1 - p_{f|\delta})].$$