Genetic Algorithm of Network Graph Multi-Objective Optimization as an Instrument of Project Monitoring

Oleg Prokopyev

1 Kazan Federal University, Institute of Economics and Finance, 420012, Kazan, Russia

Correspondence: Oleg Prokopyev, Kazan Federal University, Institute of Economics and Finance, 420012, Kazan, Russia.

Received: July 8, 2014   Accepted: October 8, 2014   Online Published: December 2, 2014
doi:10.5539/ass.v11n1p204          URL: http://dx.doi.org/10.5539/ass.v11n1p204

Abstract
Proper tracking of progress remains a vital part of modern project management, defining prospects of successful implementation of planned tasks. There are several popular concepts of project monitoring, such as logical framework approach (LFA), earned value management (EVM), etc., and each of them depends on properly optimized network graph that represents dependences between project tasks. Article describes the features and problems of multi-objective optimization in project management with reference to network graphs. The significant role of network graph optimization for project monitoring systems is proved and the model of multi-objective optimization of the network graph on criterion functions of duration and project cost based on NSGA-II genetic algorithm is proposed as the main purpose of research. Model takes into account the reserves of time on a critical way of the network graph, possibility of decreasing the load of available resources at the expense of time reserves on non-critical ways of the network graph, variety of used resources and options of delegation. One of its main advantages is quite low laboriousness of implementation, that depends on number of nodes on the network graph of the project and on number of possible options of delegation for the project tasks with several alternatives of delegation. Model has been tested on sample project with real data and results have been analyzed.

Keywords: network graphs, project monitoring, project scheduling, multi-objective optimization, genetic algorithms

1. Introduction
Successful implementation of any project is impossible without the rational organization of tracking the progress in the field of achieving its common goals through certain tasks and task groups. In modern conditions traditional practices of project control based on the results of phases and stages aren't acceptable anymore as the dynamic project environment and variety of risk factors can lead to situations when the negative deviations not revealed in time on various parameters of the project will cause substantial effects even before the next planned point of control is approached (Hardy-Vallee, 2012). Bad progress tracking causes about 25 % of project fails (McDonald, 2012), another 20-25 % do not correspond to expectations regarding ROI (Gartner, 2013). 67 % of projects were out the budget mainly for this reason (KPMG, 2013). In this regard, priority value in project management at a present stage of its development is received by the function of project monitoring (Conner, 2012).

Project monitoring can be defined as independent special function of project management that includes dynamic process of collecting, processing and transferring information based on certain algorithms and concrete systems of the indicators reflecting current state of the project and promoting the process of decision-making, involving methods and tools depending on key technical and economic characteristics of the project. Monitoring in project management established as an independent concept approximately in the 1970-th (Kerzner, 2010), however only by the end of XX century it began to be considered as one of the management functions, not identical to control, but complementary in relation to it (Newton, 2013 Structure and composition of project management functions and the organization of project tracking was touched by K. Lockyer, J. Gordon, G. Diethelm, C. Gray, E. Larsson, J.P. Lewis, H. Reschke, R. Newton, M. Trotsky, B. Grucha, K. Ogonek, D.I. Cleland and so forth. Common issues of monitoring and evaluation planning were also covered by S.G. Chaplows, J. Fuller, J.G. Gerald, J.A. Hudes, J.M. Lee, J.J. O’Brien, H. Schelle.
The most significant issue that should be resolved during planning of project monitoring is providing the due dependence between monitoring and schedules being prepared. The network graph plays an important role among these schedules (Parveen & Saha, 2012). The optimized network graph can be used subsequently as an evident source of standard values regarding terms and cost of realization of separate operations and packages of works within the project (Wang et al., 2014). Solving this problem requires simultaneous time-cost optimization.

Heuristic principles of network graph optimization were researched by E.W. David, J.H. Patterson, F. Boctor. Probability distribution analysis in network graph optimization was covered by works of H.M. Ahuja, A.B. Badiru, E.M. Padilla, R.I. Carr. Modern methods and algorithms of such optimization is rather popular scope of researches in project management. These topics were touched upon by M. Bagherpour, S. Cheung, R. Babuska, S. Colak, A. Agarwal, I. Gujit, K. Janet, K. Smith, J.J. Moder, etc.

Proceeding from the aforesaid, the purpose of this research is development and approbation of multi-objective optimization model for the schedule of project tasks, providing minimization of total duration and total cost of project.

2. Materials and Methods

The model of multi-objective optimization offered below belongs to works acceleration models group, thus, base preconditions of its application are:

- Lack of reserves of time on a critical way of the network graph;
- Possibility of decreasing the load of available resources at the expense of time reserves on non-critical ways of the network graph;
- Variety of types of used resources and their low interchangeability;
- Various options of delegation on the various tasks of the project, not assuming strict functional dependence between possible values of duration for these works and cost of their realization.

There are two approaches that can be used for the solution of such problems (Schwefel, 1981). The first approach assumes transformation of multi-objective optimization problem to linear on non-linear programming (depending on specifics of structure of project expenses). The second approach is based on application of the techniques relating to multi-objective optimization, in particular, to the Pareto front or a lexicographic order by means of genetic algorithms, Gembicki’s method and so forth.

We carry out the comparative analysis of both approaches, having considered the general mathematical statement of the designated task. This task belongs to the class of the determined tasks (Steuer, 1986) with a large number of possible decisions.

Let’s assume that we deal with the project consisting of m works, each of them can be executed by one of n executors. Total expenses on each work consist from direct and overhead costs. Direct costs ($d_i$), as well as the duration of $i$-th work ($l_{ij}$) depend on executor chosen. Overhead costs are allocated to concrete operations and don’t correlate with chosen scheme of task delegation. Overheads that cannot be reckoned towards certain task can be considered in model as additional summand in criterion function (B). For ensuring compatibility with a wide range of practical tasks considered model contains also of a limit of project duration (L), i.e. desirable duration of the project for its key interseants, and the time horizon (T), i.e. the maximum theoretical duration of the project. Separation of L and T allows to consider size of a penalty or the missed benefit (P) which pays off as a rate of a penalty or alternative costs ($\mu$), multiplied by size of excess of the actual duration of the project over a limit, in the general expenses of the project (M). If any excess of a limit of duration is inadmissible, the horizon of T will be equal to a limit of L, and the model is simplified.

Dependences between works of the network graph and delegation of these works to certain executors are considered by means of introduction of Boolean variables in sets of E and C. The $E_{ijt}$ variable turns to one if at the t-th moment j-th executor is involved in i-th work. The $C_{ijt}$ variable is equal to 1 only if at the t-th moment j-th executor has finished i-th work. In this case the moment of completion for the i-th task by j-th executor may be defined as:

$$F_{ij} = \sum_{t=1}^{T} \sum_{j=1}^{n} t \cdot C_{ijt}, \quad \forall i, j$$

Thus, criterion of an optimality of considered model is minimization of the total project costs:

$$M = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{t=1}^{T} d_j \cdot E_{ijt} + \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{t=1}^{T} b_j \cdot C_{ijt} + B + P, \quad M \to \min$$

(2)
In case if specifics of the project does not allow to allocate an overhead costs for concrete operations, criterion function can be brought to the following look:

\[ M = \sum_{i,j} \sum_{t} d_i \cdot E_{i,j} + B + P, \quad M \rightarrow \min \] (3)

These are the constraints of offered model:

a) each work is delegated to one executor:

\[ \sum_{j=1}^{k} E_{i,j} + \sum_{t=1}^{n} E_{i,t} \leq 1, \quad \forall j,t \] (4)

b) each executor can be occupied with at most one work during every moment of time:

\[ \sum_{j=1}^{k} E_{i,j} \leq 1, \quad \forall j,t \] (5)

c) as the purpose of model is optimization of the planned schedule, it is not supposed that time of performance of operation will exceed the set limits. Therefore, i-th work will be carried out by the j-th executor within time interval \([t; t+\ell_{ij}-1]\);

d) each work can be begun only after all dependences following from the network graph of works are satisfied:

\[ \sum_{j=1}^{k} F_{i,j} \geq \sum_{h=1}^{n} F_{h,j} + \sum_{l=1}^{m} l_{i,j} \cdot C_{i,j}, \quad \forall h,i \in \Phi(i) \] (6)

e) each work of the project should be surely performed by one of executors.

This model allows to receive the optimum schedule of the project based on the most effective delegation of operations. Model will require these given data (Safiullin & Prokopyev, 2011):

- register of project tasks, containing data on interrelations between them, or the network graph of the project;
- data on an overheads for certain works (and for the whole project);
- data on a direct costs on works and executors;
- estimates of duration of each project task by each of executors this work can be delegated to;
- limit of project duration, the horizon of planning and amount of penalty for project failure or estimated size of the project foregone gain.

These are the key restrictions of that model:

- allocating fixed and variable parts in total expected expenses is possible not for every project task (Yang, 2007);
- exact calculation of the missed benefit, as a rule, is an exigent task;
- using the restriction demanding delegation to one executor of only one work during the concrete moment of time is not always correct, for example, in case of outsourcing, since the companies that are ready to carry out some of project tasks may be engaged in several works simultaneously (Alireza et al., 2014);
- many projects are characterised by lack of linear dependence between time and project cost, and exact expression of criterion function showing the project duration of the project through criterion function of project costs is impossible.

Thus, in most cases instead of described above model use of the second approach is necessary. This requires reconception of the task with two criterion functions – project costs and duration, also taking into account peculiarities designated above. First of all, we will introduce the \(A_{i,j}\) variables showing total expenses to i-th work in case of its delegation to the j-th performer. Amount of sets of Boolean variables will be reduced to one set, including the variables showing the fact of delegation of i-th work to the j-th executor (we will designate them as \(G_{i,j}\)). Taking into account that options of delegation can be provided not for all works, and total expenses on other works will correspond to B indicator in the initial version of model, criterion function of the total project costs can be written down as follows:

\[ M = \sum_{i,j} A_{i,j} \cdot G_{i,j} + B, \quad M \rightarrow \min \] (7)

Respectively, the second criterion function (duration of the project) may be simplified to:
\[ N = \sum_{i=1}^{m} \sum_{j=1}^{n} l_{ij} \cdot G_{ij} + R + S, \quad N \to \min \]  

where:  
R – change of project duration caused by change of duration of tasks belonging to separate contours of the network graph;  
S – total duration of works in contours of the network graph of the project that doesn't contain works with various options of delegation.

The modified system of restrictions of model is reduced to one expression concerning the requirement of performance of each task by someone from possible executors (this means that sum of \( G_{ij} \) variable must be equal to 1).

Nevertheless, despite visual simplification of statement of a task its solution is rather laborious, i.e., because of the necessity of taking into account dependences of works and events on the network graph (Ashuri, 2012). Unlike criterion function of project costs which is exclusively additive, function of project duration combines duration of tasks forming each of closed contours on the network graph, therefore, representation of this function for each project will be different. Application of the offered multi-objective optimization model on a practical example will be described below.

### 3. Results

Model will be tested on sample project including 15 events and 17 works. Table 1 shows project tasks assuming various options of delegation. Initial version of delegation of these works to executors is shown with colour. The network graph of the project constructed according to initial delegation of tasks is represented on fig. 1. For each of options of delegation duration of performing i-th work by the j-th executor (\( l_{ij} \)) and the expenses (\( A_{ij} \)) are specified.

<table>
<thead>
<tr>
<th>ID</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variable</td>
<td>Duration (days)</td>
<td>Cost (ref. units)</td>
</tr>
<tr>
<td>0-3</td>
<td>x(1)</td>
<td>9</td>
<td>140</td>
</tr>
<tr>
<td>1-5</td>
<td>x(4)</td>
<td>7</td>
<td>260</td>
</tr>
<tr>
<td>4-7</td>
<td>x(7)</td>
<td>6</td>
<td>310</td>
</tr>
<tr>
<td>6-8</td>
<td>x(9)</td>
<td>8</td>
<td>190</td>
</tr>
<tr>
<td>9-12</td>
<td>x(12)</td>
<td>14</td>
<td>320</td>
</tr>
<tr>
<td>11-13</td>
<td>x(14)</td>
<td>10</td>
<td>300</td>
</tr>
</tbody>
</table>

The solution of a problem of multi-objective optimization was carried out in a MATLAB R2013a with use of an additional plug-in, the modified genetic algorithm of NSGA-II. As well as in a case with other large problems of optimization, the most convenient way of calling the function based on NSGA-II algorithm is using m-files that describe criterion functions, restrictions and optimization parameters.

![Initial version of project network graph](image_url)
The quantity of changeable variables \((G_{ij})\) is equal to 16 because there are 16 filled cells in the matrix shown in table #1. Optimization with use of genetic algorithms is an iterative process where a sample of genotypes (population) is formed and modified on each run of the algorithm (generation). The algorithm of NSGA-II provides adjustment of crossing and mutation procedures through the two special parameters: options.crossover and options.mutation. Options of crossing include method of crossover and factor of similarity \([\rho]\) that influence crossing process as follows:

\[
\begin{align*}
C_1 &= X + \rho \cdot r \cdot (Y - X) \\
C_2 &= Y + \rho \cdot r \cdot (Y - X)
\end{align*}
\]

where: 
\(X\) – the first parental genotype,
\(Y\) – second parental genotype,
\(r\) – a random number in a range \([0; 1]\),
\(C_1\) – the first generated genotype,
\(C_2\) – the second generated genotype.

The mutation is the second characteristic operator for the genetic algorithms, generating the modified genotypes. As well as in a case with a biological mutation, some mutating individuals in the population periodically are essentially different on testimonials from average values in this generation. NSGA-II is designed to use Gaussian mutations which is described by following function:

\[
MI = C + \psi \cdot \left(1 - \zeta \cdot \frac{CG}{LG}\right) \cdot r \cdot (UB - LB),
\]

where: 
\(MI\) – mutating genotype,
\(C\) – an initial condition of a normal genotype,
\(\psi\) – factor of deviation,
\(\zeta\) – factor of convergence,
\(CG\) – serial number of the current generation,
\(LG\) – total number of generations,
\(r\) – a random number in a range \([0; 1]\),
\(UB\) – maximum size for values of a genotype,
\(LB\) – minimum size for values of a genotype.

The factor of deviation defines the degree of mutagen deviations and can be any positive number. The factor of a convergence regulates intensity of mutations in the process of transition from one generation to another. It can turn to any values in a range \([0; 1]\).

The second m-file for the proposed sample project should contain two vectors: a vector of criterion functions \((y)\) and a vector of restrictions \((cons)\). First criterion function shows total project costs, including the sum of products of the \(A_{ij}\) variables on corresponding Boolean \(G_{ij}\) variables, added with total expenses for other tasks of the project that are not assuming options of delegation. The second criterion function of the project is the duration. It should be written down not directly (otherwise the code would be too spacious), but through the auxiliary variables designating nodes of the network graph corresponding to the 7th, 8th, 12th and 14th events. They limit those contours containing tasks with various options of delegation, therefore, only these contours will influence total project duration. They are followed by 16 blocks of model restrictions.

It is known that the solution of multi-objective optimization problems with contradicting criterion functions seldom provides one optimum decision. Much more often the result is the set of the similar decisions on the Pareto front (Zheng, 2004). The points on the front of Pareto designate border of area of optimum decisions; each of them corresponds to a situation when further improvement of the decision according to first criterion function will lead to deterioration according to second criterion function (Ozkan-Deniz et al., 2012).

Testing of the offered model with the different sizes of the population, different number of generations, and also different parameters of crossing and a mutation has shown that the total decision is rather stable: various combinations of parameters sooner or later lead to it. The condition of the Pareto front at the time of completion is shown on fig. 2. Pareto front is formed with 4 optimum decisions: 73 days and 3.71 million ref. units, 74 days and 3.68 million ref. units, 79 days and 3.625 million ref. units, and also 80 days and 3.595
million ref. units. In comparison with the initial version of delegation, providing duration of the project of 76 days and the total cost of 3.63 million, this alternative is concerned with duration of the project of 74 days and is likely to become the most preferable to the initiator of the project in this case. The network graph of the project corresponding to this way of delegation is represented on fig. 3, and the description of the final version of delegation is shown in table 2.

Figure 2. Pareto front at the last generation

Table 2. Final version of project tasks delegation

<table>
<thead>
<tr>
<th>ID</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Duration (days)</td>
<td>Cost (ref. units)</td>
<td>Duration (days)</td>
</tr>
<tr>
<td>0-3</td>
<td>x(1)</td>
<td>9</td>
<td>x(2)</td>
</tr>
<tr>
<td>1-7</td>
<td>x(4)</td>
<td>7</td>
<td>x(5)</td>
</tr>
<tr>
<td>4-7</td>
<td>x(7)</td>
<td>6</td>
<td>x(8)</td>
</tr>
<tr>
<td>6-8</td>
<td>x(9)</td>
<td>8</td>
<td>x(10)</td>
</tr>
<tr>
<td>9-12</td>
<td>x(12)</td>
<td>14</td>
<td>x(13)</td>
</tr>
<tr>
<td>11-13</td>
<td>x(14)</td>
<td>10</td>
<td>x(15)</td>
</tr>
</tbody>
</table>

Figure 3. Final version of the network graph

Thus it is necessary to emphasize that within this sample project all of the four optimum decisions found are equally preferable, however dependence between duration of the project and its total costs has non-linear
character: the economy of three days (decrease from 79 to 76 days) costs only 5 thousand of reference units, and the economy of two more additional days (decrease from 76 to 74) costs 10 times more. Besides economy of time or resources, application of multi-objective optimization allowed to reduce number of the critical tasks requiring priority monitoring.

4. Discussion

The analysis of scientific literature in recent years shows considerable interest to application of various approaches, techniques and models in optimization of projects. Such authors as S. Parveen, S.K. Saha, D. Heon Jun, K. El-Rayes, W. Wang, G. Wang, A. Azaron, J. Rada, R. Parma, W. Pereira, relying on works by R. Steuer, H.P. Schwefel, P. Brucker, Z. Michalewicz investigate various versions of algorithms of multi-objective optimization for the solution of the applied tasks concerned with schedule and resource planning. Nevertheless, the insufficient attention is paid to questions of multi-objective optimization in their interrelation with the network graph; for example, works of the project are often considered as the consecutive list though similar simplification of statement of a task can lead to incorrect results in case of existence of complex dependences between project tasks.

S. Parveen and S. K. Saha researched the techniques of solving time-cost trade-off problem with genetic algorithms based on modified adaptive weights approach, but determining proper values for these weights is rather complex issue (Parveen & Saha, 2012). Project stakeholders often can’t prefer any of criterion functions, therefore, using that model with equal weights is just unwarranted.

Studies by D. Heon Jun and K. El-Rayes are devoted to multi-objective optimization concerning resource leveling. They propose some additional metrics, i.e., Release and Re-Hire (RRH), Resource Idle Days (RID), Maximum Resource Demand (MRD). Combinations of these metrics are used as the criterion functions in multi-objective optimization models with percentage weights allowing to set the priority of certain metric. D. Heon Jun carries out the comparative analysis of such genetic algorithms as Strength Pareto Evolutionary Algorithm (SPEA), Pareto-Archived Evolutionary Strategy (PAES) and Non-dominated Sorted Genetic Algorithms (NSGA-II), concluding that NSGA-II is the most robust multi-objective algorithm for network graph optimization models (Jun & El-Rayes, 2011). Then NSGA-II is implemented into models of resource allocation and leveling designed for construction projects.

W. Wang and X. Wang also developed multi-objective optimization model based on critical chain and network graphs. It takes into account overall duration, project costs and network graph robustness and allows creation of several alternative schedules (Wang et al., 2014). They are ranged by priority by using cloud genetic algorithms. That kind of optimization models is intended for strictly resource-constrained projects.

A. Azaron, H. Katagiri, M. Sakawa, K. Kato and A. Memariani used multi-objective optimization for models with up to four conflicting criterion functions (total project costs, the mean of project completion time, the variance of project completion time and the probability that the project completion time doesn’t exceed maximum allowed value). Although proposed algorithm of optimization is rather complex, it’s designed for certain category of projects. It may be implemented in PERT networks with exponentially or Erlang-distributed durations (Azaron et al., 2006).

Alternative approach to path optimization problems is proposed by J. Rada, R. Parma and W. Pereira. It assumes multi-objective optimization in the context of principles of Animat game based on SPEA-II genetic algorithm. Further elaboration on this problem and combining proposed model with arrays of Boolean variables may lead to new class of network graph optimization models that provide additional flexibility and robustness (Rada et al., 2008).

A. Senouci and H. Al-Derham developed a model that generates not only resource utilization plans that optimize construction time and cost but also visualizing the trade-offs among project time and cost in order to support decision makers in evaluating the impact of various resource utilization plans. Flexible genetic algorithm, based on parameters of string size, number of generations, population size, mutation and crossover rates, is also proposed (Senouci & Al-Derham, 2008).

S. Leu and T. Hung propose optimal resource-constrained scheduling simulation model aimed for construction scheduling. It takes into account both duration and resource constraints. Uncertainties are simulated through certain probability distribution model. Probabilistic optimal project duration is calculated under resource constraints by using UX3 и UM3 genetic algorithms. These algorithms allow adjusting crossover and mutation ratios depending on project peculiarities (Leu & Hung, 2002).
The main limitation of model suggested in this article is difficult process of defining the criterion function of maximum project duration. This issue may be resolved by involving more techniques from graph theory and decreasing the number of Boolean variables used in the criterion function. Model also needs additional testing on other configurations of network graphs and other types of genetic algorithms.

5. Conclusion

Proposed model of multi-objective optimization of the network graph allows solving the following issues:

- The most rational delegation of project tasks to various performers;
- Formation of the list of optimum decisions which can be presented to the client or to the investor of the project as possible alternatives of project implementation;
- Specification of standard values for the subsequent monitoring of project progress in aspects of time and cost.

Using the offered model is expedient after application of simpler optimization methods (for example, using reserves, alignment of resources, paralleling). Laboriousness of model implementation is rather low and depends on number of nodes on the network graph of the project and on number of possible options of delegation for the project tasks assuming such possibilities. Automation of model application by means of the modified algorithm of NSGA-II in the environment of MATLAB provides not only the fast solution of multi-objective optimization problem, but also receiving evident results in the form of matrices of recommended task delegation and Pareto front that generalize all found optimum decisions.

Monitoring of the project progress based on the application of proposed model will be more efficient in aspects of economy of resources by the means of modifying structure of critical paths on the network graph and increase of accuracy in interpretation of deviations from optimum values of time and cost indicators.

Further improvements of proposed model may be concerned with criterion function of summary project duration that may be represented in less complex form and with improvement of accuracy of genetic algorithms used for multi-objective optimization.

Acknowledgements

Author acknowledges Azat Safiullin and Marina Myznikova (Kazan Federal University) for inestimable help and encouragement in researches. Author solely takes responsibility for all of possible errors and inaccuracies in formulae.

References


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).