A Multicriteria Approach to Project Portfolio Selection

Using Multiobjective Optimization and Analytic Hierarchy Process

Everton Gomede and Rodolfo M. Barros
Computer Science Department
State University of Londrina
Londrina, Brazil
evertongomede@gmail.com, rodolfo@uel.br

Abstract—This paper presents an approach to solve project portfolio selection problem (PPSP) in the presence of limited resources, multiples criteria, software projects, constraints, functions to be optimized, interdependent projects, and scenarios with a large number of projects available. For this purpose, it is divided into two phases, one for (i) optimization using the multiobjective algorithm NSGA-II and another (ii) post-optimization using the Analytic Hierarchy Process (AHP). Among their contributions, we can name (i) a solution to the combinatorial analysis \(2^n\) and (ii) the structure of a hierarchy of criteria derived from subjective aspects.

Keywords- Software Project Portfolio Selection, Software Engineering Decision Support, NSGA-II;

I. INTRODUCTION

Project Portfolio Management (PPM) is an important issue for organizations that want to establish a process of selection and prioritization of projects focused on alignment to the corporate strategies. This means managing the set of programs and/or projects as a systemic whole, and enables appropriate allocation of resources, whether financial, human and technological, enabling an integrated investment management [1].

The PPM includes activities such as identification, evaluation, selection, prioritization, balancing, among others, whereas support the consistency of the strategy aligned to the organizational vision, mission, and values [1]. This relation is shown in Fig. 1.

Figure 1. An organizational context of portfolio management [1].

1 A Project Portfolio (PP) is a collection of programs and/or projects managed group with intention to achieve one or more strategic objectives [1].

The Project Portfolio Selection Problem (PPSP) aims to choose a set of programs and/or projects, considering not only the constraints and characteristics of each one, but also the relations among them, optimizing one or more objectives [21]. Considering the process shown in Fig. 2, the selection activity is contained in stage “Selection”, where the alternatives are evaluated and chosen [1].

Figure 2. A process to systematic management of a project portfolio [1].

To select the projects that should be implemented by an organization is the main component of a PPM [2]. Levine [3] states that many companies strive to make their projects succeed well, not knowing, if these are the right projects to be executed. Another aspect addressed by the author is the fact that companies take excessive risks with projects or continue running projects that probably will not reach their goals. Thus, valuable resources are being expended without needing instead of being targeted for more interesting projects for organizations.

This paper presents an approach to solve PPSP consisting of two phases: (i) optimizes two objective functions (risk and return) subject to a set of constraints generating a set of optimal portfolios, and (ii) a hierarchy of qualitative criteria that captures difficult mapping mathematics information [11, 12] allowing a single project portfolio to be selected.
II. RESEARCH PROBLEM & CONTRIBUTIONS

The research question that guided this work can be described as follows: “How to select an optimal project portfolio considering quantitative and qualitative criteria in a scenario with multiple objectives, multiple constraints, and a high number of possible combinations?”. This issue has generated some contributions that can be classified as:

- **Mathematical description of problem**: encode the problem in terms of their objective functions and constraints to solve combinatorial analysis $2^n$.

- **Hierarchy of criteria for project portfolio selection**: a structure with the criteria commonly present in the context of selection of project portfolios.

III. FUNDAMENTALS & RELATeD WORKs

**A. Multiobjective Optimization (MOO)**

Problems with multiple objectives arise in a natural fashion in most disciplines and their solution has been a challenge to researchers for a long time. Despite the considerable variety of techniques developed in Operations Research (OR) and other disciplines to tackle these problems, the complexities of their solution calls for alternative approaches [4].

The use of Evolutionary Algorithms (EAs) to solve problems of this nature has been motivated mainly because of the population-based nature of EAs which allows the generation of several elements of the Pareto optimal set in a single run. Additionally, the complexity of some Multiobjective Optimization Problems (MOPs) (e.g., very large search spaces, uncertainty, noise, disjoint Pareto curves, etc.) may prevent use (or application) of traditional OR MOP-solution techniques [4].

**B. Nondominated Sorting Genetic Algorithm II (NSGA-II)**

The NSGA-II Multiobjective Evolutionary Algorithm is based on a classification of hierarchical dominance frontiers. This method is employed an elitist strategy of reinsertion in population, to ensure which any solution of the hierarchy is found in any generation, it will be retained until the final population [5]. The NSGA-II works with a parent population $P$ to generate an offspring population $Q$ similar to conventional EAs. In the first iteration, generates a population $P_0$ that is subjected to the classification of dominance. Each solution has a fitness value equal to the level of their frontier. Using the operators: selection by tournament, crossover, and mutation, obtained an offspring population $Q_0$, $P_0$ and $Q_0$ are size $N$. Both populations, $P_0$ and $Q_0$, are join in a total population $R_0 = 2N$.

For subsequent generations ($t = 1, 2, ...$), the algorithm works with the total population $R_t$. Every generation is classified in hierarchical frontier of dominance, obtained frontiers $F_1, F_2, ...$, where $F_1$ is the first frontier, with all non-dominated solutions of the current $R_t$. The reintegration of the total population $R_t$, into a new population $P_{t+1}$ of parents, is made in order to select the $N$ solutions of $R_t$ that are at a higher level of dominance. Thus, the formation of $P_{t+1}$ starts with solutions $F_1$ followed by solutions $F_2$ and so forth.

Each $F_t$ set must be inserted in $P_{t+1}$ while $P_{t+1} + |F_t| \leq N$. Inserting the solutions of a frontier $F_t$ such as $|F_t| > N - P_{t+1}$, the algorithm chooses the solutions of $F_t$ that are better spread, i.e., the reinsertion of population from one generation to another is made considering the best individuals among parents and offspring. These individuals are classified into dominance frontiers and new population is formed selecting individuals of the first frontier until population size is reached. This measurement is given by the crowd distance [5]. Fig. 3 illustrates the population scheme of NSGA-II.

![Figure 3. The population scheme of NSGA-II algorithm [5].](image)

Once obtained crowd distances, the sets $F_t$ are ordered decreasingly by their distances. Finally, it generates $|Q_{t+1}|$ from $|P_{t+1}|$ using the operators of tournament selection by crowd, crossover, and mutation [5].

The multiobjective selection in NSGA-II is performed by tournament crowd. The NSGA-II incorporates a small modification in the tournament selection method (crowd tournament) [5]. A solution $i$ is considered winner of a tournament against $j$ solution if:

- Solution $i$ has a higher level of non-dominance: $i$ frontier < $j$ frontier
- If both solutions are in the same frontier, but $i$ crowd distance is greater than $j$, or $d_i > d_j$

Subsequently, the operators of crossover and mutation are applied, as employees in EAs. At the end of each generation the population $P_t$ and $Q_t$ are inserted as previously described (Fig. 3) in an elitist strategy to obtain new parents population $P_{t+1}$. After reaching a pre-specified number of generations, the algorithm is...
stopped and the frontier of non-dominated solutions of the current population is returned as the final solution of the EA.

C. The Analytic Hierarchy Process (AHP)

The AHP was first proposed by Thomas L. Saaty and their main feature is the pairwise comparison of a hierarchy of criteria and alternatives [6]. It is often used to analyze problems of multicriteria decision-making [6, 7, 11, 12]. The AHP divides the overall problem in evaluations of minor importance, while maintaining the participation of these problems small in global decision, decomposing the structure of problem into a hierarchy (containing criteria and alternatives) [6].

Saaty states that hierarchy is an abstraction of a system structure to study the functional interactions of each components and their impact on the total system [6, 7]. The most creative part of the decisions that have significant effect on the result is the modeling of the problem. In AHP, a problem is structured as a hierarchy, and subsequently undergoes a process of comparison.

The paired comparison is an important component of the AHP. Two criteria (or alternatives) are compared using a nine-point scale, where one (1) means importance “equal”, three (3) the importance is “low”, five (5) clearly indicates “superior”; seven (7) is “very” important, and nine (9) denotes “extremely” important [6, 7]. Even numbers (2, 4, 6, 8) can be used to indicate intermediate values, if necessary. If there are n criteria to be considered, then n (n-1) / 2 pairwise comparisons should be made. Afterwards, an n x n matrix is constructed and the weights of each entity (local and global) are obtained [11, 12].

The consistency of the matrix can be verified by the following indexes: consistency index (CI) and consistency ratio (CR). They are defined in equations (1) and (2) with \( \lambda_{\text{max}} \) is the main value (Eigen) and IR random consistency index, as shown in Tab. 1. For consistency, CI and CR should be less than 0.1 to AHP analysis be considered consistent [11, 12].

\[
CI = (\lambda_{\text{max}} - n) / (n - 1) \quad (1)
\]

\[
CR = CI / RI \quad (2)
\]

According to Saaty [7], the benefit of this method is that as the values of the paired comparisons judgments are based on experience, intuition, and on physical data, AHP can deal with qualitative and quantitative aspects of a decision problem.

**TABLE I.** CONSISTENCY INDEX RANDOM (RI)

<table>
<thead>
<tr>
<th>n'</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP</td>
<td>0</td>
<td>0</td>
<td>0.58</td>
<td>0.9</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
<td>1.45</td>
<td>1.49</td>
</tr>
</tbody>
</table>

a. Dimension of the matrix (n) and b. Random consistency index (RI).

A well-constructed hierarchy is a good model of reality, and can bring some advantages [6, 7]. First, the hierarchical representation of a system can be used to describe how changes in priorities at the highest levels affect the priority of the lowest levels. The hierarchy also allows obtaining an overview of a system, since lower levels of criteria to their purposes at the highest levels. Finally, hierarchical models are flexible and stable: stable because small changes are small effects; whereas flexible because additions to a well-structured hierarchy does not disturb the overall performance [6, 7, 11, 12].

D. Related Works

There are works in the literature that deal with project portfolio selection problem (PPSP). According to Wang [10] these works can be classified into:

- **Scoring Models**
- **Mathematical Models**
- **Financial Ratios Models**
- **Probabilistic Models**
- **Pricing Options Theory**
- **Strategic Approaches**
- **Hierarchical Approaches**
- **Behavioral Approaches**

In recent years, Heuristic Methods were used to solve PPSP. Iamratanakul [8] published a literature review related to this topic, classifying the portfolio selection models in a taxonomy that comprises different types of approaches, one of them is the Heuristic Approach. The evolution of published works that address the problem of portfolio selection can be seen in paper published by Metaxiotis & Liagkouras [9]. The Fig. 5 summarize the evolution of heuristics methods to portfolio selection and other subjects.

![Figure 5. Evolution of publications related to heuristic methods [9].](image)

The line with circles represents the evolution of publications related to heuristic methods for portfolio selection. The line with triangles represents the evolution of the publications of heuristic methods in several contexts.

Among recent works (2010-2014) we can mention; in [13] the authors propose an approach for multiobjective heuristic search technique to support a selection of project portfolio in scenarios with a large number of available projects. In [14] the authors propose an alternative that uses fuzzy logic with a heuristic to choose an optimal portfolio. The same authors present a variation of this alternative in [15] adding a data mining subsystem. The work presented in [16] uses an evolutionary algorithm for selecting an optimal portfolio based on a single objective function. In [17] the authors propose a tool that identifies a set of portfolios (Pareto Optimal) within a cost range allowing the realization of interactive analysis. In [18, 19], the authors present a tool which implements a heuristic algorithm. The Tab. 2 shows the comparison of related work with our approach.
TABLE II. RELATED WORKS COMPARISON

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Related Works</th>
</tr>
</thead>
<tbody>
<tr>
<td>[13] [14] [15] [16] [17] [18, 19]</td>
<td>Our</td>
</tr>
<tr>
<td>Objective Function</td>
<td>● O O O ● ● ●</td>
</tr>
<tr>
<td>Restrictions Set</td>
<td>● ● ● ● ● ●</td>
</tr>
<tr>
<td>Heuristic Search</td>
<td>● ● ● ● ● ●</td>
</tr>
<tr>
<td>Optimal Set</td>
<td>● O O O ● ●</td>
</tr>
<tr>
<td>Structured Decision</td>
<td>● ● ● ● ●</td>
</tr>
<tr>
<td>Post Optimization Selection</td>
<td>● ● ● ● ●</td>
</tr>
<tr>
<td>Portfolio Selection</td>
<td>● ● ● ● ● ●</td>
</tr>
<tr>
<td>Criteria Set</td>
<td>● ● ● ● ●</td>
</tr>
<tr>
<td>Process in Phases</td>
<td>● ● ● ● ●</td>
</tr>
<tr>
<td>NP-hard</td>
<td>● ● ● ● ●</td>
</tr>
</tbody>
</table>

● Meets strongly O partially meets and no symbol, no answer. Criteria developed according to elements often present in publications [8, 9, 13, 14, 15, 16, 17, 18, 19].

The work presented in this article is intended to cover the gaps between the current models.

IV. TWO PHASES MULTICRITERIA APPROACH

A. The Project Portfolio Selection Problem (PPSP)

PPSP is to determine in what ways the available designs can be combined to maximize the return, considering a set of constraints while minimizing the risks involved [21].

Harry Markowitz[20] defines two fundamental characteristics of a portfolio: their expected return and their variance, representing the risk of the portfolio. PPSP can be formally defined as:

\[
\text{Maximize } \bar{R}_p = \sum_{i=0}^{n} X_i \cdot E(R)_i \tag{3}
\]

\[
\text{Minimize } \sigma_p = \sqrt{\sum_{i=0}^{n} w_i^2 \sigma_i^2 + \sum_{i=0}^{n} w_i w_j \rho_{ij} \sigma_i \sigma_j} \tag{4}
\]

Subject \( \sum_{i=0}^{n} C_i \leq \text{budget} \)

max(payback) \( \leq \text{payback} \)

\( P_i \) depends \( P[1..n] \)

\( P_i \) exclude \( P[1..n] \)

\( P_i \) \#\#\# E(R)

\( E(R) > 0 \) \( \tag{10} \)

B. Function Return

Equation (3) represents the objective function to be maximized [20, 21]. The first feature of the portfolio, its return expected \( \bar{R}_p \), is simply the weighted average of the returns of individual projects that comprise it, where:

- \( X_i \) is the percentage invested in the project \( i \).
- \( E(R)_i \) is the percentage invested in the project \( i \).

C. Risk Function

Equation (4) represents the function that must be minimized [20, 21]. The key feature of this equation is the risk, as measured by their variance, where:

- \( W_i W_j \) represent, respectively, the share of \( x \) and \( y \) in the project portfolio
- \( \sigma_i^2 \sigma_j^2 \) represent the variance of \( x \) and \( y \) projects, respectively, with respect to the risks identified
- \( \rho_{xy} \) covariance between \( x \) and \( y \) projects

The covariance \( \rho_{xy} \) is given by Person Covariance [15] in Equation (11).

\[
\rho_{xy} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{(n \sum x_i^2 - (\sum x_i)^2) \cdot (n \sum y_i^2 - (\sum y_i)^2)}} \tag{11}
\]

D. Constraints

Equations (5) and (6) ensure that the total of investment on portfolio and the return period (payback) are not higher than expected [21]. Equations (7) and (8) show the dependencies and exclusionary between projects. This means that when selecting a project, you must also select their dependents and/or eliminate mutually exclusive [21]. In (9), it ensures that a project should exist only once within the portfolio. Finally, Equation (10) ensures that the final return is greater than zero [21]. The search space is given by \( 2^n \), where \( n \) is the number of projects available for selection[3].

E. Structure of Two Phase Approach

Considering the different stages of the project portfolio selection process [21], a decision structure of two phases was created (Fig. 6), aiming (i) to generate a set of optimal solutions and (ii) allowing that one of the solutions is selected by a structured method.

MOO

AHP

- Define the objective functions
- Define the restrictions
- Get a set of optimal solutions
- Define the hierarchy of criteria
- Define the importance of the criteria
- Select an alternative

Figure 6. Structure of the two-phase approach.

The first phase (MOO) uses the NSGA-II algorithm with the objective functions and constraints explained above to generate a set of optimal solutions (Pareto Front). After that, a hierarchical structure is used to select a single solution (AHP) considering a set of criteria.

F. Hierarchy of Criteria

A total of 34 criteria to assess the values of project portfolios in the relevant literature were researched. These criteria were

---

2 Was awarded with Prize in Economic Sciences in Memory of Alfred Nobel 1990.

3 This search space disregards the restrictions imposed on the problem. This is one of the types of problems that can be classified as NP-hard [21], polynomial methods which would take considerable time to test all solutions.
organized into two groups: (i) endogenous criteria with focus on creating internal value to the organization (these criteria should express factors that are within the control of the organization), (ii) exogenous criteria intrinsically related with environment being beyond control of the organization [21]. This classification (and criteria) can be seen in Tab. 3.

![Hierarchical tree diagram](image)

**TABLE III. SELECTION CRITERIA PROJECT PORTFOLIO**

<table>
<thead>
<tr>
<th>Endogenous</th>
<th>Exogenous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit with corporate strategic objectives</td>
<td>Position of the related technology in its own life cycle</td>
</tr>
<tr>
<td>Profitability</td>
<td>Environmental and safety consideration</td>
</tr>
<tr>
<td>Capability of research team</td>
<td>Dealing with international Sanctions</td>
</tr>
<tr>
<td>Financing capacity</td>
<td>Public support for development</td>
</tr>
<tr>
<td>Impact on enhancing Innovation</td>
<td>Barriers to copy or imitation</td>
</tr>
<tr>
<td>Contents of technical plan</td>
<td>Market volume opened by Research result</td>
</tr>
<tr>
<td>Serving as infrastructure</td>
<td>Competition intensity</td>
</tr>
<tr>
<td>Technological connections</td>
<td>Benefits for human life</td>
</tr>
<tr>
<td>Extensibility of results and Span of application</td>
<td>Impact on firm prestige</td>
</tr>
<tr>
<td>Appropriateness for research cost</td>
<td>Potential for progress</td>
</tr>
<tr>
<td>Equipment support</td>
<td>Market Dynamics</td>
</tr>
<tr>
<td>Appropriateness for research project timing</td>
<td>Potential for research product growth</td>
</tr>
<tr>
<td>Impact on enhancing Firm Productivity</td>
<td>Impact on societal stakes</td>
</tr>
<tr>
<td>Advancement of related Technology</td>
<td>Number of stakeholders</td>
</tr>
<tr>
<td>Research gap to corporate core business</td>
<td>Impact of related technology on competitive issues</td>
</tr>
<tr>
<td>Quality Improvement</td>
<td></td>
</tr>
<tr>
<td>Impact on employees learning and growth</td>
<td></td>
</tr>
<tr>
<td>Experience accumulated in the field</td>
<td></td>
</tr>
<tr>
<td>Synergy with other projects</td>
<td></td>
</tr>
</tbody>
</table>

The criteria bold has a higher representation [21].

The criteria of Tab. 3 are used often in specialized literature [21], providing a structure to allow for a meaningful analysis. An example of the result of the hierarchy can be seen in Fig. 7.

The initial population size was estimated empirically, starting with 10 individuals, increasing by 10 until the result of the algorithm was not changed reaching number (rounded) of 100 individuals. This size remains the same for population during the iterations.

The stopping criterion of the genetic algorithm might vary according to user’s choice. One way is to (i) define a number of generations that must be created. Another way is (ii) run it until it is a population where individuals have the evaluation function to be reached [5]. The criterion used was (iii) convergence, i.e., there is no significant improvement in the solution for a given number of generations. The Tab. 4 summarizes these parameters.

**TABLE IV. PARAMETERS USED IN NSGA-II ALGORITHM**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of Projects</td>
<td>42 software projects</td>
</tr>
<tr>
<td>Initial Population</td>
<td>100 individuals</td>
</tr>
<tr>
<td>Population Size</td>
<td>100</td>
</tr>
<tr>
<td>Function to be Maximized</td>
<td>Equation (3)</td>
</tr>
<tr>
<td>Function to be Minimized</td>
<td>Equation (4)</td>
</tr>
<tr>
<td>Constraints</td>
<td>Equations (5), (6), (7), (8), (9) and (10)</td>
</tr>
<tr>
<td>Representations Scheme</td>
<td>Binary Encoding</td>
</tr>
<tr>
<td>Selection Operator</td>
<td>Crowding Tournament</td>
</tr>
<tr>
<td>Crossover Operator</td>
<td>Partially Mapped Crossover (PMX)</td>
</tr>
<tr>
<td>Mutation Rate</td>
<td>0.01</td>
</tr>
<tr>
<td>Stop Condition</td>
<td>100 springs without improve over solutions</td>
</tr>
</tbody>
</table>

**B. Results and Analysis of Experiments**

The data used in experiments was got from a set of 42 software projects originated from strategic planning of a midsize company. These experiments were carried out considering the real scenario and scenario simulations, where we tested variations in values of constraints.

The Pareto optimal set is displayed in Fig. 8. This set contains portfolios that were constructed by the combination of projects available, optimizing the functions represented by Equations (3) and (4) and respecting constraints expressed by equations (5), (6), (7), (8), (9) and (10). The solutions contained in the set can be considered optimum (from the viewpoint of each objective function). Variations on the constraints (Equations (5) and (6)) alter the set of solutions (i) reducing the amount of available portfolios and (ii) reducing the efficient frontier (solutions of lower return and lower risk), confirming the theory of Markowitz [20].
The Fig. 8 shows two adjacent goals of heuristic search (i) search for new frontiers and (ii) increase the biodiversity of the solutions in order to obtain a greater number of alternatives available for post-optimization step [5].

With a set of optimal solutions, one can choose among them. This step is important because you can use the tacit knowledge, experience, and intuition of experts [6, 11, 12]. For experiment, we used the criteria of Tab. 3 with 7 people (including functional managers, project managers, and directors) and the paired comparison obtained by Delphi method [11].

The result was the selection of a portfolio contained in the Pareto optimal set, with a tendency for higher risk/return and being accepted, by those involved, as the best portfolio.

VI. CONCLUSIONS & FURTHER WORK

The models and algorithms [5] discussed here aimed to bring an approach to organizational decision-making and give a new dimension to the project portfolio selection. This approach has better results when there are several constraints to be satisfied and/or the problem is large (several projects available) to be solved in deterministic or polynomial way (NP-hard) [21].

Importantly, the selection of a project portfolio assumes a broader and more complex understanding than the single use of a particular method [11, 12, 22, 23]. It presupposes that the decision on a portfolio is the result of negotiation, human aspects and strategic analysis. The approach of this work encourages and guides decision-making, but should not be used as the sole method.

Among the contributions of the approach can be mentioned (i) a solution to the combinatorial analysis 2^N and (ii) the structure of a hierarchy of criteria derived from subjective aspects that allow the selection of a single portfolio.

In future work, one can explore a larger set of constraints and other functions to be optimized, such as minimizing the costs of the portfolios generated by transforming PPSP from 2 to 3 or N goals, added data sources, such as Master Data Management (MDM) [22] to get more precision information.

REFERENCES