

Project Ranking-Based Portfolio Selection Using Evolutionary Multiobjective Optimization of a Vector Proxy Impact Measure

S. Samantha Bastiani¹, Laura Cruz¹, Eduardo Fernández², Claudia Gómez¹, Victoria Ruiz¹

¹Madero Institute of Technology, Ciudad Madero, México

²Autonomous University of Sinaloa, Culiacan, México

cgs71@hotmail.com, lcruzeves@prodigy.net.mx, eddyf@uas.edu.mx, b_shulamith@hotmail.com, victoria.rzrz@gmail.com

Abstract

Selecting project portfolios Decision-Maker usually starts with limited information about projects and portfolios. One of the challenges involved in analyzing, searching and selecting the best portfolio is having a method to evaluate the impact of every project and portfolio in order to compare them.

This paper develops a model for composing public-oriented project portfolios. Information concerning the quality of the projects is in the form of a project-ranking, which can be obtained by the application of a proper multi-criteria method; however the ranking does not assume an appropriate evaluation. A best portfolio is primarily found through a multi-objective optimization that regards the impact indicators that reflect the quality of the projects in the portfolio and competent portfolios' cardinalities. Overall good solutions are obtained by developing an evolutionary method, which is found to perform well in some test examples.

Keywords: Project portfolio selection; Multi-objective optimization; Multi-criteria analysis

1. Introduction

Project portfolio selection is one of the most difficult, yet most important decision-making problems faced by many organizations in government and business sectors. To carry out the project selection, the decision maker usually starts with limited information about projects and portfolios. His/her time is often the most critical scarce resource. In multiple situations the decision maker feels more comfortable employing simple decision procedures, because of lack of available information, lack of time, aversion to more elaborated decision methods, and even because of his/her fondness for established organizational practices. Cooper et al. ([1]) argues about popularity of scoring and ranking methods in R&D project management in most American enterprises.

Methods of scoring and ranking are used by most of the government organizations that fund R&D projects.

Usually, methods for scoring, ranking or evaluating projects contain some way of aggregating multi-criteria descriptions of projects (e.g. [2]). Validity of these methods depends on how accurately ranking and scores reflect decision maker preferences over portfolios. In fact, the portfolio's score should be a value function on the portfolio set, but this requires a proper elicitation of decision maker preferences inside the portfolio's space.

Ranking is also used in problems where a "Participatory budgeting" is involved. "Participatory budgeting" can be defined as a public space in which government and society agree on how to adapt priorities of citizenship to public policy agenda. The utility of these participatory exercises is that the government obtains information about priorities of the participating social sectors, and might identify programs with a consensual benefit.

Ranking of public actions given by the participants is an expression of their preferences on projects, not on portfolios. Let us assume that a method of integrating the individual ranking on a collective order is applied, as the Borda score or a procedure based on the exploitation of collective fuzzy preference relations (e.g. [3, 4, 5]). With the obtained group order, the decision maker has more information about social preferences regarding the different actions to budget. The decision maker should use that information to find the best portfolio.

This paper proposes a new model for project portfolio selection, which makes use of the ranking of a set of projects according to the preferences of a decision maker. The model is formulated by a set of indirect indicators that reflect the impact of the portfolio in terms of the number of projects and the positions they occupy in the ranking. This paper is structured as follows: the background is briefly described in the second section. It is also shown the algorithm which lead to optimize the proposed impact model. Section 3 presents such model, followed by a description of the solution algorithm (Section 4). Finally in Section 5 we give empirical evidence that supports our results and some conclusions are given in Section 6.

2. Background

A project is a temporary process, singular and unrepeatable, pursuing a specific set of objectives ([6]). In this paper, it is not considered that the projects can be divided into smaller units, such as assignment or activities.

A portfolio consists of a set of projects that can be performed in the same period of time ([6]). Due to that, projects in the same portfolio can share the available resources of the funding organization. Therefore, it is not sufficient to compare the projects individually; the decision maker should compare project groups to identify which portfolio makes a major contribution to the objectives of the organization.

Selecting portfolios integrated by properly selected projects is one of the most important decision problems for public and private institutions [7, 8]. The Decision Maker (DM) (a person or a collective entity) is in charge for selecting a set of projects to be supported ([9]).

2.1. Related works

Gabriel et al. ([10]) proposed an additive function as a portfolio's score. This function aggregates the rank of projects. The simplest model is to assign project priorities in correspondence to the project rank (the highest priority is assigned to the best ranked project and so on). The portfolio's score is the sum of priorities associated with its projects. 0-1 mathematical programming is used to maximize the score.

Mavrotas et al. ([11]) proposed an additive function depending on a project's augmented score. This augmented score is built according to the project's specific rank. The augmented score of a project A holds that no combination of projects with worse ranking positions and a lower total cost can have a score bigger than A. The augmented score is obtained by solving a knapsack problem for each project. The portfolio's score is the sum of its projects' augmented scores.

Validity of these methods depends on how accurately the ranking scores reflect the decision maker preferences over portfolios. In fact, the portfolio's score should be a value function on the portfolio set, but this requires a proper elicitation of the decision maker preferences in the portfolio's space. In order to illustrate limitations of those methods, consider the following example: Let us suppose a 20-project strict ranking; priority 20 is assigned to the best project; 19 is assigned to the second one; 1 is assigned to the worst ranked project. Considering a score given by the sum of priorities, the portfolio containing the best and the worst projects (score = 21) should be indifferent to the portfolio containing the second best project and the second-to-last one (score = 21). The DM could hardly agree with such a proposition.

It is necessary to compare impact of possible portfolios in order to find the best one. The information provided by the simple project ranking is very poor for portfolio optimization purposes. Hence, some proxy impact

measures should be defined. This problem was approached by [12] under the assumption that "the portfolio impact on a decision maker mind is determined by the number of supported projects and their particular rank". If project A is clearly better ranked than B, then A is admitted to have "more social impact" than B.

The DM should consider this information from the ranking. The appropriateness of a portfolio is not only defined by the quality of the included projects, but also by the amount of contained projects. The purpose is to build a good portfolio by increasing the number of supported projects and controlling the possible disagreements regarding decision maker preferences, which are assumed as incorporated in input ranking. For Fernandez and Olmedo ([12]), a discrepancy is the fact that given a pair of projects (A,B) (being B worse ranked than A), B belongs to portfolio and A does not. Different categories of discrepancy are defined according to the relative rank of the concerning projects. Some discrepancies might be acceptable between the information provided by the ranking and the decisions concerning the approval (hence supporting) of projects, whenever this fact increases the number of projects in the portfolio. However, this inclusion should be controlled because the admission of unnecessary discrepancies is equivalent to underestimating the ranking information. A multi-objective optimization problem is solved by using NSGA-II, in which the objective functions are the number of supported projects and the number of discrepancies (separately in several functions, in regard to the importance of each kind of discrepancy) ([12]).

Main drawback: a portfolio quality measure model based solely on discrepancies and in the number of supported projects is highly questionable; more information is required about project impacts. If decision maker thinks in terms of priority and relatively important projects, their numbers and ranks should be considered.

2.2. NSGA-II (Non-dominated Sorting Genetic Algorithm-II)

The research problem implicates the use of techniques of multi-objective optimization, particularly multi-objective evolutionary algorithms (MOEAs). One advantage of these algorithms is its capacity of handling problems with an exponential complexity. Other advantage is their ability to generate an approximation to the Pareto optimal set in a single run instead of having to perform many runs as in conventional multi-objective optimization. Several works have reported successful results with this kind of algorithms ([12]).

One of the most used algorithms for solving multi-objective problems is the NSGA-II (Non-dominated Sorting Genetic Algorithm), which has gained much popularity solving problems efficiently. It is shown in Figure 1.

Fig 1. Structure of the algorithm NSGA-II.

```

1  Rt= Pt∪Qt
2  F= fast-nondominated-sort (Rt)
3  until |Pt+1|< N
4    crowding-distance-assignment (Fi)
5  Pt+1=Pt+1∪ Fi
6  Sort (Pt+1, ≥n)
7  Pt+1=Pt+1[0:N]
8  Qt+1=make-new-pop(Pt+1)
9  t=t+1

```

The procedure "Fast non-dominated sorting" (shown in Figure 2), optimizes the algorithm NSGA-II.

Finally this algorithm has a diversity indicator whose evaluation is shown in Figure 3. This indicator favors solutions in less populated regions of the search space; these solutions will be advantaged by the selection mechanism ([13]).

Fig 2. Structure of the Fast-nondominated-sort procedure

```

1  for each p∈ P
2  for each q∈ P
3    if(p < q) then
4      Sp= Sp∪ {q}
5    else if(q < p) then
6      np = np + 1
7    if np=0 then
8      F1=F1∪ {p}
9    i=i+1
10 while Fi≠ 0
11   H= 0
12   for each p∈ Fi
13   for each q∈ Sp
14     nq= nq - 1
15   if nq = 0 then H= H∪ {q}
16   i=i+1
17   Fi = H

```

Fig 3. Structure of the algorithm of the Crowding-distance-assignment (I).

```

1  l= |I|
2  for each i, set I [i]distance=0
3  for each objective m
4    I=sort (I, m)
5    for i =2 to(l-1)
6  I [i]distance = I [i]distance + (I [i + 1]).m -
  I [i-1].m)

```

3. The proposed model

The new model overcomes the idea proposed in [12, 14]. In this model the optimization is performed over indicators, which positively provide indirect information on the impact of the portfolio. This model handles three categories for projects: priority, satisfactoriness and acceptability, besides incorporating a ranking in descending order.

Once data has been established, solution sets are evaluated through a set of indicators of impact that form the model proposed in this paper. The following functions are defined:

$$I_1(\vec{x}) = \sum_{i=1}^n x_i F(i,1) \quad (1)$$

$$F(i,1) = \begin{cases} 1 & \text{if } i \in G_1 \\ 0 & \text{Otherwise} \end{cases}$$

where the binary variable x_i indicates whether the i th project belongs to the portfolio or does not. That is, $x_i = 1$ if the i th project belongs to portfolio; otherwise $x_i = 0$. Note that Function I_1 counts how many projects belonging to the priority category (Group 1) are contained in the portfolio.

$$I_2(\vec{x}) = \sum_{i=1}^n x_i (n - i) F(i,1) \quad (2)$$

where $x_i (n-i)$ is a value that reflects the rank order of the supported i th project. I_2 increases with the rank ordering of the supported projects of the priority category. This function measures (in proxy way) how good the supported priority projects are.

$$I_3(\vec{x}) = \sum_{i=1}^n x_i F(i,2) \quad (3)$$

$$F(i,2) = \begin{cases} 1 & \text{If } i \in G_2 \\ 0 & \text{Otherwise} \end{cases}$$

where the binary variable x_i has the same above meaning. Note that Function I_3 counts how many projects belonging to the satisfactory category (Group 2) are contained in the portfolio.

Besides

$$I_4(\vec{x}) = \sum_{i=1}^n x_i (n - i) F(i,2) \quad (4)$$

measures (in proxy way) how good the supported satisfactory projects are.

Similarly we define

$$I_5(\bar{x}) = \sum_{i=1}^n x_i F(i,3) \quad (5)$$

$$F(i,3) = \begin{cases} 1 & \text{If } i \in G_3 \\ 0 & \text{Otherwise} \end{cases}$$

Function I_5 counts how many projects belonging to the acceptable category (Group 3) are contained in the portfolio.

Finally

$$I_6 = \sum_{i=1}^n x_i \quad (6)$$

represents the portfolio cardinality.

We assume that the DM "feels" the potential impact of the portfolio in terms of the numbers of projects for category and the positions they occupy.

The best portfolio should be the best solution of the multi-objective problem:

$$\text{Max}_{C \in R_F} (I_1, I_2, I_3, I_4, I_5, I_6) \quad (7)$$

where R_F is the feasible region determined by budgetary constraints.

In this case the DM, based on his/her preferences, should select the best portfolio.

4. The proposed algorithm

The algorithm developed in this research work is called Evolutionary algorithm for Solving the public Portfolio problem from Ranking Information (ESPRI). It is inspired by the NSGA-II algorithm developed by Deb et al. ([13]), which successfully manages exponential complexity ([12]). ESPRI uses the indicator vector from Equation 7 for evaluating solutions.

To illustrate ESPRI algorithm process, a set of n projects is taken as example, with its respective total budget as well as necessary budget for each project. Previously, such projects were ranked according to decision maker preferences. Heuristically, the projects were separated in three categories: priority, satisfactoriness and acceptability. Once this process is complete, the algorithm generates random portfolios, which form the NSGA-II initial population

Later, the following procedures are applied: fast-nondominated, crowding distance and genetic operators. Finally, the algorithm shows the found non-dominated solutions for the decision maker. Figure 4 shows ESPRI algorithm.

Fig 4. Structure of the algorithm

```

1  Rt = Pt ∪ Qt
2  quality assessment: Impact Indicators Model.
3  F = fast-nondominated-sort (Rt)
4  while: |Pt+1| < N
5    crowding-distance-assignment (Fi)
6  Pt+1 = Pt+1 ∪ Fi
7  Sort (Pt+1, ≥n)
8  Pt+1 = Pt+1[0:N]
9  Qt+1 = Create-new-pop(Pt+1)
10 t = t + 1

```

5. Computational Experiments

This section describes conducted experiments with the proposed evolutionary algorithm (ESPRI).

The aim of this experiment is to study indicator new model capacity, as well as to compare ESPRI solutions against the state of the art.

5.1. Experimental Environment

The following configuration corresponds to experimental conditions required for tests described in this paper:

1. Software: Operating System, Mac OS X Lion 10.7.5 (11G63b) Java Programming Language, Compiler NetBeans 7.2.1.
2. Hardware: computer equipment, Intel Core i7 2.8 GHz CPU and 4 GB of RAM.
3. Instances: An instance used for this study was taken from the state of the art, reported by Fernández et al. in [12, 14].
4. Performance Measure: In this case the performance is measured through the aforementioned six objectives (Eq. 7).

5.2. An illustrative example

Within the public portfolio problem, an instance for ranking strategy is formed by four attributes: Id, total amount to be distributed, project cost and ranking. The test example is taken from the state of the art, ([12]), which works with an instance that consists of 100 projects. The projects are separated into three categories: priority, satisfactoriness and acceptability, approximately uniform.

For the experiment, ESPRI algorithm was run 20 times; in each run 200 iterations were performed. The experiment reported was held with an instance of 100 projects with a total amount of 2.5 billion to be distributed; this instance can be seen in Table 1.

The ESPRI algorithm was set as: one-point crossover, Mutation probability = 0.5, population size = 200 and Number of Generations = 100.

Note that if the resources were distributed strictly following the ranking order, the resulting portfolio would have 22 projects, all belonging to the priority category.

Table 2 shows a representative sample of the approximation to Pareto frontier, which our proposal might reach. Red marks represent a set of solutions preferred by a decisions maker interested in increasing the number of priority projects that are supported, as well as the total number of projects, but with emphasis on those considered satisfactory (category 2).

One of the best compromise solutions obtained by Fernandez et al. in [12] is shown in Table 3. This solution contains a total of 24 projects, all belonging to the priority category. Compared to it, our red-marked solutions in Table 2 seem to be of greater impact and have equal or greater number of priority projects (24, 25, or 26), and contain much more total projects.

Our solutions would be preferred by every decision maker whose preferences are identified with the number of priority needs to attend to, and the total amount of needs (projects) addressed. Table 3 allows comparing the best solution by Fernández et al. ([12]) with our solution in the project space.

The impact indicator model is more flexible. The comparison shows that the proposal of [12] is a rigid model. This does not find several solutions that would have greater social benefit.

Table 4 shows the results of the instance that was used in [12]. As can be seen, the obtained non-dominated solutions seem to be satisfactory for the DM. The solutions obtained by our proposal should be more preferred than the best solution in [12] because this contains less projects and less priority projects.

6. Conclusions

The proposed model of impact indicators of the portfolio can explore the solution space and generate potential best portfolios, besides reasonably modelling decision maker preferences on portfolios under limited information about projects.

It was also proposed an evolutionary algorithm based on the NSGA-II that seems to be capable of obtaining solutions near to the Pareto frontier. The obtained solutions are more satisfactory than those obtained by the state of the art.

The quality of the solutions indicates that the algorithm converges close to the true Pareto frontier where best portfolios lie; this helps the decision maker to analyze his/her own preferences and to clarify his/her decisions.

We have obtained some evidence in favor to our proposal, which allows helps the DM in finding a rational compromise between the quality of the projects in the portfolio and the number of projects approved.

Table 1. Instance of 100 projects.

P	Budget	P	Budget	P	Budget	P	Budget
1	84.00	26	31.25	51	27.50	76	46.50
2	124.50	27	26.50	52	41.25	77	44.00
3	129.75	28	36.25	53	29.50	78	25.75
4	147.75	29	50.00	54	25.25	79	38.25
5	126.00	30	34.75	55	40.00	80	40.75
6	137.25	31	48.25	56	30.75	81	42.75
7	96.00	32	46.00	57	39.00	82	43.00
8	84.75	33	36.75	58	44.50	83	32.25
9	93.00	34	34.00	59	47.50	84	37.75
10	121.50	35	26.00	60	36.00	85	44.75
11	102.75	36	31.75	61	28.50	86	27.00
12	141.75	37	29.75	62	29.00	87	39.50
13	105.75	38	37.25	63	30.25	88	30.00
14	98.25	39	26.75	64	49.50	89	37.50
15	101.25	40	43.75	65	33.00	90	49.00
16	83.25	41	27.25	66	38.50	91	41.75
17	109.50	42	47.00	67	33.50	92	39.25
18	107.25	43	41.00	68	48.50	93	34.50
19	135.00	44	30.50	69	35.00	94	49.75
20	97.50	45	45.25	70	28.75	95	48.00
21	127.50	46	26.25	71	25.50	96	29.25
22	114.00	47	45.50	72	40.25	97	47.75
23	106.50	48	44.25	73	38.75	98	42.25
24	94.50	49	48.75	74	46.75	99	46.25
25	43.50	50	33.25	75	37.00	100	39.75
Total							5542.00

*P: Rank ordering Identifier

Table 4. Comparison of experimental results with the objectives evaluated with impact indicator model.

Papers	(I1, I2, I3, I4, I5, I6)
Fernández et al	24, 1995, 0, 0, 0, 24
Our proposal	24, 1995, 15, 859, 1, 40
Our proposal	25, 2051, 11, 644, 3, 39
Our proposal	25, 2079, 11, 609, 1, 37
Our proposal	26, 2151, 10, 577, 1, 37
Our proposal	26, 2163, 9, 514, 1, 36
Our proposal	25, 2102, 8, 464, 2, 35

References

- [1] Cooper, R., Edgett, S., Kleinschmidt, E., (2001): "Portfolio Management for New Product Development: Results of an Industry Practices Study", *R&D Management* 31 (4), 361-380.
- [2] Henriksen A.D., Traynor A.J. (1999): "A practical R&D project selection scoring tool", *IEEE Transactions on Engineering Management* 46 (2), 158-170.
- [3] Macharis C., Brans, J.P., and Mareschal, B. (1998): "The GDSS PROMETHEE Procedure. A PROMETHEE-GAIA based procedure for group decision support", *Journal of Decision Systems*, 7, 283-307.
- [4] Leyva, J.C., Fernandez, E., (2003). A new method for group decision support based on ELECTRE-III methodology. *European Journal of Operational Research*, 148 (1), 14-27.
- [5] Fernández, E., López, E., Bernal, S., Coello Coello, C. A., and Navarro, J. Evolutionary multiobjective optimization using an outranking-based dominance generalization. *Computers & Operations Research*, 37(2):390–395.(2010a).
- [6] Carazo, A. F., Gómez, T., Molina, J., Hernández-Díaz, A. G., Guerrero, F.
- [7] M., and Caballero, R. Solving a comprehensive model for multiobjective project portfolio selection. *Computers & Operations Research*, 37(4):630–639.(2010).
- [8] Castro, M. Development and implementation of a framework for I&D in public organizations. Master's thesis, Universidad Autónoma de Nuevo León. (2007).
- [9] García, R., Hyper-Heuristic for solving social portfolio problem. Master's Thesis, Instituto Tecnológico de Cd. Madero. (2010).
- [10] Fernández, E. and Navarro, J., A genetic search for exploiting a fuzzy preference model of portfolio problems with public projects. *Annals OR*, 117(191-213):191–213.(2002).
- [11] S. Gabriel, Kumar, S., Ordoñez, J., Nasserian, A. (2006): "A multiobjective optimization model for project selection with probabilistic consideration", *Socio-Economic Planning Sciences* 40 (4), 297-313.
- [12] Mavrotas, G., Diakoulaki, D., Koutentis, A. (2008): "Selection among ranked projects under segmentation, policy and logical constraints", *European Journal of Operational Research* 187 (1), 177-192, 2009.
- [13] Fernández, E., Olmedo R. Public Project Portfolio Optimization Under A Participatory Paradigm. *Applied Computational Intelligence and Soft Computing*. Archive Volume 2013, January 2013. Article No. 4
- [14] Deb, K., *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, Chichester-New York-Weinheim-Brisbane-Singapore-Toronto. (2001)
- [15] Fernández E., Luz Florida Félix, Gustavo Mazcorro., Multi-objective optimization of an outranking model for public resources allocation on competing projects. *Int. J. Operational Research*, Vol. 5, No. 2, pp. 190-210. (2009).
- [16] Coello Coello, C. A., Lamont, G. B., and Van Veldhuizen, D. A. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Genetic and Evolutionary Computation. Springer, 2nd edition. (2007).
- [17] Fernández Eduardo R., Navarro Jorge A., Olmedo Rafael A. Modelos y Herramientas Computacionales para el Análisis de Proyectos y la Formación de Carteras de I&D. *Revista Iberoamericana de Sistemas, Cibernética e Informática*. Volumen 1 - Número 1 – Año 2004, páginas: 59-64.
- [18] Fernández, E., López, E., López, F., and Coello Coello, C. A. Increasing selective pressure towards the best compromise in evolutionary multiobjective optimization: The extended NOSGA method. *Information Sciences*, 181(1):44–56.(2010b).
- [19] Ghasemzadeh, F., Archer, N., and Iyogun, P., A zero-one model for project portfolio selection and scheduling. *Journal of the Operational Research Society*, 50(7):745–755.(1999).

- [20] Nebro Antonio J., Alba Enrique, Luna Francisco. Optimización Multi-Objetivo y Computación Grid. Departamento de Lenguajes y Ciencias de la Computación. Universidad de Málaga. E.T.S. Ingeniería Informática. Campus de Teatinos, 2004.
- [21] Peñuela, C. and Granada, M., Optimización multiobjetivo usando un algoritmo genético y un operador elitista basado en un ordenamiento no dominado (NSGA-II). *Scientia Et Technica*, 8(35):175–180. (2007).
- [22] Roy, B. (1990). “The Outranking Approach and the Foundations of ELECTRE methods”, in Bana e Costa, C.A. (ed.), *Reading in multiple criteria decision aid*, Springer- Verlag, Berlin , 155-183.
- [23] Roy, B. and Slowinski, R., Handling effects of reinforced preference and counter-veto in credibility of outranking. *European Journal of Operational Research*, 188(1):185–190.(2008).
- [24] Tenorio Rodríguez Gilberto Javier. Optimización de carteras formadas por proyectos interdependientes en organizaciones públicas. Tesis para obtener el grado de maestro en ciencia en Ingeniería de sistemas, 2010.