Krassimir Markov, Vladimir Ryazanov,
Vitalii Velychko, Levon Aslanyan
(editors)

# New Trends in Classification and Data Mining

I T H E A
SOFIA
2010

Krassimir Markov, Vladimir Ryazanov, Vitalii Velychko, Levon Aslanyan (ed.)

New Trends in Classification and Data Mining

ITHEA®

Sofia, Bulgaria, 2010

First edition

Recommended for publication by The Scientific Concil of the Institute of Information Theories and Applications FOI ITHEA

This book maintains articles on actual problems of classification, data mining and forecasting as well as natural language processing:

- new approaches, models, algorithms and methods for classification, forecasting and clusterisation. Classification of non complete and noise data;
- discrete optimization in logic recognition algorithms construction, complexity, asymptotically optimal algorithms, mixed-integer problem of minimization of empirical risk, multi-objective linear integer programming problems;
- questions of complexity of some discrete optimization tasks and corresponding tasks of data analysis and pattern recognition;
- the algebraic approach for pattern recognition - problems of correct classification algorithms construction, logical correctors and resolvability of challenges of classification, construction of optimum algebraic correctors over sets of algorithms of computation of estimations, conditions of correct algorithms existence;
- regressions, restoring of dependences according to training sampling, parametrical approach for piecewise linear dependences restoration, and nonparametric regressions based on collective solution on set of tasks of recognition;
- multi-agent systems in knowledge discovery, collective evolutionary systems, advantages and disadvantages of synthetic data mining methods, intelligent search agent model realizing information extraction on ontological model of data mining methods;
- methods of search of logic regularities sets of classes and extraction of optimal subsets, construction of convex combination of associated predictors that minimizes mean error;
- algorithmic constructions in a model of recognizing the nearest neighbors in binary data sets, discrete isoperimetry problem solutions, logic-combinatorial scheme in high-throughput gene expression data;
- researches in area of neural network classifiers, and applications in finance field;
- text mining, automatic classification of scientific papers, information extraction from natural language texts, semantic text analysis, natural language processing.

It is represented that book articles will be interesting as experts in the field of classifying, data mining and forecasting, and to practical users from medicine, sociology, economy, chemistry, biology, and other areas.

General Sponsor: Consortium FOI Bulgaria (www.foibg.com).

Printed in Bulgaria



**ISBN 978-954-16-0042-9**

C\o Jusautor, Sofia, 2010

# REFERENCE-NEIGHBOURHOOD SCALARIZATION FOR MULTIOBJECTIVE INTEGER LINEAR PROGRAMMING PROBLEMS

**Krassimira Genova, Mariana Vassileva**

**Abstract**: Various real problems can be modeled as multicriteria optimization problems (MOP). In the general case, there is no single solution that optimizes all the criteria, but there is a set of solutions where improvement in the value of one criterion leads to deterioration in the value of at least another criterion. This set is known as a Pareto optimal set and any element of this set could be the final solution of the MOP. In order to select the final solution, additional information is necessary and it is supplied by the so-called decision maker. The quality of the interactive algorithms for solving MOP depends mainly on the scalarizing problems they are designed and based on. The scalarizing problems of the reference neighborhood, which are presented in the paper, are especially appropriate for solving multiobjective linear integer programming problems.

***Keywords**: Multicriteria Linear  Integer Optimization, Scalarizing  problem.*

***ACM Classification Keywords**: G.1.6. Optimization – Integer Programming*

## Introduction

Several criteria (or objective functions) are simultaneously optimized in the feasible set of solutions (or alternatives) in the problems of multicriteria optimization (MO). In the general case, there is no single solution that optimizes all the criteria. Instead, there is a set of solutions where improvement in the value of one criterion leads to deterioration in the value of at least another criterion. This set is known as a Pareto optimal set. Any element of this set could be the final solution of the multicriteria optimization problem. In order to select the final solution, additional information is necessary and it is supplied by the so-called decision maker (DM). The information that the DM gives reflects his/her global preferences with respect to the quality of the solution obtained. Generally, multicriteria optimization has to combine two aspects: optimization and decision support.

There are two main approaches in solving MO problems: a scalarizing approach and an approximation approach. The major representatives of the scalarizing approach are the interactive algorithms. Scalarization means transformation of the multicriteria optimization problem into one or several single-criterion optimization problems, called scalarizing problems. . The quality of the interactive algorithms depends mainly on the scalarizing problems they are designed and based on. The main property of every scalarizing problem is that each optimal solution generated is a Pareto (or weakly Pareto) optimal solution of the original MO problem. The properties of the scalarizing problems of the reference neighborhood, presented in the paper, make them especially appropriate in realizing a "continuous-integer" approach in interactive algorithms for solving a general multiobjective linear integer programming problem /MOILP/. Instead of generating integer solutions for evaluation at each iteration, the DM may evaluate linear continuous solutions at most of the iterations. In the criteria space the PO integer solutions are placed relatively close to PO continuous solutions. Thus, in the learning phase, the DM could be trained on the basis of linear continuous solutions, instead on the basis of integer solutions, which is particularly important for MOILP problems with large dimensions. In addition, the DM may also learn on the basis of approximate weak PO solutions, found relatively near to a weak PO surface.

The present paper describes scalarizing problems of the reference neighborhood, called RNP1, RNP1e, RNP1-L, RNP1-Le, RNP1-L and RNP3. The rest of rest of the paper is organized as follows: the second section describes the scalarizing problems of the reference neighborhood, called RNP1, RNP1e, RNP1-L, RNP1-Le, RNP1-L and RNP3. The third section presented the basic properties of the scalarizing problems of the reference neighborhood. Finally, the conclusions are given in the last section.

## Description of the scalarizing problems

Let us consider the multiobjective integer linear programming (MOILP) problems:

$$\text{"max"}\{f_k(x) = c^k x\},\ k \in K \tag{1}$$

$$\text{s.t. } Ax \le b \tag{2}$$

$$0 \le x \le d \tag{3}$$

$$x - \text{ integer} \tag{4}$$

where $x$ is an n-dimensional vector of variables, $d$ is an n-dimensional vector of variables upper bounds, A is an m x n matrix, b is the RHS vector and the vector $c^i\ (i = 1,.....k)$ represents the coefficients of the objective functions. Constraints (2)-(4) define the feasible set of the variables (solutions) $X_1$ . Problem (1)-(3) is a multiobjective linear programming (MOLP) problem, which is the weakened problem of a MOILP problem. The feasible set of the continuous solutions will be denoted by $X_2$ .

Let Z denotes the feasible region in the criteria space, i.e. the set of points $z \in \Re^k$ such that $z_i = f_i(x), i=1,...,k, x \in X_1$. $x^* \in X_1$ is an *efficient* or *Pareto optimal* solution if there is no $x \in X_1$ such that $c^i x \ge c^i x^*$ for all *i* and $c^i x > c^i x^*$ for at least one *j*. $x^{**} \in X_1$ is said to be *weakly efficient / Pareto optimal* solution if there is no $x \in X_1$ such that $c^i x > c^i x^{**}$ for all *i* . Although the term "*efficient*" is more often used for points *x* and the term *Pareto optimal* (PO)  for points *z*, they can be used interchangeably.

*Reference neighborhood* is the area around the currently preferred solution in the feasible criteria space of MOLP/MOILP problems, determined by the local preferences of the DM, in which the next currently preferred solution will be sought.

When solving a MOILP problem, the DM evaluates and compares the currently found (weak) PO solutions. In case he wants to look for a better solution, he/she sets his/her preferences for desired or feasible alterations of the values of a part or of all the criteria. Depending on these preferences, the criteria set can be implicitly divided into seven or less than seven classes: $K^>, K^\ge, K^=, K^<, K^\le, K^{><}$, and $K^0$. Every criterion $f_k(x),\ \ k \in K$ may belong to one of these classes: $K^>$ - improvement as a desired direction of change; $K^\ge$ - improvement by desired (aspiration) values $\Delta_{hj}$ ; $K^=$ - to either preserve or improve the current value of the criteria; $K^<$ - acceptable deterioration as a desired direction of change; $K^\le$ - acceptable deterioration by no more than $\delta_{hj}$ ; $K^{><}$ - the criteria value to lie within an interval, ( $a_{hj} - t_{hj}^- \le a_{hj} \le a_{hj} + t_{hj}^+$ ) around the current value $a_{hj}$ ; $K^0$ - the DM is indifferent about the value of these criteria and as such they may be altered freely.

On the basis of this criteria division by the DM, the following scalarizing problem, denoted as RNP1, is suggested for finding a PO solution of MOILP problem:

$$\min_{x\in X_1} T(x)=\min_{x\in X_1}\max\left(\begin{array}{l}\max\limits_{k\in K^{\geq}}\left(\dfrac{\bar{f}_k-f_k(x)}{\left|\bar{f}_k-f_k\right|}\right),\\ \max\limits_{k\in K^{\leq}}\left(\dfrac{\tilde{f}_k-f_k(x)}{\left|\tilde{f}_k-f_k\right|}\right)\end{array}\right)+\max\left(\begin{array}{l}\max\limits_{k\in K^{<}}\left(\dfrac{f_k-f_k(x)}{\left|f_k^*-f_k\right|}\right),\\ \max\limits_{k\in K^{>}}\left(\dfrac{f_k-f_k(x)}{\left|f_k^*-f_k\right|}\right)\end{array}\right)+\delta\left(\sum_{k\in K^{\geq}}\left(\bar{f}_k-f_k(x)\right)+\right. \tag{5}$$

$$\left.+\sum_{k\in K^{\leq}}\left(\tilde{f}_k-f_k(x)\right)+\sum_{k\in K^{<}\cup K^{>}}\left(f_k-f_k(x)\right)-\sum_{k\in K^{=}\cup K^{><}\cup K^{0}}f_k(x)\right)$$

$$\text{s.t. } f_k(x)\geq\tilde{f}_k,\quad k\in K^{>}\cup K^{><}\cup K^{=}\cup K^{\leq} \tag{6}$$

$$f_k(x)\leq\tilde{\tilde{f}}_k,\quad k\in K^{><} \tag{7}$$

$$x\in X_1, \tag{8}$$

where $\delta$ is an arbitrary small number, $f_k$ is the value of the criterion with an index $k\in K$ for the current preferred solution, $\overline{f_k}=f_k+\Delta_k$ is the desired (aspiration) level of the criterion with an index $k\in K^{\geq}$,

$$\tilde{f}_k=\begin{cases}f_k,\ \text{if } k\in K^{=}\cup K^{>},\\ f_k-t_k^{-},\ \text{if } k\in K^{><},\\ f_k-\delta_k,\ \text{if } k\in K^{\leq}.\end{cases}$$

$$\tilde{\tilde{f}}_k=f_k+t_k^{+},\ \text{if } k\in K^{><}$$

The current preferred solution of MOILP problem is a feasible solution of the current scalarizing problem RNP1, i.e., the scalarizing problem RNP1 has an initial feasible solution. This is a very important feature, because finding a feasible solution of integer problems is also a NP-problem. In addition, the feasible solutions of the scalarizing problem RNP1 are located near to the non-dominated surface of the multicriteria problem in the criteria space Z. They belong to the reference area, defined by DM's preferences.

**Theorem 1**: The optimal solution of the scalarizing problem RNP1 is an efficient/PO solution of MOILP problem.

**Proof:** The scalarizing problem RNP1 is solved, when the DM wants improvement with respect to one criterion at least, or when $K^{\geq}\neq\varnothing$ or $K^{>}\neq\varnothing$.

Let $x^*$ be the optimal solution of the scalarizing problem RNP1. Then the following conditions are satisfied:

$$T(x^*)\leq T(x),\quad x\in X_1, \tag{9}$$

and constraints (6)-(7).

Let us assume that $x^*\in X_1$ is not an efficient/PO solution of the initial MOILP problem. In this case there must exist another $x'\in X_1$, that is an efficient/PO solution of MOILP problem, for which:

$$f_k(x')\geq f_k(x^*)\text{ , for } k\in K\text{ , } f_k(x')>f_k(x^*)\text{ for at least one } k\in K \tag{10}$$

and constraints (6)-(7) are satisfied:

After transformation of the objective function $T(x)$ of the scalarizing problem RNP1, using inequalities (10), the following relation is obtained:

$$(11)\quad T(x') = \max\left( \begin{array}{l} \max\limits_{k\in K^{\geq}}\left(\dfrac{\bar{f}_k - f_k(x')}{|\bar{f}_k - f_k|}\right), \\ \max\limits_{k\in K^{\leq}}\left(\dfrac{\tilde{f}_k - f_k(x')}{|\tilde{f}_k - f_k|}\right) \end{array} \right) + \max\left( \begin{array}{l} \max\limits_{k\in K^{<}}\left(\dfrac{f_k - f_k(x')}{|f_k^* - f_k|}\right), \\ \max\limits_{k\in K^{>}}\left(\dfrac{f_k - f_k(x')}{|f_k^* - f_k|}\right) \end{array} \right) + \delta\left( \sum_{k\in K^{\geq}}(\bar{f}_k - f_k(x')) + \right.$$

$$\left. + \sum_{k\in K^{\leq}}(\tilde{f}_k - f_k(x')) + \sum_{k\in K^{<}\cup K^{>}}(f_k - f_k(x')) - \sum_{k\in K^{=}\cup K^{><}\cup K^{0}} f_k(x') \right) =$$

$$= \max\left( \max_{k\in K^{\geq}}\left( \frac{\bar{f}_k - f_k(x^*)}{|\bar{f}_k - f_k|} + \frac{f_k(x^*) - f_k(x')}{|\bar{f}_k - f_k|} \right), \max_{k\in K^{\leq}}\left( \frac{\tilde{f}_k - f_k(x^*)}{|\tilde{f}_k - f_k|} + \frac{f_k(x^*) - f_k(x')}{|\tilde{f}_k - f_k|} \right) \right) + \qquad (11)$$

$$+ \max\left( \max_{k\in K^{<}}\left( \frac{f_k - f_k(x^*)}{|f_k^* - f_k|} + \frac{f_k(x^*) - f_k(x')}{|f_k^* - f_k|} \right), \max_{k\in K^{>}}\left( \frac{f_k - f_k(x^*)}{|f_k^* - f_k|} + \frac{f_k(x^*) - f_k(x')}{|f_k^* - f_k|} \right) \right) +$$

$$+ \delta\left( \sum_{k\in K^{\geq}}\left((\bar{f}_k - f_k(x^*)) + (f_k(x^*) - f_k(x'))\right) + \sum_{k\in K^{\leq}}\left((\tilde{f}_k - f_k(x^*)) + (f_k(x^*) - f_k(x'))\right) + \right.$$

$$\left. + \sum_{k\in K^{<}\cup K^{>}}\left((f_k - f_k(x^*)) + (f_k(x^*) - f_k(x'))\right) - \sum_{k\in K^{=}\cup K^{><}\cup K^{0}} f_k(x^*) + (f_k(x') - f_k(x^*)) \right) <$$

$$< \max\left( \begin{array}{l} \max\limits_{k\in K^{\geq}}\left(\dfrac{\bar{f}_k - f_k(x^*)}{|\bar{f}_k - f_k|}\right), \\ \max\limits_{k\in K^{\leq}}\left(\dfrac{\tilde{f}_k - f_k(x^*)}{|\tilde{f}_k - f_k|}\right) \end{array} \right) + \max\left( \begin{array}{l} \max\limits_{k\in K^{<}}\left(\dfrac{f_k - f_k(x^*)}{|f_k^* - f_k|}\right), \\ \max\limits_{k\in K^{>}}\left(\dfrac{f_k - f_k(x^*)}{|f_k^* - f_k|}\right) \end{array} \right) + \delta\left( \sum_{k\in K^{\geq}}(\bar{f}_k - f_k(x^*)) + \right.$$

$$\left. + \sum_{k\in K^{\leq}}(\tilde{f}_k - f_k(x^*)) + \sum_{k\in K^{<}\cup K^{>}}(f_k - f_k(x^*)) - \sum_{k\in K^{=}\cup K^{><}\cup K^{0}} f_k(x^*) \right) = T(x^*)$$

It follows from (11) that $T(x') < T(x^*)$ and constraints (6)-(7), that contradicts to (9). Hence $x^*$ is an efficient solution, and $f(x^*)$ is a PO solution in the criteria space of MOILP problem.

In order to find a PO solution of MOLP problem, scalarizing problem RNP1 may be used, in which the constraint requiring integers (8) is removed. The relaxing problem obtained is denoted as RNP1-L.

Since the objective functions of the scalarizing problems RNP1 and RNP1-L are non-differentiable, each one of them may be reduced to the equivalent optimization problem, on the account of additional variables and constraints, but with a differentiable objective function [7]. The equivalent linear integer problem of RNP1 problem, denoted as RNP1e, can be presented as follows:

$$\min_{x\in X_1}\left(\alpha + \beta + \delta\sum_{k\in K} y_k\right) \qquad (12)$$

$$\text{s.t. } \alpha \geq (\bar{f}_k - f_k(x))/|\bar{f}_k - f_k|,\ k \in K^{\geq} \qquad (13)$$

$$\alpha \geq (\tilde{f}_k - f_k(x))/|\tilde{f}_k - f_k|,\ k \in K^{\leq} \qquad (14)$$

$$\beta \geq (f_k - f_k(x))/|f_k^* - f_k|,\ k \in K^{<} \qquad (15)$$

$$\beta \geq (f_k - f_k(x))/|f_k^* - f_k|,\ k \in K^{>} \qquad (16)$$

$$\bar{f}_k - f_k(x) = y_k,\ k \in K^{\geq} \qquad (17)$$

$$\tilde{f}_k - f_k(x) = y_k, \ \ k \in K^{\leq} \tag{18}$$

$$f_k - f_k(x) = y_k, \quad k \in K^{<} \cup K^{>} \tag{19}$$

$$f_k(x) = y_k, \quad k \in K^{=} \cup K^{0} \cup K^{><} \tag{20}$$

$$f_k(x) \geq \tilde{f}_k, \qquad k \in K^{>} \cup K^{><} \cup K^{=} \cup K^{\leq} \tag{21}$$

$$f_k(x) \leq \tilde{\tilde{f}}_k, \qquad k \in K^{><} \tag{22}$$

$$x \in X_1 \tag{23}$$

$$\alpha, \beta\ , y_k\ , \ k \in K \ \text{ - arbitrary.} \tag{24}$$

Problems RNP1 and RNP1e have one and the same feasible set of solutions, and the values of their objective functions are equal. This comes from the following affirmation:

**Theorem 2:** The optimal values of scalarizing problems RNP1 and RNP1e are equal, i.e.

$$\min_{x \in X_1}(\alpha + \beta + \delta \sum_{k \in K} y_k) = \min_{x \in X_1} \left( \begin{array}{c} \max \left( \begin{array}{l} \max\limits_{k \in K^{\geq}} \left( \dfrac{\bar{f}_k - f_k(x)}{\left|\bar{f}_k - f_k\right|} \right), \\ \max\limits_{k \in K^{\leq}} \left( \dfrac{\tilde{f}_k - f_k(x)}{\left|\tilde{f}_k - f_k\right|} \right) \end{array} \right) + \max \left( \begin{array}{l} \max\limits_{k \in K^{<}} \left( \dfrac{f_k - f_k(x)}{\left|f_k^* - f_k\right|} \right), \\ \max\limits_{k \in K^{>}} \left( \dfrac{f_k - f_k(x)}{\left|f_k^* - f_k\right|} \right) \end{array} \right) + \delta \left( \sum\limits_{k \in K^{\geq}} \left(\bar{f}_k - f_k(x)\right) + \right. \\ \left. + \sum\limits_{k \in K^{\leq}} \left(\tilde{f}_k - f_k(x)\right) + \sum\limits_{k \in K^{<} \cup K^{>}} \left(f_k - f_k(x)\right) - \sum\limits_{k \in K^{=} \cup K^{><} \cup K^{0}} f_k(x) \right) \end{array} \right).$$

**Proof:** It follows from (13) that $\alpha \geq \left(\overline{f}_k - f_k(x)\right) / \left|\overline{f}_k - f_k\right|, k \in K^{\geq}$. Since this inequality is true for every $k \in K^{\geq}$, then it is also true that

$$\alpha \geq \max_{k \in K^{\geq}} \left(\overline{f}_k - f_k(x)\right) / \left|\overline{f}_k - f_k\right| \tag{25}$$

From (14) it could be noticed that $\alpha \geq \left(\tilde{f}_k - f_k(x)\right) / \left|\tilde{f}_k - f_k\right|, k \in K^{\leq}$. Since this inequality is true for every $k \in K^{\leq}$, then it is also true that

$$\alpha \geq \max_{k \in K^{\leq}} \left(\tilde{f}_k - f_k(x)\right) / \left|\tilde{f}_k - f_k\right| \tag{26}$$

Using (25) and (26) the following is derived:

$$\alpha \geq \max \ \left( \max_{k \in K^{\geq}} \left(\overline{f}_k - f_k(x)\right) / \left|\overline{f}_k - f_k\right|, \max_{k \in K^{\leq}} \left(\tilde{f}_k - f_k(x)\right) / \left|\tilde{f}_k - f_k\right| \right) \tag{27}$$

In the same way it follows from (15):

$$\beta \geq \max_{k \in K^{<}} \left( (f_k - f_k(x)) / \left|f_k^* - f_k\right| \right), \tag{28}$$

and from (16):

$$\beta \geq \max_{k \in K^{>}} \left( (f_k - f_k(x)) / \left|f_k^* - f_k\right| \right), \tag{29}$$

From (28) and (29) it can be written for $\beta$ :

$$\beta \geq \max \ \left( \max_{k \in K^{<}} \left( (f_k - f_k(x)) / \left|f_k^* - f_k\right| \right), \max_{k \in K^{>}} \left( (f_k - f_k(x)) / \left|f_k^* - f_k\right| \right) \right) \tag{30}$$

In case the left and right sides of inequalities (27) and (30) are summed up, the following relation is obtained:

$$(\alpha+\beta)\geq \max\left(\max_{k\in K^{\geq}}\left(\overline{f}_k-f_k(x)\right)/\left|\overline{f}_k-f_k\right|,\ \max_{k\in K^{\leq}}\left(\widetilde{f}_k-f_k(x)\right)/\left|\widetilde{f}_k-f_k\right|\right)+ \\ +\max\left(\max_{k\in K^{<}}\left(\left(f_k-f_k(x)\right)/\left|f_k^*-f_k\right|\right),\max_{k\in K^{>}}\left(\left(f_k-f_k(x)\right)/\left|f_k^*-f_k\right|\right)\right) \tag{31}$$

If the term $\delta\sum_{k\in K}y_k$ is added to both sides of (31), the following inequality is obtained:

$$(\alpha+\beta+\delta\sum_{k\in K}y_k)\geq \max\left(\max_{k\in K^{\geq}}\left(\overline{f}_k-f_k(x)\right)/\left|\overline{f}_k-f_k\right|,\ \max_{k\in K^{\leq}}\left(\widetilde{f}_k-f_k(x)\right)/\left|\widetilde{f}_k-f_k\right|\right)+ \\ +\max\left(\max_{k\in K^{<}}\left(\left(f_k-f_k(x)\right)/\left|f_k^*-f_k\right|\right),\max_{k\in K^{>}}\left(\left(f_k-f_k(x)\right)/\left|f_k^*-f_k\right|\right)\right)+\delta\sum_{k\in K}y_k \tag{32}$$

Let $x^*$ be the optimal solution of RNP1e problem. Then

$$\min_{x\in X_1}(\alpha+\beta+\delta\sum_{k\in K}y_k)=\min_{x\in X_1}\left\{\begin{array}{l}\max\left(\max_{k\in K^{\geq}}\left(\dfrac{\overline{f}_k-f_k(x^*)}{\left|\overline{f}_k-f_k\right|}\right),\max_{k\in K^{\leq}}\left(\dfrac{\widetilde{f}_k-f_k(x^*)}{\left|\widetilde{f}_k-f_k\right|}\right)\right)+ \\ +\max\left(\max_{k\in K^{<}}\left(\dfrac{f_k-f_k(x^*)}{\left|f_k^*-f_k\right|}\right),\max_{k\in K^{>}}\left(\dfrac{f_k-f_k(x^*)}{\left|f_k^*-f_k\right|}\right)\right)+\delta\sum_{k\in K}y_k\end{array}\right\} \tag{33}$$

for in the opposite case $\left(\alpha+\beta+\delta\sum_{k\in K}y_k\right)$ could be still decreased. Since after summing up inequalities (17-20) the relation given below is obtained

$$\left(\sum_{k\in K^{\geq}}\left(\bar{f}_k-f_k(x)\right)+\sum_{k\in K^{\leq}}\left(\widetilde{f}_k-f_k(x)\right)+\sum_{k\in K^{<}\cup K^{>}}\left(f_k-f_k(x)\right)-\sum_{k\in K^{=}\cup K^{><}\cup K^{0}}f_k(x)\right)=\sum_{k\in K}y_k\ ,$$

then the right side of (33) can be written down also as:

$$\min_{x\in X_1}(\alpha+\beta+\delta\sum_{k\in K}y_k)=\min_{x\in X_1}\left(\begin{array}{l}\max\left(\begin{array}{l}\max_{k\in K^{\geq}}\left(\dfrac{\bar{f}_k-f_k(x)}{\left|\bar{f}_k-f_k\right|}\right), \\ \max_{k\in K^{\leq}}\left(\dfrac{\widetilde{f}_k-f_k(x)}{\left|\widetilde{f}_k-f_k\right|}\right)\end{array}\right)+\max\left(\begin{array}{l}\max_{k\in K^{<}}\left(\dfrac{f_k-f_k(x)}{\left|f_k^*-f_k\right|}\right), \\ \max_{k\in K^{>}}\left(\dfrac{f_k-f_k(x)}{\left|f_k^*-f_k\right|}\right)\end{array}\right)+\delta\left(\sum_{k\in K^{\geq}}\left(\bar{f}_k-f_k(x')\right)+\right. \\ \left.+\sum_{k\in K^{\leq}}\left(\widetilde{f}_k-f_k(x')\right)+\sum_{k\in K^{<}\cup K^{>}}\left(f_k-f_k(x')\right)-\sum_{k\in K^{=}\cup K^{><}\cup K^{0}}f_k(x')\right)\end{array}\right),$$

which proves the theorem.

The continuous scalarizing problem RNP1-L has a similar equivalent problem of linear programming RNP1-Le, in which the feasible set of the variables $x\in X_2$ is expanded. In order to find more than one weak PO solutions of the continuous scalarizing problem RNP1-Le, its parametric extension, called RNP1-Lp, may be applied:

$$\min\left(\alpha+\beta\right)$$

s.t.

$$f_k(x)+\left|\overline{f}_k-f_k\right|\alpha\geq f_k+(\overline{f_k}-f_k)t,\quad k\in K^{\geq},$$

$$f_k(x)+\left|\widetilde{f}_k-f_k\right|\alpha\geq f_k+(\widetilde{f}_k-f_k)t,\ k\in K^{\leq}$$

$f_k(x)+\left|f_k^*-f_k\right|\beta \geq f_k - t, \quad k \in K^{<}$ ,

$f_k(x)+\left|f_k^*-f_k\right|\beta \geq f_k + t, \quad k \in K^{>}$ ,

$f_k(x) \geq \tilde{f}_k, \quad k \in K^{>} \cup K^{><} \cup K^{=} \cup K^{\leq}$ ,

$f_k(x) \leq \tilde{\tilde{f}}_k, \quad k \in K^{><}$

$x \in X_2$ , $t \geq 0$ , $\alpha, \beta$ - arbitrary.

With the help of this parametric problem more new (weak) PO solutions of MOLP problem are sought, shifting the reference neighbourhood to direction of the preferences, set by the DM for desired improvement and acceptable deterioration of some of the criteria. In thus way it will not be necessary for the DM to make one more iteration in direction of improvement of the selected criteria, in order to acquire knowledge about the compromises that he has to accept with respect to the rest of the criteria. Let us assume that a (weak) PO solution of MOLP problem has been found with the help of scalarizing problems RNP1-Le or RNP1-Lp, evaluated by the DM as satisfactory. Let it be denoted by $\hat{f}=(\hat{f}_1,...,\hat{f}_p)^T$ . In order to find a (weak) Pareto optimal solution of MOILP problem, close to the solution $\hat{f}_k$ , the following Tchebycheff's problem RNP3 [7] may be used:

$$\min_{x \in X_1} S(x) = \min_{x \in X_1} \max_{k \in K}(\hat{f}_k - f_k(x)) / | \hat{f}_k^{'} |,$$

where

$\hat{f}_k^{'}$ = $\begin{cases} \hat{f}_k, & \text{if } |\hat{f}_k| > \varepsilon, \\ \varepsilon, & \text{if } |\hat{f}_k| \leq \varepsilon. \end{cases}$ , $\varepsilon$ is a small positive number.

This problem is equivalent to the following problem of mixed integer programming RNP3e:

$$\min \alpha$$

s.t.

$$\alpha \geq (\hat{f}_k - f_k(x)) / | \hat{f}_k^{'} |,$$

$$x \in X_1,$$

MOILP problems belong to the class of NP-problems. The computational efforts, connected with optimal solution finding in them depend considerably on the form of the feasible solutions area. That is why the search for new scalarizing problems, formulated in a way that could direct faster to the optimal solution, close to DM's preferences, is a constant task which all the researchers face.

## Basic properties of the scalarizing problems of the reference neighborhood

The scalarizing problems of the reference neighborhood (RNP) are formulated on the basis of inexplicit classification of the criteria in accordance with DM's preferences, aimed at improvement of the current preferred solution. Similar scalarizing problems have been offered in [3], [4], [5], [6], which also use inexplicit criteria division into groups. An open communication protocol to interact with the DM, which enables free exploration of the problem and progressive learning of the non-dominated solution set [2], is applied in the interactive algorithms, based on classification-oriented scalarizing problems. Four features of RNP problems facilitate the dialogue with the DM with respect to the information required about his/her local preferences, with respect to decrease of the time the DM has to wait for finding a new solution and also with respect to easing DM's efforts in the evaluation of more than one solutions:

- the DM has variable possibilities to express his/her preferences, concerning the alteration of the values of some or all the criteria with respect to the current solution found. He/she may apply the most appropriate and comprehensible approach for every criterion, setting levels of attainability or a feasible trade-off, define the direction of improvement only, or specify the range of a feasible alteration;
- the current preferred solution of MOILP problems, found at the previous iteration, is a feasible solution of the integer scalarizing problems RNP1e, which are to be solved at the next iteration. This reduces considerably the computational effort, since the discovery of a feasible solution of a single-criterion integer problem is an NP-problem;
- the feasible solutions of the scalarizing problems of the reference neighborhood are located near to PO surface of MOILP problems. The feasible area of RNP scalarizing problems is a part of the feasible area of MOILP problems, unlike the feasible area of the scalarizing problems of the reference point, which coincides with it. Depending on the way, in which the DM sets his/her preferences (by values of desired improvement and acceptable deterioration), this feasible area may be comparatively narrow and the feasible solutions in the criteria space, found with the help of approximate algorithms of integer programming, might be located close to the non-dominated surface of MOILP problems. The use of approximate weak PO solutions leads to considerable decrease of the time duration, when the DM is expecting a new solution for evaluation and choice. In this way, on the account of slight worsening in the quality of the solutions obtained, the dialogue with the DM can be improved;
- the strategy “not high profits - small losses” is realized. This is achieved, since the optimal solution of RNP scalarizing problems aspires to minimize the maximal Tchebycheff distance to the current preferred solution, both in direction of an improvement and also in direction of a compromise deterioration of the criteria, determined by DM’s preferences. The solutions from the reference neighborhood, that are obtained, are comparatively close, which facilitates the DM in his/her comparing of several solutions and also in selecting of the new currently preferred solution;
- the application of the interactive algorithm in RNP1Lp scalarizing problem will result in finding more than one (weak) PO solutions of MOLP problems. It is obvious, that in case the DM evaluates at each iteration more (weak) PO solutions, then he/she will learn faster in problem specifics and will find quicker the most preferred solution to MOILP problems.

## Conclusion

In the area of multicriteria decision making the interactive approach has found wide application in a specific and well defined class of algorithms, which aid the DM in the study of a set of solutions with the purpose to select one or a limited set from them. These algorithms, built on the basis of open communication protocols with the DM, prove to be some of the most promising ways of research to develop adequate MOILP tools for decision aiding in many complex practical situations [1]. Undoubtedly, in recent years the interest towards explicit multiple objectives inclusion in different real life application areas of integer programming models, has grown significantly.

The quality of the interactive algorithms depends mainly on the scalarizing problems they are designed and based on. The properties of the scalarizing problems of the reference neighborhood herein proposed make them especially appropriate in realizing a “continuous-integer” approach in interactive algorithms for solving a general MOILP problem. Instead of generating integer solutions for evaluation at each iteration, the DM may evaluate linear continuous solutions at most of the iterations. In the criteria space the PO integer solutions are placed relatively close to PO continuous solutions. Thus, in the learning phase, the DM could be trained on the basis of linear continuous solutions, instead on the basis of integer solutions, which is particularly important for MOILP problems with large dimensions. In addition, the DM may also learn on the basis of approximate weak PO solutions, found relatively near to a weak PO surface. Another advantage of this class of scalarizing problems is that in them the DM operates in the criteria space. In real life problems the criteria have their quantitative and financial aspects, so that the DM is able through their definition to express easier his/her preferences in the search for a trade-off solution.

## Bibliography

[1] Alves M.J. and Climaco J. A review of interactive methods for multiobjective integer and mixed-integer programming. European Journal of Operational Research, 180, 99-115, 2007.

[2] Alves M.J. and Climaco J. A note on a decision support system for multiobjective integer and mixed-integer programming problems. European Journal of Operational Research, 155, 258-265, 2004.

[3] Narula, S. C. and Vassilev, V. An Interactive Algorithm for Solving Multiple Objective Integer Linear Programming Problems. European Journal of Operational Research, 79, 443-450, 1994.

[4] Vassilev V., Genova K., Vassileva M., Narula S. Classification-Based Method of Linear Multicriteria Optimization. International Journal on Information Theories and Applications, 10, 3, 266-270, 2003.

[5] Vassilev V., Narula S. and Gouljaski V. An Interactive Reference Directions Algorithm for Solving Multiobjective Convex Nonlinear Integer Programming Problems, International Transactions in Operational Research, 8, 367-380, 2001.

[6] Vassileva, M. A Learning –oriented Method of Linear Mixed Integer Multicriteria Optimization, Cybernetic and Information Technologies, 1,13-25, 2004.

[7] Wolsey, L. A.. Integer Programming, Wiley-InterScience, 1998.

## Acknowledgements

The paper is partially supported by the Bulgarian National Science Fund, Grant No DTK02/71 "Web-Based Interactive System, Supporting the Building Models and Solving Optimization and Decision Making Problems" and by IIT-BAS by the project "Research, development and application of effective methods for solving NP-hard single and multiobjective programming problems".

## Authors' Information

**Mariana Vassileva –** Assoc.Prof., PhD, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria; e-mail: mvassileva@iinf.bas.bg

**Krasimira Genova -** Assoc.Prof., PhD, Institute of Information Technologies, BAS, Acad. G. Bonchev St., bl. 29A, Sofia 1113, Bulgaria; e-mail: kgenova@iinf.bas.bg