A Generalized Framework for Multi-Criteria Classifiers with Automated Learning: Application on FLIR Ship Imagery

KHALED JABEUR
ADEL GUITOUNI

This paper reviews Multi-Criteria Classifiers (MCCs) or commonly multi-criteria classification methods. These methods have many advantages including flexibility, the integration of human judgments and prevention of black box syndrome. However, these advantages come with a price: large number of parameters to be setup. In particular, this paper focuses on Nominal Concordance/Discordance-based MCCs (NCD-MCCs). A generalized framework is proposed to synthesize the underlying computation algorithm for each MCC. In order to address MCCs disadvantages, an Automated Learning Method (ALM) based on Real-Coded Genetic Algorithm (RCGA) is proposed to infer these parameters. The empirical results of some MCCs are compared with those obtained by other classifiers (e.g. Bayes and Dempster-Shafer classifiers). A military dataset of 2545 Forward Looking Infra-Red (FLIR) images representing eight different classes of ships is therefore used to test the performance of these classifiers. In this paper, we argue the benefits of cross-fertilization of MCCs and information fusion algorithms.

Manuscript received October 2, 2007; revised March 6, 2009; released for publication March 26, 2009.

Refereeing of this contribution was handled by Dr. Pierre Valin.

Authors' addresses: kK. Jabeur, Institut Supérieur de Commerce et de Comptabilité de Bizerte (ISCCB), 7021 Zarzouna Bizerte—Tunisia, E-mail: (khaled.jabeur@fsa.ulaval.ca); A. Guitouni, Defence R&D Canada Valcartier, 2459, Pie-XI Blvd North, Val-Bélair, QC, Canada, G3J 1X5, E-mail: (adel.guitouni@drdc-rddc.gc.ca).

1557-6418/09/$17.00

1. INTRODUCTION

Supervised classification often consists in assigning a set of entities (e.g. alternatives, images, projects, subjects) into pre-defined and homogeneous categories. Categories are known a priori either by defining profiles limit between them or by a set of typical profiles (reference prototypes or elements) for each category. Ordinal Classification (or Sorting) usually refers to an order relationship between the categories, and nominal classification otherwise. Recently, a variety of classification methods—based on Artificial Intelligence (AI) and Operations Research (OR) techniques—have been proposed to solve classification problems [41]. Neural Networks (NN), Machine Learning (ML), Rough Sets (RS), Fuzzy Sets (FS) and Multi-Criteria Decision Analysis (MCDA) were used for the development and the validation of these methods. This paper focuses on classification methods based on MCDA methodology.

In this paper we use Multi-Criteria Classifiers (MCCs) to designate supervised classification methods based on MCDA methodology. The most MCCs are based on either outranking or multi-attribute utility approaches. Roy and Moscarola [35], Masaglia and Ostanello [24], Yu [42], Perny [31], Belacel [3] and Henriet [15] have proposed MCCs based on the outranking approach, while M.H.DIS (Multi-group Hierarchical DIScrimination) method [40] and UTADIS (UTilités Additives DIScriminantes) method and its variants ([21], [39], [10]) are typical methods based on multi-attribute utility theory. This paper focuses essentially on outranking-based nominal MCCs where there is no order relationship between the categories. These MCCs are based on concordance/discordance concepts.

Limitation of outranking-based methods is due to the large number of parameters (e.g. discrimination thresholds, weights, reference alternatives, etc.) required. In MCDA context, these parameters are generally elicited using interactive approaches from the decision-maker to articulate his relational preference system: it's the Direct Elicitation Approach (DEA). However, it is difficult for the decision-maker to provide such information in a coherent way when the number of these parameters is considerable. Indirect Elicitation Approach (IEA) or Automatic Learning Methods (ALMs) might be the solution to elicit automatically the values of these parameters based on a training set of pre-assigned examples. These two elicitation approaches will be discussed in Section 3.

This paper makes three main contributions. First, we propose a generalized framework for Nominal Concordance/Discordance-based MCC (NCD-based MCCs). Second, we develop an ALM based on Real-Coded Genetic Algorithm (RCGA) to estimate the parameters of NCD-based MCCs. Then we illustrate and assess the performance of the proposed approach on selected NCD-based MCCs. Even if the purpose of the comparison might be seen limited, we present exper-

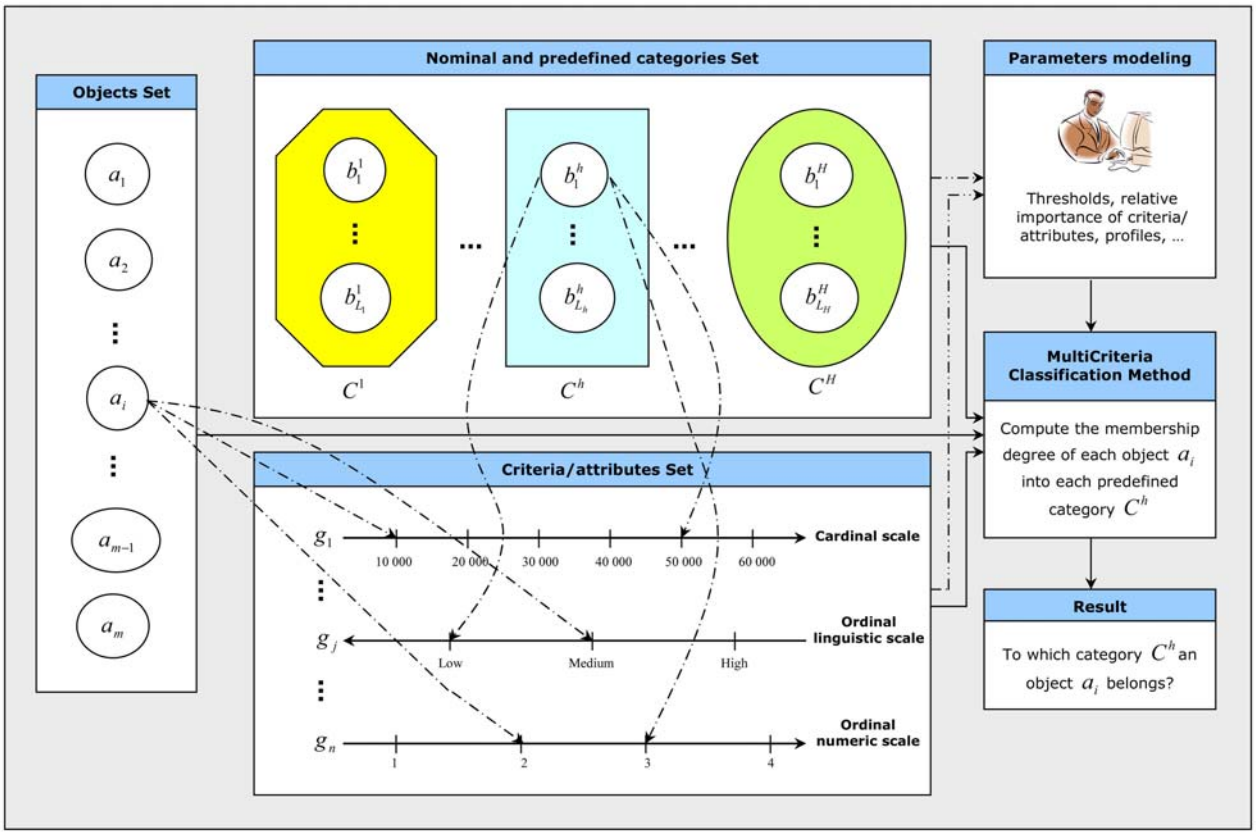


Fig. 1. NCD-based MCCs concepts.

imental results by comparing NCD-based MCCs with other classifiers such as Bayes and Dempster-Shafer classifiers. A dataset of 2545 Forward Looking Infra-Red (FLIR) images representing eight different classes of ships is used for the empirical validation.

This paper is organized as follows. Section 2 presents a generalized framework for NCD-based MCCs. Section 3 proposes an ALM based on RCGA to infer the parameters of the NCD-based MCCs. Section 4 presents a brief description of the dataset used to test the performance of the different classifiers. In Section 5, computational results of some NCD-based MCCs are presented and compared with those obtained by other classifiers. Finally, conclusions, discussions and future works are presented in Section 6.

2. A GENERALIZED FRAMEWORK FOR NCD-BASED MCCs

In our opinion, there are at least three major advantages which distinguish MCCs from the other classifiers:

1. MCCs are designed to incorporate objective and subjective information and deal "*correctly*" with quantitative and qualitative data. In fact, it is possible to take into account human judgments and compute information obtained on conflicting and heterogeneous dimensions [10]. Therefore, these methods are essential when it is important to explicitly integrate human judgments (decision-maker's preferences), to consider many conflicting criteria, and to deal with data obtained on heterogeneous measurement scales (see Fig. 1);
2. MCCs allow pairwise comparisons between the objects to be assigned and the profiles (or reference elements). The pairwise comparison might be seen as a projection isomorphism for each pair of alternatives from the attributes spaces to the preferences spaces. The result of the comparison is a valued function between each pair of alternatives/elements. The aggregation and exploitation of these valued functions avoid computing distance measures obtained on heterogeneous measurement scales (as in K-NN classifier) and allow handling qualitative and/or quantitative information;
3. MCCs avoid the black box situation: it is easy to explain the classification result in natural language.

The advantages or strengths of MCCs could also be seen as weaknesses if subjective human judgments are to be avoided and the information is not heterogeneous and highly correlated. Moreover, these methods require a quite large set of parameters to be determined, which could also be seen as a strength (offering many degrees of freedom). To overcome these limitations, an IEA may be used to infer automatically these parameters. To our knowledge, there exist in the literature four NCD-based MCCs: (i) TRI-NOMFC classifier [23], (ii) PROAFTN

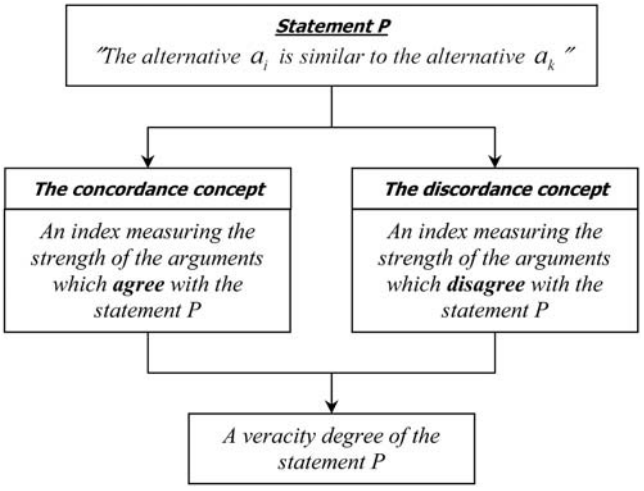


Fig. 2. Indifference relation principle.

classifier [3], (iii) PIP and K-PIP classifiers [15] and (iv) FBI classifier [31].[1] The concordance and discordance concepts were introduced by Bernard Roy [32] when he defined the outranking relation in ELECTRE I method. Indeed, according to many authors ([7], [33], [38]), an outranking relationship can be defined as follows: *an alternative a_i outranks an alternative a_k if and only if there are enough arguments to decide that a_i is at least as good as a_k (concordance concept) while there is no essential reason to refute that statement (discordance concept).* Recently, Perny [31], Belacel [3] and Henriet [15] extended the above definition by developing an indifference relation measuring the similarity between two alternatives a_i and a_k. The main idea of this relation is to compare the strength of the arguments supporting a proposition to the strength of the arguments opposing the same proposition (see Fig. 2). In fact, it's on the basis of this relation that NCD-based MCCs assign an object (project, alternative, image, etc.) into a predefined category. Hence, according to the NCD-based MCCs, assigning an object consists in computing its membership degree into a predefined category. Membership degree is an aggregate of the evaluation of the indifference between the object to be assigned and each prototype characterizing each category. Each indifference evaluation considers parameters such as discrimination thresholds and criteria/attributes weights.

Let's consider the following notations:

—Let $A = \{a_i\}_{i=1\ldots m}$ be a set of m objects to be assigned;
—Let $C = \{C^h\}_{h=1,2,\ldots,H}$ be a set of H nominal and predefined categories or classes. Each category C^h is characterized by a set of profiles or reference objects $B^h = \{b_k^h\}_{k=1\ldots L_h}$. The set of all profiles is noted by $B = \bigcup_{h=1}^{H} B^h$;
—Let $F = \{g_j\}_{j=1\ldots n}$ be a set of n criteria. We assume that the criteria are to be maximized (transforming a minimization to a maximization is a straight forward). To each criterion g_j, we assign a weight w_j^h ($j = 1\ldots n$) which expresses its relative importance in the category C^h;
—Each object a_i (respectively profile b_k^h) is evaluated on all criteria by the vector: $\mathbf{a}_i = (g_1(a_i), g_2(a_i), \ldots, g_n(a_i))$ (respectively by $\mathbf{b}_k^h = (g_1(b_k^h), g_2(b_k^h), \ldots, g_n(b_K^h))$).

Most NCD-based MCCs compute for each object a_i a fuzzy number called $\mu(a_i, C^h) \in [0,1]$, which measures the membership degree of a_i to a given category or class C^h. Hence, if $\mu(a_i, C^h) = 1$ we say that a_i belongs perfectly to the category C^h. However, when $\mu(a_i, C^h) = 0$, we say that a_i has no common ground with the category C^h. Fig. 1 presents a functioning schema of NCD-based MCCs. Fig. 3 shows a generalized framework used by NCD-based MCCs to compute $\mu(a_i, C^h)$.

The membership degree of an object a_i to a category C^h is computed using the concordance and discordance concepts. A local concordance $C_j(a_i, b_k^h)$ and discordance $D_j(a_i, b_k^h)$ indices are computed for each object a_i to be assigned, for each criterion g_j ($j = 1\ldots n$) and for each profile b_k^h ($k = 1\ldots L_h$) characteristic of C^h.[2] For instance, PROAFTN method [3] proposes linear functions—similar to those used in ELECTRE III method [34]—for the local concordance and discordance indices (see Fig. 4). In TRINOMFC method [23], only local concordance indices are computed using criteria functions similar to those used in PROMÉTHÉE method [6]. According to [23], it's not appropriate to consider discordance concept when dealing with nominal classification. The computation of local concordance and discordance indices of all NCD-based MCCs are summarised in Tables I and II (see pages 80 and 81).

The computation of local concordance and discordance indices is based on the following types of thresholds: indifference, preference and veto thresholds. The indifference (q_j) and the preference (p_j) thresholds are used to nuance the distinction between two objects into weak and strong preference relationships. The veto threshold (v_j) represents the limit of the tolerance for partial compensation between evaluations. In other word, if the evaluation of a_k is at least v_j different than the evaluation of a_i on criterion g_j, then we may refuse/confirm some propositions about a_i and a_k without regarding their evaluations on the other criteria. Note that these thresholds are established: i) for each criterion in FBI classifier, ii) for each criterion and category in PIP, K-PIP and TRI-NOMFC classifiers and iii) for each criterion, category and profile in PROAFTN classifier. The aggregation operators η_C and η_D are used in order to compute respectively the global concordance and discordance indices (see Ta-

[1] Figueira *et al.* [11] have recently proposed an extension of PROMETHEE method for classification purposes. This method will not be considered in this paper because, according to these authors, some improvements must be done to finalize it.

[2] In PROAFTN classifier [3] each profile b_k^h is defined, for each criterion g_j, by an interval $[S_j^1(b_k^h), S_j^2(b_k^h)]$ which is an exception with regards the other MCCs.

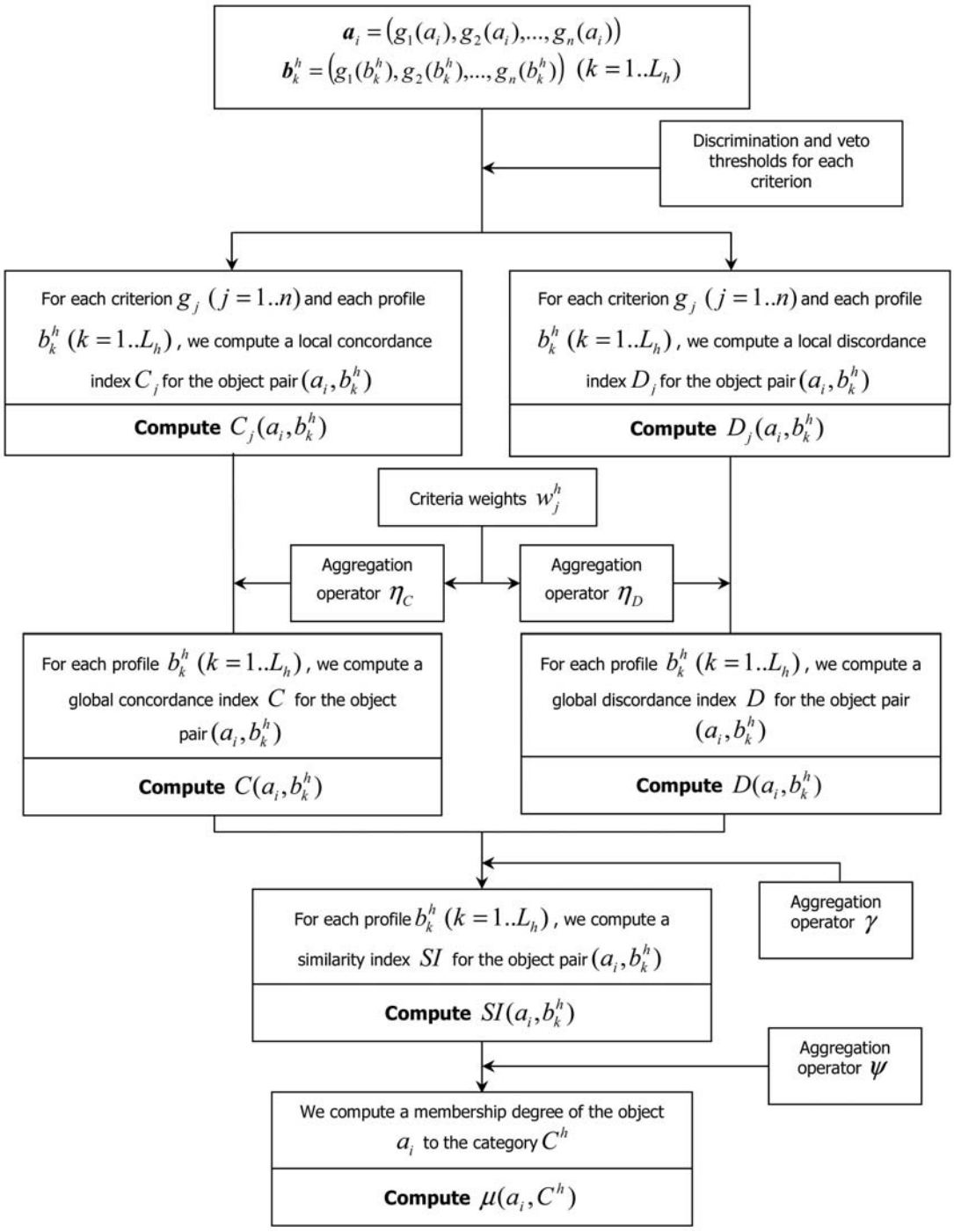


Fig. 3. Generalized framework of NCD-based MCCs to compute $\mu(a_i, C^h)$.

bles I and II). For all NCD-based MCCs, the operator η_C (respectively η_D) is often equal to the weighted sum (respectively weighted geometric mean). The computation of these global indices takes into account the criteria weights. For PROAFTN, TRI-NOMFC, PIP and K-PIP classifiers these weights are determined for each criterion and category whereas for FBI classifier these weights are computed only for each criterion.

The similarity index *SI* for each pair of objects (a_i, b_k^h) is computed as shown in Table III (see page 81). In general, two kinds of aggregation operator γ are used to combine the quantities $C(a_i, b_k^h)$ and $(1 - D(a_i, b_k^h))$: the *Product* and the *Minimum.* Since for TRI-NOMFC classifier there is no discordance, the global concordance index is equal to the similarity index; i.e. $C(a_i, b_k^h) = SI(a_i, b_k^h)$. The aggregation operator ψ computes the membership degree $\mu(a_i, C^h)$ of a_i to C^h as shown in Table IV (see page 82). Finally, based on these membership degrees $\mu(a_i, C^h)$, Table V (see page 82) presents the decision rules used by NCD-based MCCs to assign an object a_i to a category C^h.

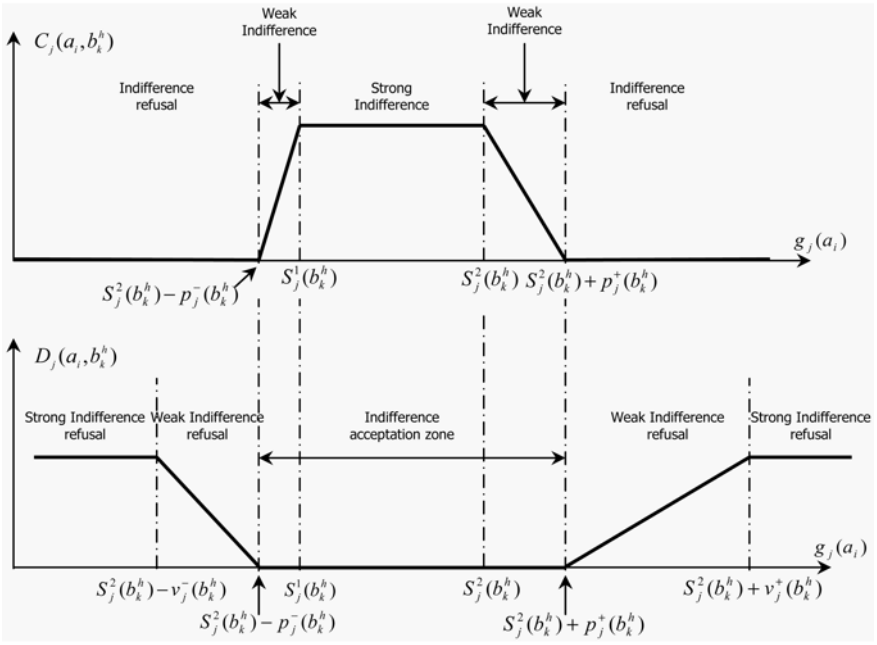


Fig. 4. Representation of local concordance/discordance functions for PROAFTN classifier.

Three interesting elements should be considered from Tables IV and V. The first one is that K-PIP classifier constitutes a generalization of the fuzzy K-NN algorithm [22] where neighbouring is defined by the similarity index SI. Thus, according to this classifier an object a_i has likelihood to be assigned to category C^h if and only if at least one of its profiles b_k^h ($k = 1 \ldots L_h$) belongs to the set of the K more similar profiles to a_i in $B = \bigcup_{h=1}^{H} B^h$. Secondly, TRI-NOMFC classifier [23] has introduced weight coefficients $w_k^h(b_k^h)$ for each profile. Indeed, we believe that the idea of assigning weights to profiles is very interesting since it is possible that, in the same category, a profile $b_{k_1}^h$ is more representative of the category C^h than another profile $b_{k_2}^h$. Finally, the third element is related to the decision rule used in TRI-NOMFC classifier [23] to assign an object a_i to a category C^h. In fact, the advantage of this rule, also known as Hurwitz rule, is to combine an optimistic (*MaxiMax*) and a pessimistic (*MaxiMin*) behaviour in order to provide a more nuanced behaviour in which the optimism level is controlled by the parameter $\alpha \in [0,1]$; When $\alpha = 1$ (respectively $\alpha = 0$) Hurwitz rule is equivalent to the optimistic rule (*MaxiMax*) (respectively to the pessimistic rule (*MaxiMin*)).

3. ELICITATION OF NCD-BASED MCCs PARAMETERS

According to [41], any multiple criteria classification methodology faces two issues: (i) The specification of the aggregation model to be used, and (ii) the assessment of the parameters of the model. In the above section the first issue is discussed. In this section, we propose first a mathematical model that provides the optimal parameters values of the aggregation model. Since, as it will be shown later, the proposed mathematical model can not be solved using classical optimization methods, an Automatic Learning Method (ALM) based on Real-Coded Genetic Algorithm (RCGA) is then proposed to approximate the optimal solution of this model and consequently to infer the parameters values of NCD-based MCCs.

3.1. A Mathematical Model to Provide the Optimal Values of NCD-Based MCC Parameters

It is essential to estimate the best values for aggregation model's parameters (e.g. criteria weights, substitution ratios, indifference, preference and veto thresholds). In MCDA literature, two approaches are proposed to elicit the parameters of MCCs: the Direct Elicitation Approach (DEA) and the Indirect Elicitation Approach (IEA). In the first approach, through an interactive questioning, the decision-maker provides the values of these parameters. The aim of this interaction is to ensure that the provided parameters values represent properly the decision-maker judgments and preference system (value or expertise). However, in many other decision-making situations, the determination of the values of these parameters represents a difficult task due to many reasons such as the size of the problem (i.e. high number of parameters), the imprecise nature of the data, the confusing meaning of the parameters, the analyst ability to perform efficiently the elicitation process, etc. Thus, the DEA is often time-consuming and consequently it may discourage the decision-maker from participating. To overcome the drawbacks of the DEA, the IEA employs ALMs to infer automatically the values of these parameters based on examples (or prototypes) known as training objects (part of a training set). In MCDA literature, this second approach is called Preference De-

TABLE I
Local and Global Concordance Indices

Method	Formula
PIP and K-PIP classifiers [15]	$C(a_i,b_k^h)=\sum_{j=1}^{n} w_j^h \times C_j(a_i,b_k^h)$ or $C(a_i,b_k^h)=\underset{j}{Min}\, C_j(a_i,b_k^h)$, where $C_j(a_i,b_k^h)=\begin{cases}0 & if \quad g_j(a_i)-g_j(b_k^h)\le -p_j^h \\ \frac{1}{2}+\frac{1}{2}\sin\left(\frac{\pi}{p_j^h-q_j^h}\right)\left(g_j(a_i)-g_j(b_k^h)+\frac{p_j^h+q_j^h}{2}\right) & if \quad -p_j^h\le g_j(a_i)-g_j(b_k^h)\le -q_j^h \\ 1 & if \quad \left\|g_j(a_i)-g_j(b_k^h)\right\|\le q_j^h \\ \frac{1}{2}-\frac{1}{2}\sin\left(\frac{\pi}{p_j^h-q_j^h}\right)\left(g_j(a_i)-g_j(b_k^h)+\frac{p_j^h+q_j^h}{2}\right) & if \quad q_j^h\le g_j(a_i)-g_j(b_k^h)\le p_j^h \\ 0 & otherwise\end{cases}$, $\sum_{j=1}^{n} w_j^h=1$ and q_j^h and p_j^h are respectively the indifference and the preference thresholds.
PROAFTN classifier [3]	$C(a_i,b_k^h)=\sum_{j=1}^{n} w_j^h \times C_j(a_i,b_k^h)$, where $C_j(a_i,b_k^h)=Min\left\{C_j^-(a_i,b_k^h), C_j^+(a_i,b_k^h)\right\}$, $C_j^-(a_i,b_k^h)=\frac{p_j^-(b_k^h)-Min\left\{S_j^1(b_k^h)-g_j(a_i),\, p_j^-(b_k^h)\right\}}{p_j^-(b_k^h)-Min\left\{S_j^1(b_k^h)-g_j(a_i),\, 0\right\}}$, $C_j^+(a_i,b_k^h)=\frac{p_j^+(b_k^h)-Min\left\{g_j(a_i)-S_j^2(b_k^h),\, p_j^+(b_k^h)\right\}}{p_j^+(b_k^h)-Min\left\{g_j(a_i)-S_j^2(b_k^h),\, 0\right\}}$, $g_j(b_k^h)$ is defined by the interval $\left[S_j^1(b_k^h), S_j^2(b_k^h)\right]$, $\sum_{j=1}^{n} w_j^h=1$ and $p_j^-(b_k^h)$ and $p_j^+(b_k^h)$ are two preference thresholds.
FBI classifier [31]	$C(a_i,b_k^h)=\sum_{j=1}^{n} w_j^h \times C_j(a_i,b_k^h)$, where $C_j(a_i,b_k^h)=Min\left\{S_j(a_i,b_k^h), S_j(b_k^h,a_i)\right\}$, $S_j(a_i,b_k^h)=\frac{p_j-Min\left\{g_j(b_k^h)-g_j(a),\, p_j\right\}}{p_j-Min\left\{g_j(b_k^h)-g_j(a),\, p_j\right\}}$, $\sum_{j=1}^{n} w_j^h=1$ and q_j and p_j are respectively the indifference and the preference thresholds.
TRI-NOMFC classifier [23]	$C(a_i,b_k^h)=\sum_{j=1}^{n} w_j^h \times C_j(a_i,b_k^h)$, where $C_j(a_i,b_k^h)=F_j(g_j(a)-g_j(b_k^h))$ and F_j is an adaptation of some PROMETHEE criterion functions and $\sum_{j=1}^{n} w_j^h=1$.

segregation Approach (PDA) (e.g. [4], [8], [9], [10] and [36]). In artificial intelligence, the IEA is known as Machine Learning (ML) (e.g. [2] and [28]).

Indeed, the mathematical model that provides the optimal values of NCD-based MCCs parameters is presented as follows:

$$(\Lambda):\begin{cases}\underset{p_l\in P}{\text{Minimize}} & \Gamma=\sum_{a_i\in Z}\sum_{h=1}^{H}(\mu_{ih}(p_1,p_2,\ldots,p_s)-\eta_i^h)^2 \\ \text{Subject to} & \text{Structural Constraints (SC)} \\ & \text{Decision-Maker's Constraints (DMC)}\end{cases},$$

where

$$\eta_i^h=\begin{cases}1 & \text{if} \quad a_i\in C^h \\ 0 & \text{otherwise}\end{cases}, \qquad \mu_{ih}\equiv\mu(a_i,C^h), \quad P=\{p_i\}_{i=1\ldots s}$$

is the parameters set and Z is the training set (i.e. a set of objects which assignments is known in advance).[3] Γ is the objective function of cumulative classification errors and should be minimised, i.e. the difference between the *estimated membership degree* of a_i (i.e. $\mu(a_i,C^h)$) obtained by applying a NCD-based MCC and the *true membership degree* η_i^h of a_i given a priori in the training set Z. Two types of constraints are considered in Λ: Structural Constraints (SC) and Decision-Maker's Constrains (DMC). In general, the first type of constrains are articulated in function of the characteristics of the parameters and their mutual relationship (e.g. the sum

[3]The training set is obtained by partitioning the entire dataset in two subsets: the first one, called training subset, is used to elicit the values of the parameters and the second subset, called test subset, is used to evaluate the performance of the MCCs.

TABLE II
Local and Global Discordance Indices

Method	Formula
PIP and K-PIP classifiers [15]	$D(a_i,b_k^h)=1-\prod_{j=1}^{n}(1-D_j(a_i,b_k^h))$ or $D(a_i,b_k^h)=\underset{j}{Max}\,D_j(a_i,b_k^h)$, where $D_j(a_i,b_k^h)=\begin{cases}1 & if \quad g_j(a_i)-g_j(b_k^h)\le -v_{jh}^- \\ \frac{1}{2}-\frac{1}{2}\sin\left(\frac{\pi}{v_{jh}^+-v_{jh}^-}\right)\left(g_j(a_i)-g_j(b_k^h)+\frac{v_{jh}^++v_{jh}^-}{2}\right) & if \quad -v_{jh}^+\le g_j(a_i)-g_j(b_k^h)\le -v_{jh}^- \\ 0 & if \quad \left\|g_j(a_i)-g_j(b_k^h)\right\|\le v_{jh}^- \\ \frac{1}{2}+\frac{1}{2}\sin\left(\frac{\pi}{v_{jh}^+-v_{jh}^-}\right)\left(g_j(a_i)-g_j(b_k^h)+\frac{v_{jh}^++v_{jh}^-}{2}\right) & if \quad v_{jh}^-\le g_j(a_i)-g_j(b_k^h)\le v_{jh}^+ \\ 1 & otherwise\end{cases}$, v_{jh}^- and v_{jh}^+ are respectively two veto thresholds.
PROAFTN classifier [3]	$D(a_i,b_k^h)=1-\prod_{j=1}^{n}\left(1-D_j(a_i,b_k^h)\right)^{w_j^h}$ where $D_j(a_i,b_k^h)=Max\left\{D_j^-(a_i,b_k^h),D_j^+(a_i,b_k^h)\right\}$ $D_j^-(a_i,b_k^h)=\frac{g_j(a_i)-Max\left\{g_j(a_i),S_j^1(b_k^h)-p_j^-(b_k^h)\right\}}{p_j^-(b_k^h)-Max\left\{S_j^1(b_k^h)-g_j(a_i),v_j^-(b_k^h)\right\}}$, $D_j^+(a_i,b_k^h)=\frac{g_j(a_i)-Min\left\{g_j(a_i),S_j^2(b_k^h)+p_j^+(b_k^h)\right\}}{-p_j^+(b_k^h)-Max\left\{-S_j^2(b_k^h)+g_j(a_i),v_j^+(b_k^h)\right\}}$ where $\sum_{j=1}^{n}w_j^h=1$, $p_j^-(b_k^h)$ and $p_j^+(b_k^h)$ are two preference thresholds and $v_j^-(b_k^h)$ and $v_j^+(b_k^h)$ are two veto thresholds.
FBI classifier [31]	$D_I(a_i,b_k^h)=1-\prod_{j=1}^{n}\left(1-D_j^I(a_i,b_k^h)\right)^{\alpha/n}$ where $D_j^I(a_i,b_k^h)=Max\left\{D_j^S(a_i,b_k^h),D_j^S(b_k^h,a_i)\right\}$, $D_j^S(a_i,b_k^h)=Min\left\{1,Max\left\{0,\frac{g_j(b_k^h)-g_j(a)-p_j}{v_j-p_j}\right\}\right\}$, $\alpha\in[1,n]$ is a technical parameter introduced to modify the degree of synergy between criteria and p_j and v_j are respectively the preference and the veto thresholds.
TRI-NOMFC classifier [23]	No discordance

TABLE III
Similarity Index Computation

Method	Formula
PIP and K-PIP classifiers [15]	- $SI(a_i,b_k^h)=C(a_i,b_k^h)\times\left(1-D(a_i,b_k^h)\right)$ - $SI(a_i,b_k^h)=Min\left(C(a_i,b_k^h),1-D(a_i,b_k^h)\right)$
PROAFTN classifier [3]	- $SI(a_i,b_k^h)=C(a_i,b_k^h)\times\left(1-D(a_i,b_k^h)\right)$
FBI classifier [31]	- $SI(a_i,b_k^h)=Min\left(C(a_i,b_k^h),1-D(a_i,b_k^h)\right)$
TRI-NOMFC classifier [23]	- $SI(a_i,b_k^h)=C(a_i,b_k^h)$ since there is no discordance.

of the criteria weights for each category is equal to 1, i.e. $\sum_{j=1}^{n} w_j^h = 1$; the indifference threshold is smaller or equal than the preference threshold for each criterion, i.e. $q_j^h \le p_j^h$ $(j = 1 \ldots n)$). The second group of constrains expresses the preferences of decision-maker with respect to the NCD-based MCC parameters. For instance, the decision-maker may specify, for a particular category, that $\sum_{j\in J} w_j^h \ge \sum_{j\in I} w_j^h$ where $I,J \subset \{1 \ldots n\}$ and $I \cap J = \emptyset$.

It's noteworthy that Belacel [3] and Henriet [15] have proposed similar mathematical models to infer the parameters values of their respective NCD-based MCC. When the values of the different thresholds are known, Henriet [15] has shown that for specific configurations of global concordance, discordance and similarity indices, the mathematical model (Λ) can be easily relaxed into linear program and then solved by classical optimization methods. In addition, Henriet [15] has pro-

TABLE IV
Membership Degree Computation

Method	Formula
PIP and K-PIP classifiers [15]	- **PIP :** $\mu(a_i,C^h)=\begin{cases}\underset{b_k^h\in V_K(a_i)\cap B^h}{Max}\ SI(a_i,b_k^h) & if \quad V_K(a_i)\cap B^h\neq\emptyset\\ 0 \quad otherwise\end{cases}$, - **K-PIP:** $\mu(a_i,C^h)=\begin{cases}1-\prod\limits_{b_k^h\in V_K(a_i)\cap B^h}\left(1-SI(a_i,b_k^h)\right) & if \quad V_K(a_i)\cap B^h\neq\emptyset\\ 0 \quad otherwise\end{cases}$, where B^h is the profile set of the category C^h, $V_K(a_i)$ is the set of the K more indifferent (or similar) profiles to a_i in $B=\bigcup_{h=1}^{H}B^h$.
PROAFTN classifier [3]	- $\mu(a_i,C^h)=\underset{k}{Max}\ SI(a_i,b_k^h)$
FBI classifier [31]	- $\mu(a_i,C^h)=\underset{k}{Max}\ SI(a_i,b_k^h)$
TRI-NOMFC classifier [23]	- $\mu(a_i,C^h)=\underset{k}{Min}\ SI(a_i,b_k^h)$ - $\mu(a_i,C^h)=\underset{k}{Max}\ SI(a_i,b_k^h)$ - $\mu(a_i,C^h)=\sum_{k=1}^{L_h}w_k^h(b_k^h)\times SI(a_i,b_k^h)$ where $w_k^h(b_k^h)$ is the likelihood degree of object b_k^h and $\sum_{k=1}^{L_h}w_k^h(b_k^h)=1$.

TABLE V
Decision Rules to Assign a Given Object

Method	Decision rule
PIP and K-PIP classifiers [15]	- $a_i\in C^*\Leftrightarrow\mu(a_i,C^*)=\underset{h}{Max}\,\mu(a_i,C^h)$
PROAFTN classifier [3]	- $a_i\in C^*\Leftrightarrow\mu(a_i,C^*)=\underset{h}{Max}\,\mu(a_i,C^h)$
FBI classifier [31]	- $a_i\in C^*\Leftrightarrow\mu(a_i,C^*)=\underset{h}{Max}\,\mu(a_i,C^h)$
TRI-NOMFC classifier [23]	- $a_i\in C^*\Leftrightarrow\mu(a_i,C^*)=\underset{h}{Max}\left\{\begin{array}{l}\alpha\times\left(\underset{k\in\{1..L_h\}}{Max}\left\{SI(a_i,b_k^h)\right\}\right)\\ +(1-\alpha)\times\underset{k\in\{1..L_h\}}{Min}\left\{SI(a_i,b_k^h)\right\}\end{array}\right\}$, where $\alpha\in[0,1]$ is a coefficient of relative optimism. - $a_i\in C^*\Leftrightarrow\mu(a_i,C^*)=\underset{h}{Max}\,\mu(a_i,C^h)$

posed two methods based respectively on K-Means and Genetic algorithms to identify the profiles which characterize each category. In this same perspective, Belacel *et al.* [4] have presented a methodology to infer the parameters of PROAFTN classifier. In their methodology, these authors have made some simplifications: only concordance concept is considered, the criteria weights are assumed to be equal and each category is characterized by only one profile. Hence, the parameters that are inferred in their learning process are: the upper and the lower bounds of the interval $[S_j^1(b_k^h),S_j^2(b_k^h)]$ and the two preference thresholds $p_j^-(b_k^h)$ and $p_j^+(b_k^h)$. Belacel *et al.* [4] solved a mathematical model similar to (Λ) by using a training set and the Reduced Variable Neighbour-

hood Search (RVNS) meta-heuristic recently proposed by [29]. Belacel *et al.* [4] have reported that the Average Identification Rate (AIR)[4] of PROAFTN is, in general, better than the AIR of other classification methods reported on the same datasets.

Since the objective function Γ of (Λ) is neither convex nor concave and may have many local optima, it will be difficult to find a global optimum for (Λ). Therefore, it's not possible to use classical optimization methods (e.g. gradient algorithms and interior-point algorithms) to solve (Λ). To overcome this difficulty, an ALM based on RCGA is proposed to approximate the optimal solution of (Λ) and consequently to infer the parameters values of NCD-based MCCs.

3.2. An ALM Based on RCGA to Infer the Values of NCD-Based MCC Parameters

Genetic Algorithms (GAs) are stochastic algorithms based on the mechanism of the genetic evolution (selection, cross-over and mutation) to solve complex and large optimization problems. GAs were initially introduced by John Holland [19], but they were popularized thanks to the book of Goldberg [12]. The main idea of GAs is to start with an initial population of potential solutions (or chromosomes) arbitrarily selected. Then, evaluate the relative performance of each solution through a fitness function. Then, on the basis of solutions performances, generate a new population using three evolutionary operators: selection, crossover and mutation. The selection operator identifies both the relatively "good" solutions that will be used to generate the new population and the relatively "bad" solutions that will be removed from the current population. The crossover operator swaps the structures of two "parent solutions" in order to form two similar "offspring solutions" that will be involved in the new population. The mutation operator alters arbitrarily the features of one or more solutions in order to increase the structural variability of the population. The above three operators are repeated until a stopping condition is met. A simplified structure of genetic algorithm is shown in Fig. 5.

The application fields of GAs are considerable. For instance, these algorithms are used in:

—Optimization, when the functions to optimize are complex, irregular and with high dimensionality;
—Physics, as optimization methods for real problems (e.g. structures optimization);
—Artificial intelligence, where the adaptive abilities of GAs are exploited;
—Economy, to model the behaviour of agents for instance;
—Image recognition, for example to classify the unknown objects to pre-defined categories;

Begin
- $t = 0$
- Initialize *Population* (t)
- Evaluate *Population* (t)

While (not stopping condition) do
- $t = t + 1$
- Select *Population* (t) from *Population* ($t-1$)
- Recombine (Crossover and Mutation) *Population* (t)
- Evaluate *Population* (t)

End do
End

Fig. 5. A simplified architecture of a GA.

—Graph and game theories, to solve for example the Traveling Salesman Problem (TSP) or some problems in repetitive and differential games;
—...

The success of GAs is mainly due to their ability to exploit vast unknown search spaces in order to orient subsequent searches into useful subspaces. This feature makes GAs more efficient and effective search technique to explore large, complex, and poorly understood search spaces, where classical search tolls are inappropriate. Since many years, binary coded solutions (or chromosomes) have dominated GAs research. However, Michalewicz [25] has showed that this kind of representation for optimization problems with continuous variables may involves at least three drawbacks:

- It's difficult to use binary coding for optimization problems with high dimension and numerical precision. For instance, with 100 variables belonging to the interval $[-500, 500]$ and a precision of 6 decimal numbers, the size of a binary coded solution is 3000. This generates a search space of about 10^{1000}. For this kind of binary coded problems GAs will have weak performance [27];
- The Hamming[5] distance between two neighborhood real numbers may be large in binary coding. For example, the Hamming distance between 0111 (which is equal to 7) and 1000 (which equal to 8), is equal to 4;
- When the crossover and the mutation operators are applied on binary coded continuous chromosomes they may generate new infeasible solutions.

The above drawbacks of the binary coding have motivated the development of other coding types. Real coding is particularly natural when optimization problems involve real variables. GAs with this type of coding are called real-coded GAs (RCGAs) [18]. In the recent years, RCGAs have been used to solve various continu-

[4]This is a performance measure of a classifier. It's defined by the following ratio: the number of objects that are correctly classified divided by the total number of objects.

[5]The Hamming distance between two binary coded strings is defined as the number of bits which are different in the two strings.

ous optimization problems (e.g. [26], [16], [17], [1]). In RCGAs, each solution (or chromosome) is treated as a vector of real numbers. Since the conventional crossover and mutation operators for binary coding are not applicable for real coding, many other adapted operators are proposed in the literature for real coding [18].

In this work, RCGAs will be used to infer the parameters of the NCD-Based MCCs since all of them are real numbers (e.g. thresholds, weights, etc.). To implement the RCGA, some technical choices have been made on its parameters (e.g. selection methods, crossover and mutation operators, etc). These choices will be specified in Section 5. In next section, we briefly describe the military dataset that will be used to test the performance of the different classifiers.

4. MILITARY DATASET DESCRIPTION

The military dataset that will be used in this work includes 2545 Forward Looking Infra-Red (FLIR) images belonging to eight different classes of ships. These images were provided by the U.S. Naval Weapons Center and Ford Aerospace Corporation. Typical silhouettes of the best image of each class and other related information about classes are listed in Table VI (see page 85).

Based on these 2545 FLIR images, Park and Sklansky [30] proposed to extract 11 features[6] (or attributes/criteria). These attributes are obtained as follows:

- The first seven (7) attributes are represented by Hu's [20] moments m_i. These moments are invariant under scaling (different zoom factors), rotation (different look angles) and translation (silhouette not necessarily centered). The moments m_i are computed by using the second and the third order moment formula, let:

$$\mu_{nm} = \sum_{(x,y)\in S} (x-\bar{x})^n (y-\bar{y})^m,$$

  where $(n+m)$ is the moment order; x (respectively y) is the horizontal (respectively the vertical) coordinates in the silhouette S; $\bar{x}$ and $\bar{y}$ are the coordinates of the centroid of S. For instance, the first four (among the seven attributes) Hu's [20] invariant moments m_i $(i = 1\ldots4)$ are given as follows:

—$g_1 = m_1 = r/B$, where $r = \sqrt{\mu_{20}+\mu_{02}}$ the radius of gyration and B is the distance between the camera and the ship.

—$g_2 = m_2 = \dfrac{(\mu_{20}-\mu_{02})^2 + 4\mu_{11}^2}{r^4}$

—$g_3 = m_3 = \dfrac{(\mu_{30}-3\mu_{12})^2 + (3\mu_{21}-\mu_{30})^2}{r^6}$

—$g_4 = m_4 = \dfrac{(\mu_{30}+\mu_{12})^2 + (\mu_{21}+\mu_{30})^2}{r^6}$

  The seven Hu's [20] invariant moments are noted by $g_i = m_i$ for $i = 1\ldots7$. It's worth noting that the weakness of invariant features is that they contain only information that deals with the general shape of the ship and thereby they represent poorly the other details of the observed object. To overcome this disadvantage, Hu [20] proposed four other attributes which provide more information details about ship;
- The last four (4) attributes represent the parameters of an Auto Regressive (AR) Model. They were extracted by fitting an AR model to one-dimensional sequence which represents the projection of a ship image onto horizontal axis. Let $r(i)$, $i = 1\ldots N$, denote the sequence of the projected ship image sampled at N equally spaced points. Based on these sequences, an AR model is defined recursively by:

$$r(i) = \sum_{j=1}^{m} \theta_j r(i-j) + \alpha + \sqrt{\beta}\varepsilon(i).$$

  The above model expresses the projection r_i $(i = 1\ldots N)$ as a linear combination of the previous projections $r(i-j)$ $(j = 1\ldots m)$,[7] plus a bias α and the error $\varepsilon(i)$ associated with the model. The parameters are estimated by a least square fit of the model to the one-dimensional sequence $r(1), r(2), \ldots, r(N)$. Thus, if $\hat{\theta}$, $\hat{\alpha}$ and $\hat{\beta}$ denote the least squares estimates of θ, α and β respectively, the four (4) AR parameters ($m = 3$, $N = 30$) are presented as follows:

—$g_{i+7} = \hat{\theta}_i,\ i = 1\ldots3$

—$g_{11} = \dfrac{\hat{\alpha}}{\sqrt{\hat{\beta}}}$

Park and Sklansky [30] have shown that all AR parameters are invariant to rotation, translation and scaling, so that they may be used as features for classification purpose.

5. COMPUTATIONAL RESULTS

Only two NCD-based MCCs are implemented in this work: PROAFTN classifier [3] and K-PIP classifier [15]. This choice is justified by the two following facts. The first one is that K-PIP classifier is an enhanced version of the FBI classifier [31] (see [5]). Second, since we want to experiment the effects of both concordance and discordance concepts on classification results, TRINOMFC classifier [23] is removed from our list.

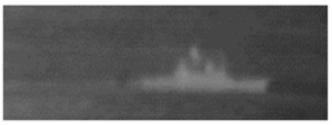
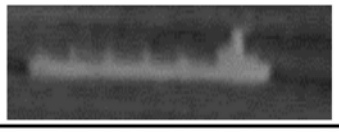
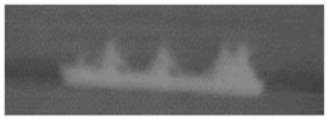
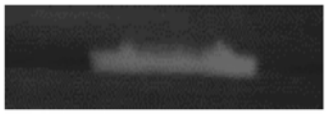
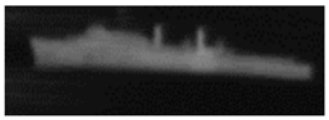
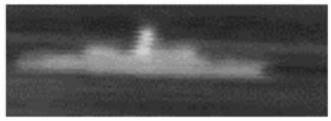
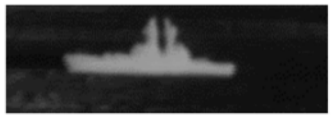
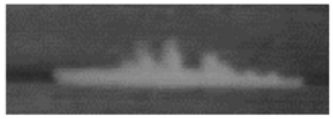
In the ALM proposed in this paper, only thresholds will be estimated for both PROAFTN and K-PIP classifiers:

—For PROAFTN classifier [3], we consider the upper and the lower bounds of the interval $[S_j^1(b_k^h), S_j^2(b_k^h)]$, the two preference thresholds $p_j^-(b_k^h)$ and $p_j^+(b_k^h)$ and

[6] A feature is an abstraction of the raw data in order to represent the original information.

[7] Here m denotes the number of weight parameters.

TABLE VI
Military Dataset Description

Class	Class of ship	Number of images	Typical silhouette
1	Destroyer (D)	340	
2	Container (CO)	455	
3	Civilian Freighter (CF)	186	
4	Auxiliary Oil Replenishment (AOR)	490	
5	Landing Assault Tanker (LAT)	348	
6	Frigate (F)	279	
7	Cruiser (CR)	239	
8	Destroyer with Guided Missile (DGM)	208	

the two veto thresholds $v_j^-(b_k^h)$ and $v_j^+(b_k^h)$ for $j = 1 \ldots 11$, $h = 1 \ldots 8$ and $k = 1 \ldots L_h$;

—For K-PIP classifier[8] [15], we infer the indifference threshold q_j^h, the preference threshold p_j^h and finally the two veto thresholds v_{jh}^- and v_{jh}^+ for $j = 1 \ldots 11$ and $h = 1 \ldots 8$.

It's obvious that the dimensionality (or the number of parameters to infer) of the ALM for the above two classifiers is not the same. For instance, in K-PIP classifier there is only 352 parameters $(4 \times 11 \times 8)$ to estimate while in PROAFTN classifier there is 5280 parameters $(6 \times 11 \times 8 \times 10)$ to estimate if we assume that each category is represented by only 10 profiles.

Since the criteria weights are not included in the ALM, they are estimated by the Entropy method [43]. Hence, the more the criterion discriminates between images the more it will be important. In the other hand, the profiles of each category are identified by using an improved version of K-Means algorithm.[9] The number of profiles in each category is determined by a percentage of the total number of objects in this category. This percentage varies from 1% to 10%. Many others technical choices have been made to implement the RCGA on which is based the

[8]In this computational experiment, the number K of K-PIP classifier (see Table IV) is fixed to 5.

[9]In this improved version of K-means algorithm, we remove each profile that forms an empty group.

ALM:

—Four selection methods are implemented: (1) Roulette Wheel Selection (RWS), (2) Stochastic Remainder Without Replacement Selection (SRWRS), (3) Linear Rank based Selection (LRS) and (4) Tournament Selection (TS)[10];
—Five crossover operators are implemented: (1) Flat Crossover (FC), (2) Arithmetical Crossover (AC), (3) BLX-α Crossover (BLXC), (4) Extended Line Crossover (ELC) and (5) Simple Crossover (SC)[10];
—Four mutation operators are implemented: (1) Random Uniform Mutation (RUM), (2) Non Uniform Mutation (NUM), (3) Mühlenbein Mutation (MM) and (4) Gaussian Mutation (GM)[10];
—The crossover and mutation probabilities vary respectively from 0.6 to 0.8 and from 0.05 to 0.1;
—The size of the generated populations varies from 30 to 80. Note that the initial population (or the set of initial chromosomes) in generated at random. However, the random values of parameters that constitute each chromosome are generated within specific intervals. These intervals are determined based on some statistical measures on the training dataset for each class and each feature (some examples of these statistical measures are presented in Fig. 6). The aim of these measures is to limit the variation domains of the parameters and thereby to make easy the random generation of the initial population;
—The maximum iteration number—fixed to 100—is used as stopping criteria for the RCGA. In fact, this number is not fixed at random. Indeed, by testing many data splits, we have observed that—in general—beyond 100 iterations the improvement of the classification accuracy of the tested classifiers becomes insignificant regarding the computational effort (in time) provided to execute an additional iteration;
—The method used for the evaluation of the classification accuracy is a **cross-validation** method called **repeated random sub-sampling validation**. This technique randomly splits the initial dataset into training and validation (or test) subsets. For each such split, the classifier is retrained with the training subset and validated (or tested) on the remaining subset. The results from each split are then averaged. Hence, according to this cross-validation technique the military dataset is randomly divided into two subsets: a training subset (which size varies from 50% to 70% of the entire dataset) used to infer the values of NCD-based MCC parameters and a test (or a validation) subset (which size varies from 50% to 30% of the entire dataset) used to evaluate the performance of the different MCCs.[11] Hence, for each MCC, 20 different random splits are generated to test its performance.

All algorithms in the ALM—i.e. the RCGA, PROAFTN and K-PIP classifiers—are coded in Visual Basic (VB) and tested on a Pentium IV processor with 2.8 GHz and 512 Mb of RAM. The developed software involves some visualization and statistical tools on the entire, training and test datasets. For instance, Fig. 6 presents some statistical measures on the training dataset for each class and each attribute (or feature).

It's important to underline that two prior works ([37] and [30]) have used the same military dataset to test the performance of four different classifiers: Dempster-Shafer-based (DS) classifier, Modified-Bayes-based (MB) classifier, K-Nearest Neighbors (K-NN) classifier and Neural Net (NN) classifier. The results, expressed in Average Identification Rate (AIR), obtained by these works are presented in Table VII.

Note that the above works, i.e. Valin *et al.* (2006) and Park and Sklansky (1990), use the repeated random sub-sampling cross-validation method to evaluate the classification accuracy of all their tested classifiers. Valin *et al.* (2006) generate many random splits by using Monte-Carlo runs[14] whereas Park and Sklansky (1990) generate only one random split.

The application of PROAFTN and K-PIP classifiers on the military dataset provides the results presented respectively in Tables VIII and IX. An example of screen showing the application of the ALM for K-PIP classifier is presented in Fig. 7. By observing Tables VIII and IX, we conclude that both PROAFTN and K-PIP NCD-based MCCs give, in general, good results: the AIR of FROAFTN is 86.78% and the AIR of K-PIP is 80.69%. Hence, these two MCCs have an AIR better than those of MB and DS classifiers but worse than those of K-NN and NN classifiers.

Some other comments may be made on the classification results of both PROAFTN and K-PIP classifiers:

—The AIRs of both classifiers on the training and test datasets are stable since they are situated around their average (small standard deviation for both classifiers and datasets). We can also use the coefficient of variation ($CV = (\sigma_{\text{AIRs}}/\bar{X}_{\text{AIRs}})\%$)[15] as measure of the robustness (or the stability) of the obtained results. For instance, the *CV* of the PROAFTN AIRs for the test dataset is equal to $CV = 1.65/86.78 = 1.9\%$ (see Table VIII) which is small.
—The AIRs obtained on the training and the test datasets are too different for PROAFTN classifier (average range 7.04%), while these AIRs are nearly the same for K-PIP classifier (average range 2.53%). We believe that this is due both to the high number of

[10]To learn more about these evolutionary operators (crossover and mutation) and methods (selection), we refer the reader to the work of Herrera *et al.* [18].

[11]Note that each subdivision constitutes a partition of the entire dataset, i.e. the union of the training subset and the test subset form the entire dataset.

[14]The number of splits is not specified by these authors.

[15]Note that more the *CV* is small the more the observations are homogeneous, i.e. that the observations are concentrated around the mean. In this case we said that the mean is representative of its observations.

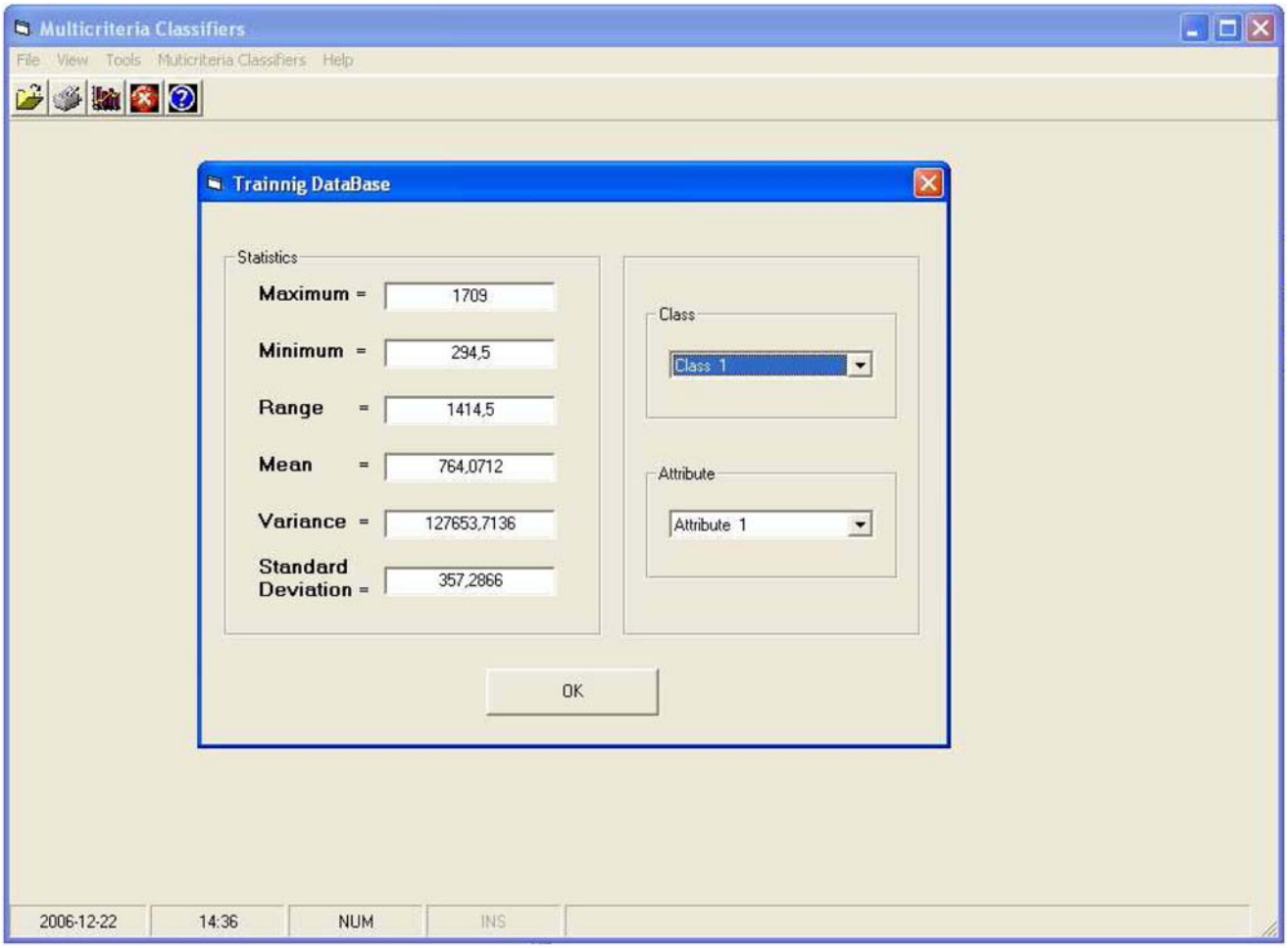


Fig. 6. Descriptive statistical tools.

TABLE VII
The Different Results of Prior Works on the Same Military Dataset

Papers	MB classifier	DS classifier	K-NN classifier	NN classifier
Valin *et al.* [37]	77.7%	74.5%	94.8%[12]	92.7%
Park and Sklansky [30]	***	***	88.3%[13]	***

[12]In this work, the K-NN classifier is applied with K = 3 and by using an Euclidean distance weighted by the inverse of the inter-categories covariance matrix.

[13]In this work, the K-NN classifier is applied with a simple Euclidean distance. The value of K that produces the highest AIR is chosen among the values of K between 1 and 17.

parameters in PROAFTN classifier and to the over-specification (or over-fitting) problem. This problem occurs when the parameters of the classifier became much specific to the data set from which they are assessed. Hence, when these parameters are used to classify another data set, the classification results obtained on this latter will be much different from those obtained on the first data set. Since PROAFTN classifier use more parameters than K-PIP classifier, the over-specification problem will be more apparent with the former classifier;

—According to our computational experiments, all selection methods and evolutionary operators (mutation and crossover) seems to perform equally. However, we have observed that, in general, the AIRs of both MCCs increase when the population size increases;

—Some AIRs obtained by PROAFTN classifier for the training dataset exceed the best AIR obtained by all other classifiers (i.e. 94.8%). This is shows the ability of this classifier to provide better results and thereby we believe that it constitutes a promising classifier which merits to be improved (see Section 6 for eventual improvements);

—The AIR obtained by PROAFTN classifier is better than that of K-PIP classifier since, for a specific criterion, the first consider that the profiles of the same

TABLE VIII
PROAFTN Results

	Genetic Algorithm		Population size	Selection Method	Crossover Step		Mutation Step	
Problems	Training Database	Test database			Cross. Oper	Cross. Prob	Muta. Oper	Muta. Prob
Problem 1	92,62%	86,23%	30	RWS	FC	0,60	RUM	0,05
Problem 2	91,78%	85,03%	30	SRWRS	AC	0,65	NUM	0,07
Problem 3	94,33%	88,54%	30	LRS	SC	0,70	MM	0,09
Problem 4	90,44%	85,09%	40	TS	BLX	0,75	GM	0,10
Problem 5	95,22%	89,67%	40	RWS	ELC	0,80	RUM	0,05
Problem 6	93,02%	84,88%	40	SRWRS	FC	0,60	NUM	0,07
Problem 7	96,17%	88,98%	50	LRS	AC	0,65	MM	0,09
Problem 8	94,01%	86,11%	50	TS	SC	0,70	GM	0,10
Problem 9	93,28%	85,96%	50	RWS	BLX	0,75	RUM	0,05
Problem 10	93,72%	86,31%	60	SRWRS	ELC	0,80	NUM	0,07
Problem 11	95,12%	88,17%	60	LRS	FC	0,60	MM	0,09
Problem 12	91,73%	84,39%	60	TS	AC	0,65	GM	0,10
Problem 13	95,09%	87,81%	70	RWS	SC	0,70	RUM	0,05
Problem 14	96,00%	89,42%	70	SRWRS	BLX	0,75	NUM	0,07
Problem 15	93,08%	86,12%	70	LRS	ELC	0,80	MM	0,09
Problem 16	94,65%	88,23%	70	TS	FC	0,60	GM	0,10
Problem 17	95,00%	87,93%	80	RWS	AC	0,65	RUM	0,05
Problem 18	93,75%	85,59%	80	SRWRS	SC	0,70	NUM	0,07
Problem 19	92,82%	85,07%	80	LRS	BLX	0,75	MM	0,09
Problem 20	94,68%	86,12%	80	TS	ELC	0,80	GM	0,10
Mean	93,83%	86,78%						
Standad deviation	1,50%	1,65%						
Mean range	7,04%							

category don't have necessarily the same thresholds, while the second assumes that the profiles of the same category have identical thresholds. Hence, for a particular category and criterion, PROAFTN classifier provides more thresholds to each profile (i.e. more degree of freedom) than K-PIP classifier.

6. DISCUSSIONS AND CONCLUSIONS

In this paper, a classification methodology that combines the advantages of multi-criteria decision analysis and automated learning algorithms has been proposed. This classification methodology uses some selected NCD-based MCCs as aggregation models and an IEA to assess the parameters values of these MCCs. To understand the implementation of NCD-based MCCs, we have proposed a generalized framework to explain how these classifiers proceed to assign an object to a given category. The strength of the MCCs could be seen along three dimensions: (i) integration of subjective information like the decision-maker knowledge and preferences, (ii) rigorous manipulation of heterogeneous, conflicting and non commensurable information, and (iii) easy to explain, and therefore are not black boxes. The IEA is implemented using a mathematical model that provides automatically the "optimal" values of the NCD-based MCCs parameters. An ALM based on RCGA has been proposed to approximate its "optimal" solution and consequently to infer the parameters values of these classifiers because the proposed model could not be solved by classical optimization tools (e.g. gradient algorithms and interior-point algorithms). The proposed ALM overcomes some simplifications made in prior works (e.g. [4]): both concordance and discordance concepts are taken into account, the criteria weights are used in the computation of the membership degree of an object to a pre-defined category and finally each category may be characterized by many profiles.

A military dataset of 2545 Forward Looking Infra-Red (FLIR) images representing eight different classes of ships is used to test the performance of two NCD-based MCCs (PROAFTN classifier [3] and K-PIP classifier [15] with respect four other classifiers (Dempster-Shafer-based (DS) classifier, Modified-Bayes-based (MB) classifier, K-Nearest Neighbors (k-NN) classifier and Neural Net (NN) classifier). The computational re-

TABLE IX
K-PIP Results

Problems	Genetic Algorithm		Population size	Selection Method	Crossover Step		Mutation Step	
	Training Database	Test database			Cross. Oper	Cross. Prob	Muta. Oper	Muta. Prob
Problem 1	83,38%	80,63%	30	RWS	FC	0,60	RUM	0,05
Problem 2	81,47%	79,84%	30	SRWRS	AC	0,65	NUM	0,07
Problem 3	82,31%	80,89%	30	LRS	SC	0,70	MM	0,09
Problem 4	82,54%	79,58%	40	TS	BLX	0,75	GM	0,10
Problem 5	84,31%	81,03%	40	RWS	ELC	0,80	RUM	0,05
Problem 6	83,25%	80,22%	40	SRWRS	FC	0,60	NUM	0,07
Problem 7	85,02%	82,41%	50	LRS	AC	0,65	MM	0,09
Problem 8	84,59%	81,19%	50	TS	SC	0,70	GM	0,10
Problem 9	83,69%	80,07%	50	RWS	BLX	0,75	RUM	0,05
Problem 10	82,58%	80,97%	60	SRWRS	ELC	0,80	NUM	0,07
Problem 11	85,17%	83,72%	60	LRS	FC	0,60	MM	0,09
Problem 12	80,77%	78,61%	60	TS	AC	0,65	GM	0,10
Problem 13	85,04%	82,51%	70	RWS	SC	0,70	RUM	0,05
Problem 14	82,54%	79,91%	70	SRWRS	BLX	0,75	NUM	0,07
Problem 15	84,00%	81,25%	70	LRS	ELC	0,80	MM	0,09
Problem 16	80,87%	77,88%	70	TS	FC	0,60	GM	0,10
Problem 17	83,36%	80,11%	80	RWS	AC	0,65	RUM	0,05
Problem 18	82,43%	80,81%	80	SRWRS	SC	0,70	NUM	0,07
Problem 19	85,65%	83,09%	80	LRS	BLX	0,75	MM	0,09
Problem 20	81,47%	79,17%	80	TS	ELC	0,80	GM	0,10
Mean	83,22%	80,69%						
Standad deviation	1,46%	1,45%						
Mean range	2,53%							

sults show that NCD-based MCCs provide AIRs better than those provided by MB and DS classifiers but worse than those obtained by K-NN and NN classifiers. Although NCD-based MCCs don't provide the best AIRs in this application, we believe that they are promising classifiers and merit to be further explored. Note that NCD-based MCCs are not optimized for this kind of dataset. In fact, if qualitative information and human judgment are introduced, we are confident that NCD-based MCCs will certainly outrank K-NN and NN classifiers. Moreover, NCD-based MCCs are not black boxes and all their results are automatically explained.

Many improvements could be made to enhance the AIRs of NCD-based MCCs like:

—Integrating the profiles and the criteria weights in the ALM;
—Using other improved versions of K-Means algorithm for the profile identification (e.g. Y-Means [13] or J-Means [14]);
—Integrating the profile weights in the membership degree computation;
—Combining the aggregation operators of different NCD-based MCCs;
—Using the concept of specified classifier, i.e. for the classification purpose we only use a subset of criteria that discriminate more between objects;
—Implementing a parallel version of NCD-based MCCs to reduce the computation time. For instance concordance and discordance indices may be computed simultaneously;
—Since Genetic Algorithms (GAs) are inefficient to exploit local information of solutions in each population, it will be benefit to integrate, in each iteration of the GA, a local search strategy (e.g. steepest descent strategy) to fine-tuning these solutions locally.

We uphold that cross fertilization of multi-criteria decision analysis and information fusion concepts could be beneficial to both domains. Multiple criteria classifiers are designed to consider human in the loop and to support human decision making.

REFERENCES

[1] M. M. Ali and A. Törn
Population set based global optimization algorithms: Some modifications and numerical studies.
Computers and Operations Research, 31, (2004), 1703–1725.

[2] N. P. Arhcer and S. Wang
Application of the back propagation neural networks algorithm with monotonicity constraints for two group classification problems.
Decision Sciences, 24 (1993), 60–75.

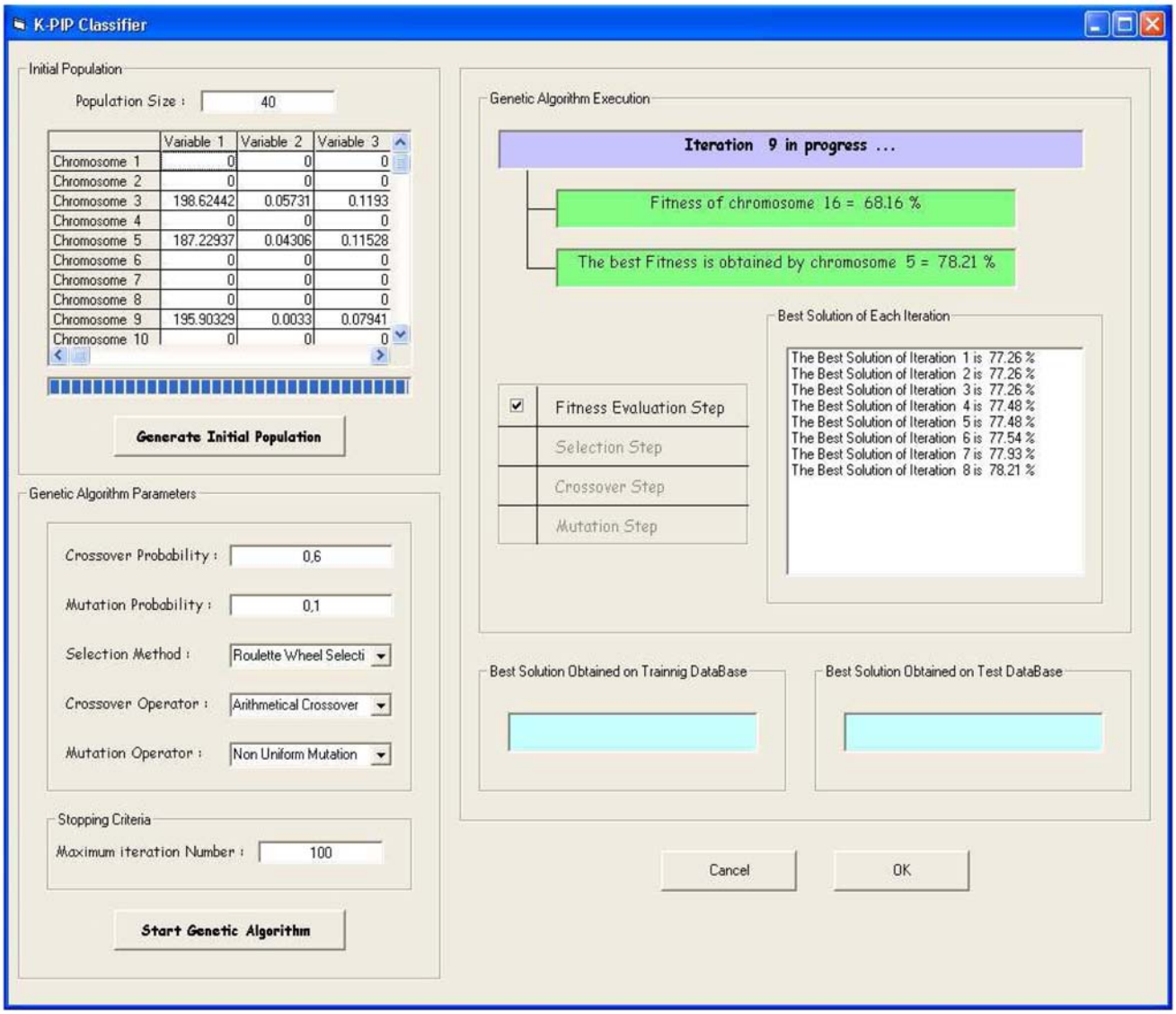


Fig. 7. Screen of the ALM for K-PIP classifier.

[3] N. Belacel
Multicriteria assignment method PROAFTN: Methodology and medical application.
European Journal of Operational Research, 125 (2000), 175–183.

[4] N. Belacel, H. B. Raval and A. P. Punnen
Learning multi-criteria fuzzy classification method PROAFTN from data.
Computer and Operations Research, 34 (2007), 1885–1898.

[5] L. Benabbou, A. Guitouni and P. Lang
Méthode de classification: Revue de littérature, essai de caractérisation et de comparaison.
Document de travail 2004-023, (2004), Faculté des Sciences de l'Administration (FSA), Université Laval, Québec, Canada.

[6] J-P. Brans and Ph. Vincke
A preference ranking organisation method: The PROMETHEE method for multiple criteria decision-making.
Management Science, 31 (1984), 647–656.

[7] D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukias and Ph. Vincke
Evaluation and Decision Models: A Critical Perspective.
Boston/London/Dordrecht: Kluwer Academic Publishers, 2000.

[8] L. C. Dias, V. Mousseau, J. Figueira and J. Climaco
An aggregation/disaggregation approach to obtain robust conclusions with ELECTRE TRI.
European Journal of Operational Research, 138 (2002), 332–348.

[9] L. C. Dias and V. Mousseau
Inferring Electre's veto-related parameters from outranking examples.
European Journal of Operational Research, 170 (2006), 172–191.

[10] M. Doumpos and C. Zopounidis
The use of the preference disaggregation analysis in the assessment of financial risks.
Fuzzy Economic Review, 3 (1998), 39–57.

[11] J. Figueira, Y. De Smet and J-P. Brans
MCDA Methods for Sorting and Clustering Problems: PROMETHEE TRI and PROMETHEE CLUSTER.
20th European Conference on "Operational Research and the Management of Electronic Services," Rhodes Island, Greece, July 4–7, 2004.

[12] D. Goldberg
Genetic Algorithm in Search.
Optimization and Machine Learning, Addison-Wesley, 1989.

[13] Y. Guan, A. Ghorbani, N. Belacel and Y. Means
A Clustering Method for Intrusion Detection.
The Canadian Conference on Electrical and Computer Engineering, Montréal, Québec, Canada, May 4–7, 2003.

[14] P. Hansen and N. Mladenovic
J-MEANS: A new local search heuristic for minimum sum of squares clustering.
Pattern Recognition Journal, 34 (2001), 405–413.

[15] L. Henriet
Systèmes D'évaluation et de Classification Multicritère pour l'aide à la Décision.
Thèses de Doctorat, Université Paris Dauphine, France, 2000.

[16] F. Herrera and M. Lozano
Two-loop real-coded genetic algorithms with adaptive control of mutation step sizes.
Applied Intelligence, 13 (2000), 187–204.

[17] F. Herrera, M. Lozano and A. M. Sanchez
A Taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study.
International Journal of Intelligent Systems, 18 (2003), 309–338.

[18] F. Herrera, M. Lozano and J. L. Verdegay
Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis.
Artificial Intelligence Review, 12 (1998), 265–319.

[19] J. H. Holland
Adaptation In Natural And Artificial Systems.
University of Michigan Press, 1975.

[20] M. K. Hu
Visual pattern recognition by moment invariant.
IEEE Transactions on Information Theory, 8 (1962), 179–187.

[21] E. Jacquet-Lagrèze
An application of the UTA discriminant model for the evaluation of R&D projects.
In P. M. Pardalos, Y. Siskos, C. Zopounidis (Ed.), Advances in Multicriteria Analysis, Dordrecht: Kluwer Academic Publishers, 1995, 203–211.

[22] J. M. Keller, M. R. Gray and J. A. Givens
A fuzzy K-nearest neighbour algorithm.
IEEE Transactions on Systems, Man and Cybernetics, 15 (1985), 580–585.

[23] J. Léger and J-M. Martel
A multicriteria assignment procedure for a nominal sorting problematic.
European Journal of Operational Research, 138 (2002), 349–364.

[24] M. Massaglia and A. Ostanello
N-TOMIC: A decision support for multicriteria segmentation problems.
In P. Korhonen, (Ed.), International Workshop on Multicriteria Decision Support, Lecture Notes in Economics and Mathematics Systems, Berlin: Springer, 356 (1991), 167–174.

[25] Z. Michalewicz
Genetic Algorithms+Data Structures=Evolution Programs.
Berlin: Springer-Verlag, 1992.

[26] Z. Michalewicz
Genetic Algorithms+Data Structures=Evolution Programs.
Berlin: Springer-Verlag, 1996.

[27] Z. Michalewicz, T. Logan and S. Swaminathan
Evolutionary operators for continuous convex parameters spaces.
In A. Sebald, L. Fogel (Ed.), Proceedings of the 3rd Annual Conference on Evolutionary Programming, World Scientific Publishing, 1994, 84–107.

[28] R. S. Michalski
A theory and methodology of inductive learning.
Artificial Intelligence, 20 (1983), 11–116.

[29] N. Mladenovic and P. Hansen
Variable neighbourhood search: Principles and applications.
European Journal of Operational Research, 130 (2001), 449–467.

[30] Y. Park and J. Sklansky
Automated design of linear tree classifiers.
Pattern Recognition, 23 (1990), 1393–1412.

[31] P. Perny
Multicriteria filtering methods based on concordance and non-discordance principles.
Annals of Operations Research, 80 (1998), 137–165.

[32] B. Roy
Classement et choix en présence de points de vue multiples (la méthode ELECTRE).
R.I.R.O. Recherche Opérationnelle, 2 (1968), 57–75.

[33] B. Roy
Critères multiples et modélisation des préférences: l'apport des relations de surclassement.
Revue d'Économie Politique, 1 (1974).

[34] B. Roy
ELECTRE III: Algorithme de classement basé sur une représentation des préférence en présence de critères multiples.
Cahiers du CERO, 21 (1978), 3–24.

[35] B. Roy and J. Moscarola
Procédure automatique d'examen de dossiers fondée sur une segmentation trichotomique en présence de critères multiples.
RAIRO Recherche Opérationnelle, 11 (1977). 145–173.

[36] Y. Siskos, E. Grigoroudis, C. Zopounidis and O. Saurais
Measuring customer satisfaction using a survey based preference disaggregation model.
Journal of Global Optimization, 12 (1998), 175–195.

[37] P. Valin, F. Rhéaume, C. Tremblay, D. Grenier, A-L. Jousselme and E. Bossé
Comparative implementation of two fusion schemes for multiple complementary FLIR imagery classifiers.
Information Fusion Journal, 7 (2006), 197–206.

[38] P. Vincke
L'aide multicritère à la décision.
Éditions de l'Université de Bruxelles, Belgique, 1989.

[39] C. Zopounidis and M. Doumpos
A multicriteria decision aid methodology for sorting decision problems: The case of financial distress.
Computational Economics, 14 (1999), 197–218.

[40] C. Zopounidis and M. Doumpos
Building additive utilities for multi-group hierarchical discrimination: The M.H.DIS method.
Optimization Methods and Software, 14 (2000), 219–240.

[41] C. Zopounidis and M. Doumpos
Multicriteria classification and sorting methods: A literature review.
European Journal of Operational Research, 138 (2002), 229–246.

[42] W. Yu
Aide Multicritère à la Décision Dans le Cadre de la Problématique du Tri: Concepts, Méthodes et Applications.
Thèse de doctorat, Université de Paris-Dauphine, France, 1992.

[43] M. Zeleny
Multiple Criteria Decision Making.
McGraw-Hill, 1982.

Khaled Jabeur received his B.S. degree in computer engineering from Faculté des Sciences de Tunis in 1993. He received his Ph.D. and M.B.A. in operations and system decision from Université Laval in 2004 and 1997, respectively.

Dr. Jabeur is currently an assistant professor at l'Institut Supérieur de Commerce et de Comptabilité de Bizerte (ISCCB). His research interests include multicriteria decision analysis, operational research and group decision-making. He has published various papers in international journals (e.g. *European Journal of Operational Research*, *Group Decision and Negotiation*, *Information Systems and Operational Research*). In 2004, Dr. Jabeur has obtained the Governor General's Academic Gold Medal—which is the most prestigious award that students in Canadian universities can receive—to recognize his outstanding doctoral research.

A. Guitouni is a senior defence scientist with Defence R&D Canada–Valcartier. He specializes in decision support systems, operational research and artificial intelligence. His work includes planning and optimization, automated classification, decision analysis, multiple criteria decision aids, net-enabled dynamic resource management, and collaborative decision making (planning). He is an adjunct professor with Laval University and University of Victoria. He has designed and taught several graduate and undergraduate courses. He has supervised Ph.D., M.Sc. and MBA students.