Fuzzy Assignment Procedure based on Categories’ Boundaries

George Rigopoulos, John Psarras, and Dimitrios Th. Askounis
School of Electrical & Computer Engineering,
National Technical University of Athens, 9 Iroon Polytechneiou Str, 15773, Athens, Greece

Abstract: We presented a procedure for multicriteria assignment problems. Assignment to a category is based on the comparison between an action and categories’ boundaries. Boundary is represented by the least typical action that can be included in the category, and is considered as the category’s inclusion threshold. Comparison is executed considering performance on evaluation criteria and calculation of a fuzzy inclusion relation, generalizing preference relations model with concordance and non-discordance concepts as used in ELECTRE methods. An illustrating example is also presented to demonstrate method’s application to assignment problems. The proposed procedure can be applied to a variety of domains, such as production and operations, as well as financial and human resources.

Keywords: Multicriteria decision analysis; nominal classification;

INTRODUCTION

Assignment of actions (numbers, people, etc) to categories, often referred as classification, is a common objective in decision problems including financial decisions, medical diagnosis, human resources management, and marketing and production management [23, 2]. Operations research, neural networks, mathematics, and machine learning, contribute with algorithms and methodologies to solve such problems. Multicriteria analysis also provides methodologies to solve sorting and classification problems, as well as choice and ranking ones [21]. Classification can generally be distinguished in:

1. Supervised, which refers to predefined categories and requires decision maker’s contribution and,

2. Unsupervised, which does not require decision maker’s contribution.

Depending on whether categories are ordered or not, supervised approach is referred as ordinal sorting or nominal classification respectively, while in unsupervised approach categories are not predefined and it is referred as clustering.

We can further distinguish ordinal sorting multicriteria approaches in direct judgment methods and case based reasoning ones

1. Direct judgment methods. The majority of direct judgment methods belongs to ELECTRE methods family [18] and utilizes outranking relation for the aggregation of criteria scores. NTOMIC [14] and ELECTRE TRI [22] are typical methods of this approach, with ELECTRE TRI being the most well known one. In ELECTRE TRI categories are defined by their limits and the upper limit of a category is the lower limit of the next one. Classification is based on the comparison result between the action and the category limit, which is executed according to the outranking aggregation procedure. Lately, an extension has been proposed allowing sorting without defining category limits [16]. Within outranking approach sorting methods based on PROMETHEE methodology [3] have also been proposed, such as PROMETHEE TRI [6], PAIRCLASS [5] and PROMSORT method [1].

2. Case based reasoning methodologies utilize either utility functions [9] or rough set method [20] to sort actions. Methodologies using UTA approach [8] are M.H.DIS [24] and UTADIS [25], where categories are defined according to utility limits and classification result is derived by comparison between an action and the utility category limit. Rough set methods are based on rough set theory where classification is executed according to specific rules which are defined by decision maker. Methodologies using this approach are ORCLASS [11], Koksalan-Ulu method [10], TOMASSO [13], and work of Greco et al [7].
Compared to ordinal sorting, multicriteria analysis approaches for nominal classification problems are limited. In these approaches, categories are usually defined by their best representative action, which belongs to the category and acts as category profile. The profile is defined by decision maker and classification is based on the comparison result between this and the action to be classified. Nominal classification approaches include multicriteria filtering method [15], PROAFTN method [2] which is based on fuzzy indifference relation, Scarelli and Narula’s [19] multicriteria assignment method and the work of Malakooti and Yang [12].

A detailed review of multicriteria sorting methodologies as well as applications in financial problems can be found in works of Zopounidis and Doumpos [26, 4]. From the analysis of existing approaches, we notice that multicriteria analysis focuses mostly in ordinal sorting problems, while there is relative lack in approaches for nominal classification problems, despite the presence of such problems in several domains such as financial or human resources.

Within this framework we present a novel fuzzy assignment procedure based on multicriteria analysis. Our approach focuses on the case of non-ordered categories, or nominal classification. Assignment of an action to a category is based on the comparison between action and category’s boundary, which is defined by its least typical representative and is considered as category’s inclusion threshold. Evaluation is based on the calculation of a fuzzy inclusion relation for each action, based on generalization of preference relations model and concordance/non-discordance concepts as used in ELECTRE methods [18]. The aim of this paper is to present the methodology and an illustrative example demonstrating its usage. The proposed methodology can be utilized to support assignment problems in a variety of fields, such as production and operations, as well as financial and human resources.

Starting from the introduction, where we present an overview on multicriteria classification approaches, we continue with presentation of the procedure, and an illustrative example demonstrating its applicability. Results and discussion on the methodology are also presented concluding with a short overview considering contributions and future directions.

MATERIALS AND METHODS

In this section we present the proposed methodology for multicriteria assignment problems. The methodology is based on the concept of inclusion/exclusion of an action \( a \) from a category \( C \), which is determined by evaluating the fuzzy inclusion degree of action for the specific category, following the concordance/non-discordance concepts. Categories are defined by an entrance threshold (or more than one in the general case), which can be considered as the least typical representative action that satisfies the inclusion requirements. The objective of the methodology is to classify actions to categories in a way to consider inclusion/exclusion concept.

The following notations and basic definitions will be used next:
1. \( A = \{a_1, a_2, \ldots, a_n\} \) is a set of \( n \) actions to be assigned to predefined categories,
2. \( C = \{C^1, C^2, \ldots, C^h\} \) is a set of \( h \) categories, defined by their least typical representatives referred as entrance thresholds,
3. \( G = \{g_1, g_2, \ldots, g_m\} \) is a set of \( m \) evaluation criteria,
4. \( W = \{w(g_1), w(g_2), \ldots, w(g_m)\} \) is the set of criteria importance weights,
5. \( B = \{b^1_h, b^2_h, \ldots, b^h_h\} \) is a set of \( b \) entrance thresholds for each category \( C^h \in C \), in the general case a category can be defined by more than one threshold,
6. Each action’s \( a \in A \) score on the evaluation criteria \( g_j \) is defined as \( g_j(a_i), \; i = 1, \ldots, n \; / \; j = 1, \ldots, m \),
7. Each category’s \( C^h \in C \) threshold score on the evaluation criteria \( g_j \) is defined as \( g_j(b^h_k), k = 1, \ldots, b \), where \( b^h_k \) is the \( k_{th} \) threshold of category \( C^h \in C \),
8. An action \( a_i \in A \), which is classified to category \( C^h \in C \), is defined as \( a_i \in C^h \).
Based on the above notations the problem that we want to solve is defined as follows:

Having a set of actions \( A = \{a_1, a_2, \ldots, a_n\} \), a set of non-ordered categories \( C = \{C^1, C^2, \ldots, C^h\} \) which are defined by their entrance thresholds \( B = \{b^1_k, b^2_k, \ldots, b^h_k\} \), and a set of evaluation criteria \( G = \{g_1, g_2, \ldots, g_m\} \), assign actions to categories with respect to their score on the evaluation criteria and the inclusion/exclusion concept.

Definition of Fuzzy inclusion relation

The assignment procedure is based on the concept of inclusion/exclusion, which defines at what degree an action can be included in a category or excluded from it. In order to utilize this concept for classification of actions, we define the fuzzy inclusion relation \( P(a, b) \) as a binary relation between an action and a category threshold, where \( a_i \in A \) and \( b^h_k \in B \) utilizing concordance/non-discordance principle.

According to the concept of inclusion/exclusion and considering category thresholds, an action \( a_i \in A \) is preferred over a threshold \( b^h_k \in B \) (and can be thus included in the category \( C^h \in C \)) iff there is a majority of “reasons” supporting preference of action \( a_i \in A \) over threshold \( b^h_k \in B \) and there is no strong opposition to this.

Following the above, fuzzy inclusion relation \( P(a, b) \) can be expressed formally as

\[
P(a, b) \iff C(a, b) \land \neg D(a, b) \tag{1}
\]

where

- predicate \( C(a, b) \) expresses that there is a majority of “reasons” to support action’s \( a_i \in A \) preference (or inclusion) over threshold \( b^h_k \in B \) and,
- predicate \( D(a, b) \) expresses that there is a strong opposition to action’s \( a_i \in A \) preference (or inclusion) over threshold \( b^h_k \in B \).

Associating the “reasons” to evaluation criteria we get the following expressions instead:

- \( C(a, b) \) expresses that there is a significant coalition of criteria for which action \( a_i \in A \) is preferred over threshold \( b^h_k \in B \) and,
- \( D(a, b) \) expresses that there is a significant opposition to action’s \( a_i \in A \) preference over threshold \( b^h_k \in B \).

Partial inclusion relations

According to the definition of fuzzy inclusion relation \( P(a, b) \) (Eq. 1), an action is classified to a category \( a_i \in C \) if the action is preferred on category’s threshold with respect to scores on every criterion.

In order to evaluate the contribution of each criterion, we define partial inclusion relations for every criterion \( g_j \) as follows:

\[
P_j(a, b) \iff g_j(a_i) \in [g_j(b^h_k), \infty) \tag{2}
\]

where \( g_j(a_i) \) and \( g_j(b^h_k) \) are the scores of action \( a_i \in A \) and threshold \( b^h_k \in B \) on criterion \( g_j \) respectively. This definition states that an action \( a_i \in A \) is preferred over a threshold \( b^h_k \in B \) if \( g_j(a_i) > g_j(b^h_k) \).

In order to overcome imprecision in definition of data, we define two discrimination thresholds \( q(g_j) \) and \( p(g_j) \) for each criterion, which consequently leads to three areas of preference (inclusion). These thresholds act as indifference \( q(g_j) \) and preference \( p(g_j) \) thresholds.

The three areas, which are defined with the introduction of the thresholds, are the following,

1. Area of no preference (no inclusion). For \( g_j(a_i) \leq g_j(b^h_k) + q(g_j) \) there is no preference (or inclusion) of action \( a_i \in A \over threshold b^h_k \in B \).

2. Area of weak preference (weak inclusion). For \( g_j(b^h_k) + q(g_j) \leq g_j(a_i) \leq g_j(b^h_k) + p(g_j) \) there is weak preference (or inclusion) of action \( a_i \in A \over threshold b^h_k \in B \).
3. Area of strong preference (strong inclusion). For $g_j(a_i) \geq g_j(b_k^h) + p(g_j)$ there is strong preference (or inclusion) of action $a_i \in A$ over threshold $b_k^h \in B$.

Partial inclusion index

Having defined the partial inclusion relations (Eq. 2) and the two discrimination thresholds, we define an index $C_j(a_i, b_k^h)$ for every partial relation, in order to measure the degree of validity of the above three preference (inclusion) situations. We name this as partial inclusion index, and for each area we define appropriate values for it as follows:

1. Area of no preference (no inclusion). For $g_j(a_i) \leq g_j(b_k^h) + q(g_j)$ the index gets its minimum value $C_j(a_i, b_k^h) = 0$.

2. Area of weak preference (weak inclusion). For $g_j(b_k^h) + q(g_j) \leq g_j(a_i) \leq g_j(b_k^h) + p(g_j)$ the index gets value between maximum and minimum values $C_j(a_i, b_k^h) \in [0,1]$.

3. Area of strong preference (strong inclusion). For $g_j(a_i) \geq g_j(b_k^h) + p(g_j)$ the index gets its maximum value $C_j(a_i, b_k^h) = 1$.

These statements can be represented with the following function

$$C_j(a_i, b_k^h) = \frac{g_j(a_i) - g_j(b_k^h) - q(g_j)}{p(g_j) - q(g_j)}, \text{ for } g_j(b_k^h) + q(g_j) \leq g_j(a_i) \leq g_j(b_k^h) + p(g_j)$$

and

$$C_j(a_i, b_k^h) = 0, \text{ for } g_j(a_i) \leq g_j(b_k^h) + q(g_j),$$

$$C_j(a_i, b_k^h) = 1, \text{ for } g_j(a_i) \geq g_j(b_k^h) + p(g_j).$$

Comprehensive inclusion index

Fuzzy partial inclusion relations as defined in Eq. 2 and the inclusion indexes respectively as defined in Eq. 3, represent the credibility degree of the following statement:

An action $a_i \in A$ is preferred over a threshold $b_k^h \in B$ (and can be thus included in the category) with respect to criterion $g_j$.

In order to evaluate the contribution of all the evaluation criteria, we define the comprehensive inclusion index for action $a_i \in A$ as

$$C(a_i, b_k^h) = \sum_{j=1}^{m} w_j \cdot C_j(a_i, b_k^h)$$

where $C_j(a_i, b_k^h) \in [0,1]$ is the partial inclusion index of action $a_i \in A$ for criterion $g_j$, and $w_j$ is the importance weight of criterion $g_j$.

Comprehensive inclusion index is the aggregated result of the partial indexes, where the aggregation function is the weighted mean considering criteria importance weights.

Partial discordance index

In some cases a criterion can express negative judgment about classification of action $a_i \in A$ to a class $C^h \in C$. More specifically, a criterion $g_j$ can express a significant opposition to action’s $a_i \in A$ preference (or inclusion) over threshold $b_k^h \in B$. In this case the criterion is discordant with the inclusion relation between action $a_i \in A$ and threshold $b_k^h \in B$. To handle this situation, we define a veto threshold $v(g_j)$ for each criterion as the minimum value which is incompatible with the assertion that An action $a_i \in A$ is preferred over a threshold $b_k^h \in B$ (and can be thus included in the category) for criterion $g_j$.

According to the definition of veto threshold $v(g_j)$, the following three areas of discordance are distinguished

1. No discordance. For $g_j(a_i) \leq g_j(b_k^h) + p(g_j)$ there is no discordance for preference (or
inclusion) of action $a_i \in A$ over threshold $b^h_k \in B$.

2. Weak discordance. For $g_j(b^h_k) + p(g_j) \leq g_j(a_i) \leq g_j(b^h_k) + v(g_j)$ there is weak discordance for preference (or inclusion) of action $a_i \in A$ over threshold $b^h_k \in B$.

3. Strong discordance. For $g_j(a_i) \geq g_j(b^h_k) + v(g_j)$ there is strong discordance for preference (or inclusion) of action $a_i \in A$ over threshold $b^h_k \in B$.

In order to measure the discordance degree we define a discordance index $D_j(a_i,b^h_k)$ for every criterion. For each of the above areas we define appropriate values for it as follows:

1. Area of no discordance. For $g_j(a_i) \leq g_j(b^h_k) + p(g_j)$ the index gets its minimum value $D_j(a_i,b^h_k) = 0$.

2. Area of weak discordance. For $g_j(b^h_k) + p(g_j) \leq g_j(a_i) \leq g_j(b^h_k) + v(g_j)$ the index gets value between maximum and minimum values $D_j(a_i,b^h_k) \in [0,1]$.

3. Area of strong discordance. For $g_j(a_i) \geq g_j(b^h_k) + v(g_j)$ the index gets its maximum value $D_j(a_i,b^h_k) = 1$.

These statements can be represented with the following function

$$D_j(a_i,b^h_k) = \frac{g_j(a_i) - g_j(b^h_k) - p(g_j)}{v(g_j) - p(g_j)} , \text{ for }$$

and

$$D_j(a_i,b^h_k) = 0 , \text{ for } g_j(a_i) \leq g_j(b^h_k) + p(g_j) ,$$

$$D_j(a_i,b^h_k) = 1 , \text{ for } g_j(a_i) \geq g_j(b^h_k) + v(g_j) .$$

Comprehensive fuzzy inclusion relation

Utilizing the concordance/non-discordance principles we define the comprehensive fuzzy inclusion relation aggregating the partial inclusion relations weakened by discordance.

The comprehensive fuzzy inclusion relation is defined as

$$P(a_i,b^h_k) = C(a_i,b^h_k) \prod_{j=1}^{m} \left( \frac{1 - D_j(a_i,b^h_k)}{1 - C(a_i,b^h_k)} \right) \tag{6}$$

where

$$C(a,b^h_k) = \sum_{j=1}^{m} w_j \cdot C_j(a_i,b^h_k) \text{ is the comprehensive inclusion index and } D_j(a_i,b^h_k) \text{ is the discordance index.}$$

Fuzzy inclusion degree

In order to measure the credibility degree of the comprehensive fuzzy inclusion relation (Eq. 6) we define the fuzzy inclusion degree of each action $a_i \in A$ for every category $C^h \in C$, which takes into account the strength of concordance weakened by discordance in the general case of multiple category thresholds as follows

$$\gamma(a_i,C^h) = \max \{ P(a_i,b^h_k), ..., P(a_i,b^h_k) \} . \tag{7}$$

Assignment to classes

As defined, each category $C^h \in C$ is represented by one (or as set) threshold that is considered as the least typical representative of the category and is defined by its score on the evaluation criteria. Utilizing the previously defined relations and indexes, in order to classify an action $a_i \in A$ to a category $C^h \in C$ we execute the following steps:

1. Calculate partial inclusion relations. For each action $a_i \in A$ we calculate the fuzzy partial inclusion relations $C_j(a_i,b^h_k)$ (Eq. 3) over all
thresholds \( b_k^h \in B \) and categories \( C^h \in C \) following concordance/non-discordance concepts.

2. Calculate comprehensive inclusion relations. For each action \( a_i \in A \) we calculate comprehensive inclusion relations \( C(a_i, b_k^h) \) (Eq. 4) over all thresholds \( b_k^h \in B \) and categories \( C^h \in C \).

3. Calculate fuzzy inclusion degree. For each action \( a_i \in A \) we calculate fuzzy inclusion degree for every category \( C^h \in C \) as:

\[
\gamma(a_i, C^h) = \max \{ P(a_i, b_k^h), ..., P(a_i, b_k^h) \}.
\]

Assign to category. Assign action \( a_i \in A \) to category for which the fuzzy inclusion degree is the maximum

\[
a_i \in C^h \iff \gamma(a_i, C^h) = \max \{ \gamma(a_i, C^i) \}.
\]

RESULTS AND DISCUSSION

In the following we present a numeric example of the proposed methodology in order to demonstrate its usage in assignment problems. The problem refers to the assignment of a set of 10 actions in 4 categories. The actions and categories' thresholds are defined by their scores on 13 criteria. Specifically

1. Four categories are defined for assignment C1, C2, C3, C4. Following the methodology, thresholds for the categories are defined setting values for each criterion in the scales defined previously. These values correspond to the least typical action that can be included in the category, and are defined by appropriate values reflecting the problem requirements. For simplicity, we also define low indifference and preference zones, and set veto thresholds in relative high values in order to restrict exclusions only for extreme cases. So, eventually thresholds are defined as depicted in Table 1, where \( b \) and \( v \) are the category entrance threshold and veto threshold respectively.

2. A set of 13 evaluation criteria are defined with scales between 1-100 and importance weights as depicted in Table 1.

3. Finally, a set of 10 actions is defined for assignment, and their scores on the evaluation criteria are depicted in Table 2.

Having defined the parameters, all necessary calculations are executed and classification result is presented in Table 3, along with the fuzzy inclusion degrees for each category.

From the results we derive that classification is in agreement with category definition since

- Alternatives with relative high performance in the majority of profitability criteria (G1, G2, G3, G4) and relative medium in the rest criteria (G5, G6, G7, G8, G9, G10, G11, G12, G13) are assigned to category C1.
- Alternatives with relative high performance in criteria G4, G5, G6, G7 and relative medium in criteria G1, G2, G3, G8, G9, G10, G11, G12, G13 are assigned to category C2.
- Alternatives with relative high performance in criteria G7, G8, G9, G10 and relative medium in criteria G1, G2, G3, G4, G5, G6, G11, G12, G13 are assigned to category C3.
- Alternatives with relative high performance in criteria G10, G11, G12, G13, G1 and relative medium in criteria G1, G2, G3, G4, G5, G6, G7, G8, G9 are assigned to category C4.

A sensitivity and robustness analysis on the results would provide details on parameter significance and indicate considerations that have to be taken at the problem formulation step. However, in real world problems decision maker has to define values for categories’ thresholds understanding the impact of wrong definitions on the final result. In order to overcome this, methodology accepts the definition of multiple thresholds for each category. Moreover, decision maker can define as threshold an existing action, which considers as the least typical representative of the category. However, if the problem has been formulated with regard to methodology’s limitations, it provides decision maker with robust classification results.

CONCLUSION

In this paper we presented a novel multicriteria procedure for nominal classification problems. Our approach is differentiated to the existing ones in the assignment procedure. While the common approach is to profile a category with its best representative action, we introduce a new concept where category is profiled with its least typical representative. Assignment thus of an action to the category is based on the comparison between action and category’s profile, which is considered as the category inclusion threshold. Both action and category threshold are scored on evaluation criteria and the calculation of fuzzy inclusion relation for every action drives the assignment procedure. The
The methodology can be applied with success to nominal classification problems to a greater extent, and not only to financial domain. Future development of the methodology includes the case of minimizing decision makers’ interference by utilizing a preference disaggregation approach, as well as development of appropriate decision support tools.
REFERENCES

17. Roy B. 1968. Classement et choix en présence de points de vue multiples (la méthode Eletre). Revue Francaise d’Informatique et de Recherche Operationnelle, 8; 57-75