ICDSST 2016 on Decision Support Systems Addressing Sustainability & Societal Challenges

A Recommender System based on MultiCriteria Aggregation

FOMBA Soumana[1,2], ZARATE Pascale[2]

KILGOUR D.Marc[3]

1 : Université des Sciences, des Techniques et des Technologies de Bamako (USTTB) - Hamdallaye ACI 2000 - Rue 405 - Porte 359 -BP E423 - Mali
2 : IRIT – Toulouse Capitole University – 2 rue du Doyen Gabriel Marty 31042 Toulouse Cedex 9 – France
3 : Wilfrid Laurier University - 75 University Avenue West, Waterloo, Ontario N2L 3C5 Canada
Soumana.Fomba@irit.fr, Pascale.Zarate@irit.fr
mkilgour@wlu.ca
web-page: http://www.irit.fr/

ABSTRACT

Recommender systems aim to support decision-makers by providing decision advice. We offer multi-criteria decision recommendations based on a performance matrix and a partial order on criteria submitted by the user. Our method is to aggregate performance measures over all criteria based on inferences about preferences from the decision-maker's input. After reviewing some multicriteria aggregation operators, we present a recommender system that uses the Choquet integral of a fuzzy measure to determine a total ordering of the alternatives.
Keywords: Recommender System, Choquet Integral, MCDA

INTRODUCTION

Research in the field of multi-criteria decision aid (MCDA) [1] has provided us with models of decision problems that are both flexible and robust. Indeed, the use of fuzzy measures [2] addresses both of these concerns, and their use in Choquet integrals [3] effectively models the preferences of the decision-maker, while taking into account both positive or negative synergies among criteria. A fuzzy measure that represents numerically these preferences can be determined based on the decision-maker's input, and can thus be used to support decision making.

In this article, we identify a fuzzy measure based on a partial order on a subset of alternatives submitted by the decision-maker. Then through the Choquet integral we establish a final ranking of the alternatives. After this introduction, we introduce the required notation and present several aggregation operators. The third section describes the application of the

Choquet integral to the evaluation of a fuzzy measure. Then, we demonstrate how the Choquet integral is implemented and then, in the final section, show how an intuitive web interface can be developed.

Notation

We start by introducing some concepts. Let $X = \{a, b, \ldots\}$ be the set of alternatives (solutions), and let $N = \{1, \ldots, n\}$ be the set of indices of criteria.

Let $\geq$ be a relation on X representing the decision-maker's preference. ($\geq$ is usually pronounced "at least as good as".) As a binary relation, $\geq$ is usually assumed reflexive. For alternatives a and b, we use both the prefix notation $\geq(a, b)$ and the infix notation $a \geq b$ to mean that a is preferred to b. As the prefix notation indicates, $\geq$ is considered to be a function on X^2. If the relation is binary, then $\geq$ takes the values 0 and 1 only; if the relation is fuzzy (blurred or valued [4]), $\geq$ takes values in [0, 1]. Note that $\sim$ is the symmetric part of the $\geq$ relation, i.e., $a \sim b$ iff $a \geq b$ and $b \geq a$ (and is pronounced "a is indifferent to b").

AGGREGATION OPERATORS

The overall objective of an aggregation operator in a decision support *problematique* is to determine an overall score for an alternative from its local performance and the user's preferences, in order to compare it to other alternatives. The overall score is used to establish a ranking to help the decision-maker in his decision. Next we will review several popular aggregation operators.

Weighted sum

The weighted sum is often used because of its simplicity. It requires a weight for each criterion, based on its degree of importance in the decision problem. It is defined by

$$\psi(\mathrm{a}_1, \ldots, \mathrm{a}_\mathrm{n}) = \sum_{i=1}^{n} w_i a_i \qquad (1.1)$$

where, for $i = 1, 2, \ldots, n$, $w_i \in [0, 1]$ represents the weight of criterion i and

$$\sum_{i=1}^{n} w_i = 1 \qquad (1.2)$$

The weighted sum is a very limited operator because it does not take into account the dependencies between the criteria. Moreover, the preferences of the decision maker are included in a simplistic way, through fixed weights assigned to each criterion.

The ordered weighted sum

The ordered weighted sum [5] is a class of aggregation operators (called OWA) that determines the weight of a criterion for an alternative based on performance on that alternative relative to others. (In contrast, in weighted sum aggregation, the weight of a criterion depends on the nature of the criterion.) OWA is defined by

$$OWA_w(a_1, \ldots, a_n) = \sum_{i=1}^{n} w_i a_{(i)} \qquad (1.3)$$

where $W = (w_1, \ldots, w_n)$ is a weight vector satisfying $w_i \in [0, 1]$ for all i and

$$\sum_{i=1}^{n} w_i = 1 \tag{1.4}$$

where the notation (A) refers to the permutation of the indices of $A = (a_1, ..., a_n)$ that satisfies $a_{(1)} \leq ... \leq a_{(n)}$. It is possible to us OWA to model many basic functions such as:

- The « Max » function, with weight vector $W = (0,0, ..., 1)$
- The « Min » function, with weight vector $W = (1, ..., 0, 0)$
- The « average», or mean, over the n criterions, with weight vector $W = (\frac{1}{n}, ..., \frac{1}{n})$.

The Choquet Integral

Aggregation operators such as weighted sum and OWA are unable to model interactions between (or among) criteria. But, in practice, synergy between criteria is normal, and a realistic representation must take these interactions into account. For this reason, methods based on weight vectors are not appropriate. Instead, a non-additive function is used to define a weight, not only for each criterion, but also for each subset of criteria. These non-additive functions model the importance of criteria as well as the positive and negative synergies between them. Sugeno proposed that these non-additive functions be called fuzzy measures [2].

Definition: *A fuzzy measure μ on N is a function μ: $2^N \rightarrow [0, 1]$ which is monotonic, that is, $\mu(S) \leq \mu(T)$ whenever $S \subseteq T$, and satisfies the limit conditions $\mu(\emptyset) = 0$ and $\mu(N) = 1$.*

The measure of a subset $S \subseteq N$ of criteria, $\mu(S)$, reflects the weight or importance of the criteria in S (compared to all others). To determine a fuzzy measure means to determine 2^n weights, corresponding to the 2^n subsets of N. The measure μ may be additive, that is, $\mu(S \cup T) = \mu(S) + \mu(T)$ for all subsets S and T, in which case, the weights of the n criteria are sufficient to calculate the fuzzy measure.

Fuzzy measures may be used as aggregation operators. By using the weights of criteria and sets of criteria, it is possible to represent interactions between criteria. These functions are called fuzzy integrals [2]. There are several classes of fuzzy integrals, of which one of the most representative is the Choquet integral [3].

The Choquet integral is defined as follows: Let μ be a fuzzy measure on N. The Choquet integral of $x \in \mathbb{R}^n$ with respect to μ is defined by:

$$C_{\mu}(x) := \sum_{i=1}^{n} x_i[\mu(A_i) - \mu(A_{i+1})] \tag{1.5}$$

where (.) denotes the permutation of the components of $x = (x_1, \ldots, x_n)$ such that $x_{(1)} \leq ... \leq x_{(n)}$. As well, $A_{(i)}=\{(i), ..., (n)\}$ and $A_{(n+1)} = \emptyset$.

In the Choquet integral, fuzzy measures represent the dependencies between (among) the criteria, as well as the relative weight of each criterion. The index of importance [6] or Shapley's value for criterion i with respect to μ is defined by:

$$\boldsymbol{\Phi}(\mu, \boldsymbol{i}) := \sum_{\boldsymbol{T} \subseteq \boldsymbol{N} \setminus \boldsymbol{i}} \frac{(n-t-1)!\, t!}{n!} [\mu(\boldsymbol{T} \cup \boldsymbol{i}) - \mu(\boldsymbol{T})] \tag{1.6}$$

If μ is additive, we have $\mu(T \cup i) - \mu(T) = \mu(i)$; otherwise, this equality is false and the criteria are dependent. To assess the degree of interaction between criteria i and j with respect to μ, the index of interaction [7] can be employed:

$$I(\mu, ij) = \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\,t!}{(n-1)!} (\Delta_{ij}\mu)(T) \quad (1.7)$$

$$(\Delta_{ij}\mu)(T) := \mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T). \quad (1.8)$$

The index of interaction $I(\mu,ij)$ is in the range [-1, 1] for all $i, j \in N$. If the index is positive, then there is a synergy between these two criteria. Conversely, if the index of interaction is negative, the criteria are called redundant.

Models that are 2-additive are economical because only interactions between two criteria need be considered. In this case, the Choquet integral is defined for all $a \in \mathbb{R}^n$ by

$$C_\mu(a) = \sum_{I_{ij}>0} (a_i \wedge a_j) I_{ij} + \sum_{I_{ij}<0} (a_i \vee a_j) |I_{ij}| + \sum_{i=1}^{n} a_i (\Phi(i) - \tfrac{1}{2} \sum_{j \neq i} |I_{ij}|) \quad (1.9)$$

DETERMINATION OF A FUZZY MEASURE

The next step is to identify a capacity so that the Choquet integral with respect to this capacity represents the preferences of the decision-maker. In practice, the decision-maker can usually indicate a subset $O \subseteq X$ of alternatives of interest, usually of low cardinality, on which the decision-maker has definite preferences. Alternatively, the decision-maker may be willing to specify partial preferences on the set of all criteria. Therefore the input from which the capacity is to be determined may consist of, for example:

- A partial preorder $\geq_O$ on the subset O;
- A partial preorder $\geq_N$ on the set of criteria, N;

In the present context, it seems natural to translate the partial preorder $\geq_O$ using the following rules:

- $a \geq_O b$ is equivalent to $C_\mu(a) > C_\mu(b)$
- $a \sim_O b$ is equivalent to $C_\mu(a) = C_\mu(b)$

where μ is the capacity to be determined. Similarly, $i \geq_N j$ **c**an be taken to be equivalent to $\Phi(\mu, i) \geq \Phi(\mu, j)$ on the set of criteria N and $i \sim_N j$ to $\Phi(\mu, i) = \Phi(\mu, j)$ on the same set.

Translating all the preferences expressed by the decision-maker using the rules above produces an optimization problem whose solution is the fuzzy measure μ on N. This optimization problem is expressed as follows:

Min or Max $F(\ldots)$

$$\text{Subject to} \begin{cases} \mu(S \cup i) - \mu(S) \geq 0, \forall i \in N, \; \forall S \subseteq N \setminus i, \\ \mu(\emptyset) = 0, \mu(N) = 1, \\ C_{\mu(a)} - C_{\mu(b)} \geq \boldsymbol{\delta} c, \\ \ldots \\ \Phi(\mu, i) - \Phi(\mu, j) \geq \boldsymbol{\delta} sh \\ \ldots \end{cases} \quad (1.10)$$

where F is an objective function that depends on the method of identification chosen. Among the main methods are:

- Approaches based on least squares [8];
- Approaches based on linear programming [9];
- Method of minimum variance [10].

We use the Kappalab package [11] to determine the capacity using the minimum variance method [10]. Other tools such as JRI (Java \ R Interface for using a Java program within R), JDK (Java Development Kit), and the libraries for the operation of a J2EE application, could also be used. The capacity determination process can be summarized in the following steps:

1. Define the set of criteria used for the decision problem, entered by the user.
2. Define the performance of each alternative for each criterion. This is the performance matrix, entered by the user.
3. Establish a partial order on the subset of alternatives specified by the user. Preference is defined for a pair of alternatives by preference value, which must be one of
   - 1, if the first alternative is preferred to the second,
   - −1, if the second alternative is preferred to the first,
   - 0, if both alternatives are indifferent or equivalent.
4. Using the preference table, create an R matrix containing all preferential information.
5. Use the function mini.var.capa.ident of package Kappalab to determine the capacity corresponding to the preferential information. This function, executed on the R platform using components JRI, uses the minimum variance identification method [10] to determine a capacity.
6. Recover the resulting capacity from the R platform.

CHOQUET INTEGRAL IMPLEMENTATION

Now, using the package Kappalab, we can aggregate the performance of each alternative using the Choquet integral to determine an overall score with which to establish a final ranking. We proceed to calculate the full Choquet integral as follows:

1. Calculate the Choquet integral value of each alternative.
2. Sort the alternatives based on the overall score of each one.

Note that Step 1 can be simplified if the capacity determined above is 2-additive, i.e., if the interaction index of all subsets of more than two criteria is 0, which can be determined by first finding the index of interaction of all 2^n subsets of criteria and then the Shapley values of all alternatives. Using 2-additivity, the Choquet integral value for each alternative can be obtained by applying the formula for the Choquet integral using the indices of interaction and Shapley values.

RECOMMENDER SYSTEM: WEB PLATFORM

We will illustrate the application through an example [12] involving four chefs. We want to evaluate the chefs based on their ability to prepare three dishes:

- Frog legs (CG)
- Steak tartare (ST)
- Scallops (SJ).

The evaluation of the 4 cooks a, b, c, d on a scale 0 to 20 is given:

	CG	ST	SJ
a	18	15	19
b	15	18	19
c	15	18	11
d	18	15	11

Reasoning of the decision maker:

- When a chef is known for his preparation of Scallops, it is more important that he prepares frog legs well, as compared to steak tartare;
- Conversely, when a cook does not do a good job preparing scallops, it is more important that he prepares steak tartare well, as compared to frog legs.
- Thus, we can conclude that the decision maker thinks : $a \geq b \geq c \geq d$

Results: one can easily check that the decision maker's preferences were taken into account. The system we implemented obtained the same final ranking, $a \geq b \geq c \geq d$.

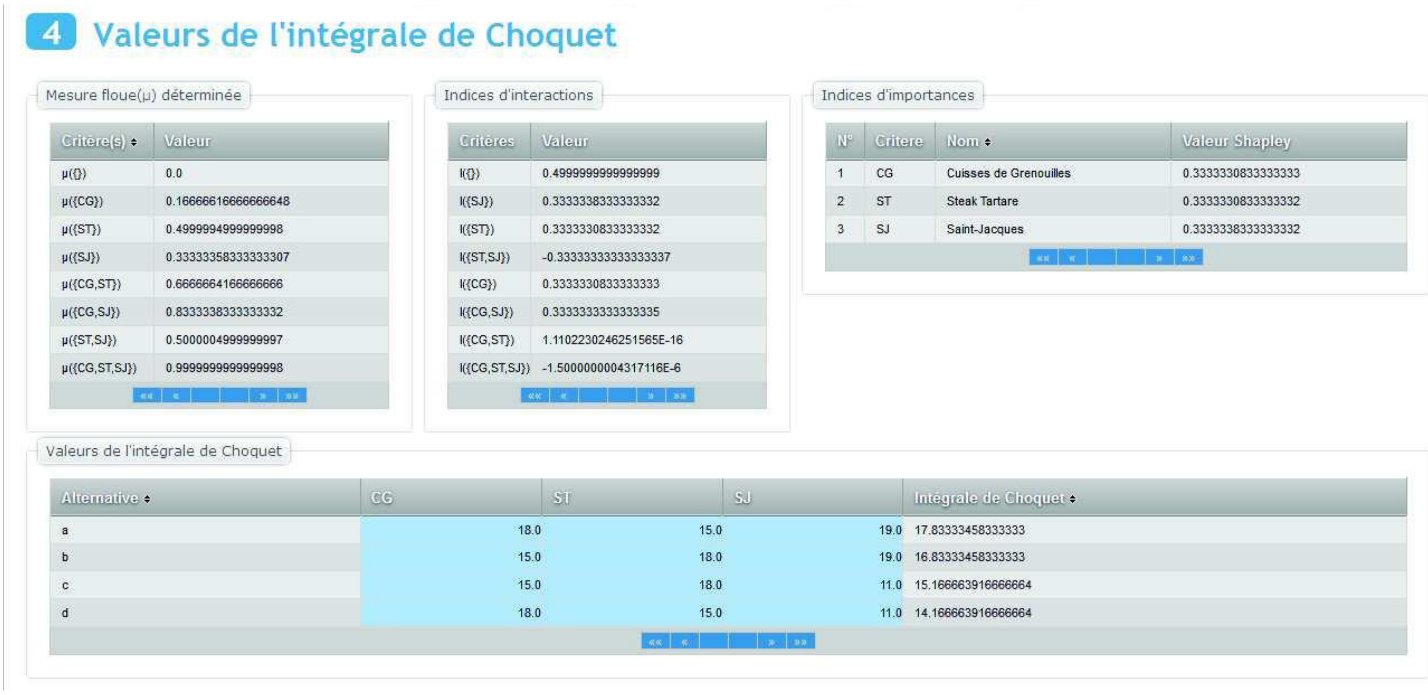

Figure 1: Results

CONCLUSIONS

We gave an overview of some aggregation operators useful for MCDA. Our aim was to show how a fuzzy measure could be used to address a problem of decision aid using as input a partial order established by the decision maker. The implemented web platform is under development, and will be strengthened by incorporating other aggregation operators and other decision aid concepts, such as the bicapacity [12] concepts and the bipolar Choquet integral [12]. Another possible development would be the automatic adjustment of aggregation techniques, based on the context and the profile of the user.

Acknowledgments. The authors would like to thank the CIMI Excellence Laboratory, Toulouse, France, for inviting Marc Kilgour on a Scientific Expert position during the period May-June 2015.

REFERENCES

[1] B. ROY, "Méthodologie multicritère d'aide à la décision", Kluwer Academic Publishers, Dordrecht, 1996.

[2] M. SUGENO, Theory of fuzzy integrals and its applications, Thèse de doctorat, Tokyo Institute of technology, Japon, 1974.

[3] G. CHOQUET, Theory of capacities. In Annales de l'Institut Fourier, 5: 131-295, 1953.

[4] J.C. FODOR and M. ROUBENS, Fuzzy Preference Modelling and Multi-Criteria Decision Aid, Kluwer Academic Publisher, 1994.

[5] R. YAGER, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Trans. Systems, Man & Cybern., 18, pages 183–190, 1988.

[6] L. SHAPLEY, A value for *n*-person games. In *Contributions to the Theory of Games*, volume II, by H.W. Kuhn and A.W. Tucker, editors. *Annals of Mathematical Studies* v. 28, pp. 307–317. Princeton University Press, 1953.

[7] T. MUROFUSHI and S. SONEDA, Techniques for reading fuzzy measures (iii): interaction index, In 9th Fuzzy System Symposium, pages 693–696, Sapporo, Japan, May 1993.

[8] T. Mori, T. Murofush, An analysis of evaluation model using fuzzy measure and the Choquet integral, in '5th Fuzzy System Symposium', Kobe, Japan, pp. 207–212, 1995.

[9] J. Marichal and M. Roubens, Determination of weights of interacting criteria from a reference set, European Journal of Operational Research 124, pp. 641–650, 2000.

[10] I. Kojadinovic, Minimum variance capacity identification, European Journal of Operational Research, in press.

[11] M. Grabisch, I. Kojadinovic, and P. Meyer., Kappalab : an R package for Choquet integral based MAUT, 63rd meeting of the EURO Working Group Multicriteria Aid for Decisions (MCDA), 2006.

[12] A. Rolland, J. Ah-Pine, and B. Mayag, Elicitation of 2-additive bi-capacity parameters, EURO Journal on Decision Processes, DOI 10.1007/s40070-015-0043-3, PP. 6-26, 2015.