About the computation of robust PROMETHEE II rankings: empirical evidence

Y. De Smet[1]

[1]SMG unit, Computer and Decision Engineering department, Ecole polytechnique de Bruxelles, Université libre de Bruxelles, Bruxelles, Belgium

(yves.de.smet@ulb.ac.be)

Abstract – **Engineering decision problems often involve the simultaneous optimization of several conflicting criteria. Among multicriteria decision aid methods, PROMETHEE has gained a lot of attention during the last three decades. Despite its successful application in different fields, some researchers have pointed out the fact that PROMETHEE does not respect the independence to third alternatives property. This leads to the so-called "rank reversal" phenomenon; the relative position of two alternatives may depend on a third one (and so a manipulation threat might exist). In this paper, we propose an alternative method to compute a complete ranking. Inspired by the ideas of robust statistics, we assess the probability of having a given alternative being ranked before another one based on different samplings of the set of alternatives. This allows us to assess a new pairwise comparison matrix that is then exploited by means of net flow scores. We show on an illustrative example that the new method does not suffer from rank reversal (while the PROMETHEE II ranking does).**

Keywords – **Multicriteria Decision Aid, PROMETHEE II Ranking, Rank Reversal.**

I. INTRODUCTION

Strategic engineering problems often rely on the simultaneous optimization of multiple conflicting criteria. For instance;

- In road design problems, one tries to minimize investment and operational costs, maximize security (for all types of users), optimize environmental impacts (greenhouse gases emissions, noise, etc.), improve traffic mobility, etc. [1].
- When facing the development of new 3D Stacked Integrated Circuits (SIC), one has to minimize the cost, the interconnection global length, the package volume, the thermal dissipation, the power consumption, etc. This problem can be modeled as a multi-objective combinatorial optimization problem leading to a huge number of possible design options [2].

Nowadays, many methods are available to address multicriteria problems. Among them, we may cite the multi-attribute utility theory [3], the Analytical Hierarchy Process [4], the ELECTRE methods [5], MACBETH [6], PROMETHEE [7], etc.

In this contribution, we will focus on PROMETHEE. These approaches have been extensively applied in finance, health care, sport, transport, environmental management, etc. [8]. To our point of view, this success is due to (1) the simplicity of the methods and (2) the availability of user-friendly software [9].

Despite their successful applications in various fields, these methods (as others) have been criticized because they suffer from the so-called "rank reversal" phenomenon [10]: the relative ranking of two alternatives may depend on the presence (or the deletion) of a third one. The method detractors emphasize the fact that such distinctive feature may lead to manipulation threats. Following these observations, researchers have investigated conditions under which rank reversal could be avoided or controlled [11; 12]. For instance, in the context of the PROMETHEE methods, they have pointed out that rank reversal could not happen if the scores of alternatives were sufficiently different.

In this contribution, instead of further studying such conditions, we propose an alternative method to compute a robust PROMETHEE II ranking. Based on ideas coming from robust statistics, we propose to compute rankings on various samples of the set of alternatives. These will allow us to assess the "global" probability of having a given alternative being preferred to another one (and vice versa). These pairwise comparison probabilities will be considered as a new preference matrix that will then be exploited by means of net flow scores. The robustness of this new method will be compared to the traditional PROMETHEE II ranking on a simplified version of the Human Development Index (2015).

Let us point out that the question of "rank reversal" is at the origin of huge scientific debates about its legitimacy (or not). In this contribution, we do not want to address this issue. On the contrary, our aim is to show that an alternative method can lead to lower (or eventually eliminate?) the number of rank reversal instances. Finally, let us emphasize that different types of rank reversal can be considered (adding/removing a set of alternatives, adding/removing a copy of a given alternative, removing a non discriminating criterion, etc.). In this paper, we will restrict ourselves to rank reversal occurrences that are induced by the deletion of a single alternative.

The paper is organized as follows. In section 2, we recall the main steps needed to compute the PROMETHEE II ranking. The new ranking method is presented in section 3. Then a comparison on an illustrative example is performed in terms of rank reversal sensitivity in section 4.

II. COMPUTING PROMETHEE II RANKINGS

In this section, we remind the basic steps to compute the PROMETHEE II ranking. The interested reader is referred to [6] for a complete presentation of the PROMETHEE and GAIA methodology.

Let $A=\{a_1,\ldots,a_n\}$ be a set of n alternatives and $F=\{f_1,\ldots,f_q\}$ be a set of q criteria. Without loss of generality we assume that these criteria have to be maximized. At first, we compute the difference between every couple of alternatives according to each criterion:

$$d_k(a_i,a_j) = f_k(a_i) - f_k(a_j) \tag{1}$$

This allows to quantify in which way alternative a_i is better than alternative a_j on a given criterion f_k. Nevertheless, $d_k(a_i,a_j)$ depends on the units of the considered criterion and does not integrate any intra-criterion preference information provided by the decision maker. Therefore the next step is to transform these differences into preference degrees using a non-decreasing preference function $P_k : \Re \rightarrow [0,1]$. An example of such a function is presented in figure 1. The parameter q_k is an indifference threshold. Below this value the difference is so low that the preference is equal to zero. The parameter p_k is a preference threshold. Beyond this value the difference is considered to be so important that a strict preference is stated. Between these two thresholds, the preference is assumed to increase linearly (let us note that other preference functions can be used).

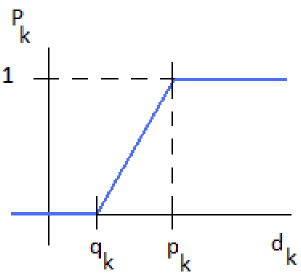

Figure 1. Example of a preference function.

Once the pairwise comparisons between a_i and a_j have been computed for every criterion, they are aggregated into a global preference degree denoted $\pi(a_i,a_j)$:

$$\pi(a_i,a_j) = \sum_{k=1}^{q} \omega_k . P_k[d_k(a_i,a_j)] \tag{2}$$

Where ω_k represents the weight associated to criterion f_k. Furthermore, we assume that $\omega_k > 0$ and $\sum_{k=1}^{q} \omega_k = 1$.

Each alternative can be characterized by its "*strengths*" and "*weaknesses*" with respect to the whole set of alternatives. This can be quantified by computing the positive and negative flow scores:

$$\emptyset^+(a_i) = \frac{1}{n-1}\sum_{a_j \epsilon A} \pi(a_i,a_j) \tag{3}$$

$$\emptyset^-(a_i) = \frac{1}{n-1}\sum_{a_j \epsilon A} \pi(a_j,a_i) \tag{4}$$

The PROMETHEE I partial ranking is obtained as being the intersection of the two rankings induced by $\emptyset^+$ and $\emptyset^-$. Finally, the PROMETHEE II score is computed as follows:

$$\emptyset(a_i) = \emptyset^+(a_i) - \emptyset^-(a_i) \tag{5}$$

This score is lying between [-1;1] and allows to rank the alternatives from the worst to the best ones. We have:

- $\emptyset(a_i) > \emptyset(a_j)$: a_i is preferred to a_j (a_i P a_j);
- $\emptyset(a_i) < \emptyset(a_j)$: a_j is preferred to a_i (a_j P a_i);
- $\emptyset(a_i) = \emptyset(a_j)$: a_i is indifferent to a_j (a_i I a_j)

III. A NEW METHOD TO COMPUTE ROBUST PROMETHEE II RANKINGS

As already stressed, the PROMETHEE II method suffer from the so-called dependence to third alternatives; the relative position of two alternatives may depend on the presence (or not) of a third one. This will be illustrated in section IV.

Based on this observation, researchers have started to study under which conditions rank reversal could occur (or not) [11,12]. For instance, when a given alternative dominates another one (i.e. is at least as good for all criteria and strictly better for a least one criterion), it is easy to prove that rank reversal could never happen between these alternatives. Mareschal et al. [10] have further proved that rank reversal could never happen between a_i and a_j as soon as:

$$|\emptyset(a_i) - \emptyset(a_j)| > \frac{2}{n-1} \tag{5}$$

From a theoretical point of view, this result shows that rank reversal can only happen between alternatives that have "close" net flow scores. Unfortunately, from a practical point of view, this bound is often too rough to be applied in practice (see section IV).

The underlying idea of the new model is to propose a way to better assess the probability of having an alternative a_i being better ranked that a_j. This will be denoted p_{ij}. For doing that, we propose to randomly select a subset of m alternatives in A denoted $A_m=\{a_{l1}, a_{l2},\ldots,a_{lm}\}$ with m<n (by using equi-distributed probabilities). The PROMETHEE II sub-ranking is computed on the basis of A_m. This step is repeated R times. The value p_{ij} is assessed as being the number of times a_i is ranked before a_j among the sub-rankings that integrate both these two alternatives. In the end, the P matrix is exploited by means of the net flow scores in order to get a complete ranking.

$$\widetilde{\emptyset}^{+}(a_i) = \frac{1}{n-1} \sum_{a_j \epsilon A} p_{ij} \quad (6)$$

$$\widetilde{\emptyset}^{-}(a_i) = \frac{1}{n-1} \sum_{a_j \epsilon A} p_{ji} \quad (7)$$

We have:

$$\widetilde{\emptyset}(a_i) = \widetilde{\emptyset}^{+}(a_i) - \widetilde{\emptyset}^{-}(a_i) \quad (8)$$

IV. ILLUSTRATION

In this section, we consider a simplified version of the Human Development Index (HDI - 2015). Only two criteria are considered : life expectancy at birth (in years – denoted LEB) and mean years of schooling (in years – denoted MYS). We only keep the 20 first ranked countries (see Table 1). Of course, this example is used to test the new method in terms of rank reversal sensitivity. The comparison between the actual HDI ranking and the ranking induced by PROMETHEE goes beyond the scope of this paper. Therefore, we do not discuss the preference parameter values in details (see table 2)

	LEB	**MYS**
Norway	81,6	12,6
Australia	82,4	13
Switzerland	83	12,8
Denmark	80,2	12,7
Netherlands	81,6	11,9
Germany	80,9	13,1
Ireland	80,9	12,2
United States	79,1	12,9
Canada	82	13
New Zealand	81,8	12,5
Singapore	83	10,6
Hong Kong, China	84	11,2
Liechtenstein	80	11,8
Sweden	82,2	12,1
United Kingdom	80,7	13,1
Iceland	82,6	10,6
Korea	81,9	11,9
Israel	82,4	12,5
Luxembourg	81,7	11,7
Japan	83,5	11,5

Table 1 – Top 20 ranked countries (HDI 2015)

	Indifference threshold	**Preference threshold**	**Weight**
LEB	0	3	0,5
MYS	0	3	0,5

Table 2 – PROMETHEE preference parameters

Based on these parameters, one can compute the rankings induced by PROMETHEE II and the new method (called Robust-PROMETHEE II)). These values are presented in table 3. Regarding the Robust-PROMETHEE II ranking, we have chosen the following values: R= 10.000 and m=5. This leads to a computation time that is approximatively equal to 0.13 seconds (on current computers).

	Scores PII	Rank	Scores RPII	
Switzerland	0,31491	1	1	**Switzerland**
Australia	0,25	2	0,89474	**Australia**
Canada	0,18246	3	0,73192	Hong Kong
Hong Kong	0,1807	4	0,68706	Canada
Japan	0,16316	5	0,58371	**Japan**
Israel	0,16228	6	0,52362	**Israel**
New Zealand	0,059649	7	0,34897	Sweden
Sweden	0,058772	8	0,2826	New Zealand
Norway	0,042105	9	0,15789	**Norway**
Germany	0,0078947	10	0,052632	**Germany**
United Kingdom	-0,025439	11	-0,070013	Korea
Korea	-0,02807	12	-0,14051	United Kingdom
Singapore	-0,071053	13	-0,28121	**Singapore**
Netherlands	-0,080702	14	-0,35036	**Netherlands**
Luxembourg	-0,098246	15	-0,47368	**Luxembourg**
Iceland	-0,13772	16	-0,59488	**Iceland**
Ireland	-0,15	17	-0,66827	**Ireland**
Denmark	-0,17632	18	-0,78947	**Denmark**
United States	-0,2886	19	-0,89474	**United States**
Liechtenstein	-0,36579	20	-1	**Liechtenstein**

Table 3 – Top 20 ranked countries (HDI 2015) – comparison between PROMETHEE II ranking (PII) and Robust-PROMETHEE II (RPII)

We see that both rankings are similar (countries marked in bold have the same rank). The correlation between the two series of net flow scores is equal to

0,9842. Only “successive” pairwise rank reversals are observed between the two rankings. For instance, Canada is ranked before Hong Kong in the PROMETHEE II ranking, while it is the opposite in the Robust-PROMETHEE II ranking.

First, let us focus ourselves on the PROMETHEE II ranking. Among the 190 (=20*19/2) possible parwise comparisons, 136 exceeded the bound 2/19 (see formula 6). In these cases, we are sure that rank reversal is impossible between these pairs of alternatives. In order, to test the other cases we ran the following experiment. An alternative is removed from *A*, the new PROMETHEE II ranking is computed and compared to the original one. This is done for each country one by one. Among the 20 tests, only 2 led to the same ranking (these correspond to the “individual” suppression of Israel and Japan respectively). All the others led to rank reversal occurrences. Of course, this observation is due to the fact that initial net flow scores were already quite similar. The “individual suppression” of the following countries led to a rank reversal between Japan and Israel: Switzerland, Austria, Canada, Hong Kong, New Zealand, Sweden, Norway, Germany, United Kingdom, Korea, Singapore, The Netherlands, Luxembourg, Iceland and Ireland. In most cases, Japan was initially (weakly) preferred to Israel. After, the suppression of the aforementioned countries, the two alternatives become indifferent. In addition;

- when suppressing Hong Kong, United Kingdom and Korea become indifferent;
- when suppressing Denmark, the ranking of Canada and Hong Kong is reversed;
- when suppressing United States, the ranking of Canada and Hong Kong is reversed – New Zealand and Sweden become indifferent;
- when suppressing Liechtenstein, the ranking of Canada and Hong Kong is reversed.

This experiment led to 20 rank reversal occurrences among the potential 54 that were not excluded by the bound (6). These rank reversals are only related to successive alternatives. Among these cases, only the relative ranking between the United States (-0,2886) and Lichtenstein (-0,36579) verifies the bound; -0,2886 – 0,36579=0,11228 > 0,1052.

By applying the same experimental procedure to the Robust-PROMETHEE method, **no** rank reversal was observed.

Of course, the Robust-PROMETHEE ranking heavily depends on the number of samples taken to compute the probability matrix. One can imagine that few samples may lead to different rankings. Therefore, we analyzed the number of rankings induced by different values of R (for the specific case m=5). Table 4 summarizes these statistics. We see that for low values of R, a lot of different rankings can be found. On the other hand, for R=5000 or R=10000, a unique ranking is determined (while the computation time is limited to 0,13 seconds on current computers).

R	Min	Max	Mean	Std
100	792	1000	938,057	87,318
200	259	997	380,6	245,19
500	54	937	96,8	158,834
1000	7	20	14,3	3,25
5000	1	1	1	0
10000	1	1	1	0

Table 2 – Robust-PROMETHEE II – statistics (based on 30 repetitions) about the number of different rankings for different number of sample sizes (R) – m=5.

VI. CONCLUSION

In this contribution, we presented an alternative way to compute a PROMETHEE II ranking. As illustrated on the example in section IV, the number of rank reversal instances is reduced to 0 (while it was equal to 20 in the traditional PROMETHEE II method). This observation constitutes a first experimental evidence that the new ranking is more robust in terms of rank reversal occurrences.

If this first result is encouraging, there are still a number of issues that need to be further investigated. The Robust-PROMETHEE II ranking depends on the sample size (m) and the number of repetitions (R). Of course, some combinations of these parameters might be more appropriate than others. An extensive empirical analysis has still to be conducted regarding this issue. This should also be linked to the probability of assessing each pair of alternatives properly.

REFERENCES

[1] R. Sarrazin and Y. De Smet “Design safer and greener road projects by using a multi-objective optimization evolutionary approach” in Internationl Journal of Multicriteria Decision Making, vol 6, n°1, 14-33, 2016

[2] N.A.V.Doan, D. Milojevic, F. Robert, F. and Y. De Smet “*A MOO-based methodology for designing 3D-stacked integrated circuits*” accepted (with minor modifications) in the Journal of Multi-Criteria Decision Analysis (2013) – vol. 21 (1-2), 43-63

[3] J.S. Dyer “*MAUT – Multiattribute Utility Theory*” in J. Figueira, S. Greco and M. Egrgott editors, Multiple Criteria Decision Analysis: State of the Art Surveys, 265-292, Springer Verlag, Boston, Dordrecht, London, 2005.

[4] T.L. Saaty “The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangible Criteria and for Decision-Making” in J. Figueira, S. Greco and M. Egrgott editors, Multiple Criteria Decision Analysis: State of the Art Surveys, 345-407, Springer Verlag, Boston, Dordrecht, London, 2005.

[5] J. Figueira, V. Mousseau, B. Roy « *ELECTRE methods* » in J. Figueira, S. Greco and M. Ergott editors, Multiple Criteria Decision Analysis: State of the Art Surveys, 133-162, Springer Verlag, Boston, Dordrecht, London, 2005.

[6] C.A. Bana e Costa, J.M. De Corte, J.C. Vansnick “*On the Mathematical Foundation of MACBETH”* in J. Figueira,

S. Greco and M. Egrgott editors, Multiple Criteria Decision Analysis: State of the Art Surveys, 409-442, Springer Verlag, Boston, Dordrecht, London, 2005.

[7] J.P. Brans and Y. De Smet "*Promethee Methods*" in J. Figueira, S. Greco and M. Egrgott editors, Multiple Criteria Decision Analysis: State of the Art Surveys, Second Edition 187-220, Springer Verlag, Boston, Dordrecht, London, 2016.

[8] M. Behzadian, R.B. Kazemzadh, A. Albadvi D. and M. Aghdasi, "*PROMETHEE: A comprehensive literature review on methodologies and applications*" European Journal of Operational Research, vol. 200(1), 198-215, 2010.

[9] Q. Hayez, Y. De Smet and J. Bonney "*D-Sight: a new decision making software to address multi-criteria problems*" the International Journal of Decision Support Systems Technologies, vol. 4, n°4, 1-23, 2012

[10] De Keyser, W. and Peeters, P. "A note on the use of PROMETHEE multicriteria methods" European Journal of Operational Research, 89 (1996), 457-461

[11] B. Mareschal, Y. De Smet and P. Nemery de Bellevaux, "*Rank reversal in the PROMETHEE II method: Some new results"* in proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, 8-11 dec., 2008, 959-963.

[12] Y. De Smet and C. Verly *"Some rank results about rank reversal instances in the PROMETHEE methods"* in the International Journal of Decision Making, vol. 3, n°4, 325-345, 2013