Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation ▮ (▮▮▮▮) ▮▮▮–▮▮▮
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

Multiple-criteria performance ranking based on profile distributions: An application to university research evaluations

Pierre L. Kunsch[a,*], Alessio Ishizaka[b]

[a] *Vrije Universiteit Brussel (VUB), BUTO Dept., Pleinlaan 2, Brussels, Belgium*
[b] *Portsmouth Business School, Centre for Operations Research and Logistics, University of Portsmouth, Portsmouth, PO1 3DE, United Kingdom*

Received 30 August 2016; received in revised form 19 May 2018; accepted 30 May 2018
Available online xxxxx

Highlights

- Profile Ranking with Order Statistics Evaluations (PROSE) is proposed.
- PROSE applies to criteria evaluated with profiles and not with single values.
- In PROSE the score values of the ordinal profile levels obey order statistics.
- Data from the UK Research Excellence Framework (REF) are used to test the approach.
- Research units are ranked with PROSE using the 2014 REF as a case study.

Abstract

This article addresses a category of multi-criteria ranking problems in which performance evaluations of the 'objects' or 'alternatives' to be ranked are not given by unique numbers but by performance profile distributions running over an ordered set of performance score levels. The numerical values to be given to score levels are not specified *a priori* using cardinal scales. A weighted sum approach is developed based on order statistics to combine the individual profile distributions. In this way, a global ranking indicator is obtained, considering not only mean distribution values but also standard deviations. As a test of feasibility, the resulting Profile Ranking with Order Statistics Evaluations (PROSE) approach has been applied to the performance profile distributions provided by the UK Research Excellence Framework 2014, evaluating the quality of UK research.

Keywords: Multi-criteria ranking; Performance profiles; Order statistics; Ranking indicator; 2014 Research Excellence Framework

1. Introduction

This article addresses a new category of multi-criteria ranking problems in which evaluations of the 'objects' or 'alternatives' are not given by unique numbers but by performance profile distributions normalised to one. Such

* Corresponding author.
E-mail addresses: pkunsch@vub.ac.be (P.L. Kunsch), alessio.ishizaka@port.ac.uk (A. Ishizaka).

https://doi.org/10.1016/j.matcom.2018.05.021

Please cite this article in press as: P.L. Kunsch, A. Ishizaka, Multiple-criteria performance ranking based on profile distributions: An application to university research evaluations, Mathematics and Computers in Simulation (2018), https://doi.org/10.1016/j.matcom.2018.05.021.

profile distributions have properties analogous to probability distributions; they run over a range of ordered ordinal score levels, e.g. for five score levels: *0 (not classified), 1 (marginal), 2 (mediocre), 3 (rather good), 4 (good)*. This type of problems is more and more common in practice because it is often difficult to give a crisp score. However, they have not been yet considered in the academic literature.

The question then arises if and how such ordinal score levels may be combined by simple arithmetic through a weighted sum to provide a global score, as is usual in many multi-criteria aggregation methods. Taking the weighted sum approach as given, we shall discuss in this paper the way in which the profile distributions are to be combined in a rigorous way: without anticipating the results, we conclude that a weighted sum mean value is not sufficient for ranking the 'objects' or 'alternatives' and thus a more elaborate ranking indicator must be considered. In addition, order statistics are introduced to value the ordinal score levels.

The applicability of the approach, called Profile Ranking with Order Statistics Evaluations (PROSE), is tested using a real-world case study: the ranking of UK research units based on data from the 2014 UK Research Excellence Framework (2014 REF) exercise [42]. The data resulting from the 2014 REF have exactly the format described, which to our knowledge is a unique case in university research evaluations. Before analysing this important case further, the problem can be made more concrete by means of a simple didactic example to present the ranking approach in PROSE. The most common and well-known example of this type is found in school evaluations. If the school teacher gives a pupil a fractional grade, e.g. 2.7, for a course on the ordinal 0–4 score scale, the thinking is: the pupil is well above the 'mediocre' grade 2, but has not entirely achieved the objectives of the 'rather good' grade 3 and therefore my subjective judgement is that the performance is at the level 2.7, i.e. 30% lower than 3, but 70% above 2. In this simple example, the normalised distribution is a binary one (i.e. 0% at 0, 0% at 1, 30% at 2, 70% at 3, 0% at 4); it is thus easy to verify that the mean grade value is indeed 2.7.

Distributions could have more than two non-vanishing components, for instance if grades are given by several school teachers for the same discipline. Considering several disciplines, the common – and probably unique school practice giving no room in general to second thoughts about its validity – is to combine the fractional grades between two ordered ordinal scores by discipline (i.e. criteria) using a weighted sum. Most of the time, all weights are taken to be equal, unless other directives are imposed by the school authorities: for instance, that more weight should be given to mathematics than to music. This does not call into question the weighted sum approach. But the weighted sum approach used as such in the specified setting, in addition to its lack of rigour in combining arithmetically ordinal level scores, which is not directly permissible, has the well-known drawback that only one value results in the form of an average final grade: therefore, compensation effects are possible between good and less good grades. Another pupil with the same average grade may be more 'regular' in the sense that the dispersion of the global grade distribution would be smaller, but this is generally not considered by school teachers. This article presents a way of avoiding this shortcoming by considering the spread in performance.

The article is structured as follows: Section 2 presents stepwise the proposed PROSE approach, using simple school examples which, as set out, have a comparable input data structure. A useful ranking indicator is introduced. Section 3 is dedicated to discussing the 2014 REF. First, a detailed introduction to how the profile data were obtained is provided. Profile distributions and weights are applied in the second part to test the PROSE approach developed in Section 2. Section 4 concludes the work and presents future research developments regarding multi-criteria ranking problems with a similar structure.

Two UK research rankings resulting from the 2014 REF data are given in Appendices A and B.

2. Ranking approach: PROSE

Mathematical notations are defined (Section 2.1); we address the basic weighted sum approach in multi-criteria problems (Section 2.2); we show by means of a school example how this method can be modified when the measures of performance are profile distributions rather than single values (Section 2.3.1). We extend the weighted sum to define the ordinal performance score levels as order statistics distributions (Section 2.3.2).

2.1. Mathematical notations and glossary

A bar above a variable indicates the mean value.
PROSE = **P**rofile **R**anking with **O**rder **S**tatistics **E**valuations
$k = 1, 2, \ldots, K$: running index of K criteria in decreasing order of importance

Please cite this article in press as: P.L. Kunsch, A. Ishizaka, Multiple-criteria performance ranking based on profile distributions: An application to university research evaluations, Mathematics and Computers in Simulation (2018), https://doi.org/10.1016/j.matcom.2018.05.021.