Information Fusion 17 (2014) 74–82

Contents lists available at SciVerse ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

A ranking procedure based on a natural monotonicity constraint

Michael Rademaker *, Bernard De Baets

KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, B-9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 29 April 2011
Received in revised form 20 December 2011
Accepted 6 January 2012
Available online 20 January 2012

Keywords:
Ranking
Social preference
Social ordering
Monotonicity

ABSTRACT

We formulate a new ranking procedure in the traditional context where each voter has expressed a linear order relation or ranking over the candidates. The final ranking of the candidates is taken to be the one which best adheres to a natural monotonicity constraint. For a ranking $a \succ b \succ c$, monotonicity implies that the strength with which $a \succ c$ is supported should not be less than the strength with which either one of $a \succ b$ or $b \succ c$ is supported. We investigate some properties of this ranking procedure and encounter some surprising preliminary results.

1. Introduction

We consider a ranking problem with r voters and k candidates, where each of the r voters has expressed a linear order relation over the k candidates, altogether making up a profile $\mathcal{R}$ (using the nomenclature from [1]). We develop a ranking procedure that chooses as the winning ranking the one which best adheres to the constraint of monotonicity. For a ranking $a \succ b \succ c$, monotonicity implies that the strength with which $a \succ c$ is supported should not be less than the strength with which either one of $a \succ b$ or $b \succ c$ is supported. The notion of strength of support is taken to be the number of voters who expressed the corresponding pairwise preference. It is inspired by the intuitive idea that the more clearly one candidate is better than another, the easier it is for voters to indicate the 'correct' preference. This could be natural when we suppose the existence of a 'true' ranking, of which the rankings expressed by the voters are imperfect observations, and the partially contradictory information thus obtained needs to be fused in a proper way. However, even if no 'true' ranking exists, in the case where personal preferences are the causal agents of the differences in rankings expressed by the voters, the approach we advocate makes sense, as we will discuss.

We describe the notations and conventions we adhere to in Section 2. In this section we also describe the main contribution, the ranking procedure based on monotonicity (in the remainder of the text, we will omit "based on monotonicity" when no confusion is possible). The underlying intuition is discussed in Section 3. Section 4 contains an investigation into some properties of the proposed ranking procedure. Some related concepts identified in previous work are mentioned in Section 5. Finally, conclusions and an outlook are provided in Section 6.

2. Problem setting and formulation of the ranking procedure

2.1. Preliminaries

We consider a set $\mathcal{C}$ of candidates $\{a, \dots, k\}$ and suppose a profile $\mathcal{R}$ of linear order relations $\{\succ_1, \dots, \succ_r\}$ on the set of candidates $\mathcal{C}$. An example $\succ$ for a three-alternative problem could be $\{(a,b), (a,c), (a,d), (b,c), (b,d), (c,d)\}$, i.e., $a \succ b \succ c \succ d$. We denote the number of times candidate a is preferred to candidate b by

$$F_{\mathcal{R}}(a,b) = |\{\succ_i \in \mathcal{R} | a \succ_i b\}|. \quad (1)$$

One could also formulate $F_i(a,b)$ for each voter i, trivially taking value 1 if $a \succ_i b$ and 0 otherwise. It then also holds that

$$F_{\mathcal{R}}(a,b) = \sum_{i=1}^{r} F_i(a,b). \quad (2)$$

In other words, $F_{\mathcal{R}} : \mathcal{C}^2 \to \{0, 1, \dots, r\}$. It is natural to call $F_{\mathcal{R}}(a,b)$ the strength of support for (a,b). Due to $\mathcal{R}$ being a collection of linear order relations, it holds that

$$\forall a \neq b \in \mathcal{C} : F_{\mathcal{R}}(a,b) + F_{\mathcal{R}}(b,a) = r. \quad (3)$$

If $F_{\mathcal{R}}(a,b) > r/2$, we will write $a \succ_{\text{m}} b$ to denote that there is a strict majority for a over b. When $F_{\mathcal{R}}(a,b) \geqslant F_{\mathcal{R}}(b,a)$, we say (a,b) is supported.

Trivially, to each linear order relation $\succ_i$ corresponds a total weak order relation $\succcurlyeq_i$. Furthermore, each linear order relation $\succ_i$

* Corresponding author.
E-mail addresses: michael.rademaker@ugent.be (M. Rademaker), bernard.debaets@ugent.be (B. De Baets).

1566-2535/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.inffus.2012.01.003

also implies a strict partial order relation $\triangleright_i$ on $\mathcal{C}^2 \setminus \{(a,a)|a \in \mathcal{C}\}$ defined by:

$$(a,d) \triangleright_i (b,c) \overset{\text{def}}{\Longleftrightarrow} ((a \succcurlyeq_i b) \wedge (c \succcurlyeq_i d)) \wedge ((a \succ_i b) \vee (c \succ_i d)). \tag{4}$$

To this strict partial order relation corresponds the trivially satisfied monotonicity constraint

$$\forall (a,d) \triangleright_i (b,c) : F_i(a,d) \geqslant F_i(b,c). \tag{5}$$

Obviously, if $a \succ_i b$, then $(a,b) \triangleright_i (b,a)$ and $1 = F_i(a,b) \geqslant F_i(b,a) = 0$.

Before we are able to define what it means for a ranking to be optimal, we finally need to introduce the natural monotonicity constraint. Condition (5) stipulates what is known about each F_i with respect to the corresponding $\succ_i$. We now will extend this to a condition on $F_\mathcal{R}$ with respect to a given ranking $\succ$. For a given ranking $\succ$ and corresponding $\triangleright$, the natural monotonicity constraint on $F_\mathcal{R}$ is the following:

$$\forall (a,d) \triangleright (b,c) : F_\mathcal{R}(a,d) \geqslant F_\mathcal{R}(b,c). \tag{6}$$

Condition (6) is now no longer trivially satisfied, as the voters need not agree on the rankings of the alternatives. Yet, even in the absence of unanimity, condition (6) can still be satisfied. When condition (6) is not satisfied for any ranking $\succ$, we will have to look for an $F : \mathcal{C}^2 \to \{0,1,\ldots,r\}$ which does satisfy condition (6). Naturally, we would prefer to have this F stay close to $F_\mathcal{R}$, with the distance between them given by $d(F_\mathcal{R}, F) = \sum_{a,b \in \mathcal{C}} |F(a,b) - F_\mathcal{R}(a,b)|$. More formally:

Definition 1. For a given $F_\mathcal{R}$ and a ranking $\succ$, a *closest monotone* $F : \mathcal{C}^2 \to \{0,1,\ldots,r\}$ is one satisfying:

(MON) $\forall (a,d) \triangleright (b,c) : F(a,d) \geqslant F(b,c)$,
(REC) $\forall a \neq b : F(a,b) + F(b,a) = r$,
(OPT) $\nexists F' : \mathcal{C}^2 \to \{0,1,\ldots,r\}$ satisfying the above conditions while $d(F_\mathcal{R}, F') < d(F_\mathcal{R}, F)$.

Observe that we only refer to F and $F_\mathcal{R}$ in the definition above, and not to $\mathcal{R}$ itself: As with many ranking procedures, the original rankings are no longer used at this point. The monotonicity condition (MON) is a straightforward analog to condition (6), and demands that F be monotone w.r.t. the ranking $\succ$. Next, the reciprocity condition (REC) guarantees the total number of votes will remain constant. Finally, (OPT) guarantees we stay as close as possible to $F_\mathcal{R}$. As $F_\mathcal{R}(a,b)$ is a quantification of the number of times voters preferred a to b, adapting $F_\mathcal{R}$ amounts to changing preferences by reversing them. Thus, we will also use the notion 'reversing changes', and 'minimizing the number of reversing changes'. By determining the number of reversing changes that is needed in order to have a given $F_\mathcal{R}$ be monotone w.r.t. a given ranking $\succ$, we are able to quantify how close a given $F_\mathcal{R}$ is to being monotone w.r.t. this ranking. This leads us to the following natural definition of a winning ranking:

Definition 2. For a given $F_\mathcal{R}$, an *optimal ranking* is a ranking $\succ$ (with a corresponding closest monotone F) for which it holds that there exists no ranking $\succ'$ (with a corresponding closest monotone F') while $d(F_\mathcal{R}, F') < d(F_\mathcal{R}, F)$.

In other words, an optimal ranking is a ranking for which only a minimal number of reversing changes need to be made to $F_\mathcal{R}$ to render it monotone w.r.t. it. In the remainder of the text, we will denote such an optimal ranking by $\succ_\mathcal{R}$.

As a final remark, it is important to point out the relation of the monotonicity condition to the strong stochastic transitivity condition (in our context, strong stochastic transitivity means that if both (a,b) and (b,c) are supported, the strength of support for (a,c) is at least as strong as the strongest support between (a,b) and (b,c)): When $F_\mathcal{R}$ is strong stochastic transitive, there exists a ranking w.r.t. it is monotone. For more information in this context, see our previous work on voting with intensities of preference [2]. We will also refer to the stochastic transitivity conditions in Section 5.

2.2. The ranking procedure

It will be clear now that the problem of minimizing the number of reversing changes is, in fact, a non-monotonicity problem: The ranking $\succ_\mathcal{R}$ yields the strict partial order $\triangleright_\mathcal{R}$, represented for a four-alternative problem in Fig. 1, where the monotonicity constraint arises due to the demand that $F_\mathcal{R}$ (or F) should not increase on any downward path in Fig. 1. After all, an increase on a downward path would mean that (MON) is violated. By now, it will be clear that we could call a thus determined $\succ_\mathcal{R}$ a 'monotonicity-based ranking'. However, we will use the less cumbersome name 'optimal ranking', as defined before.

We have exhaustively described how to solve (stochastic) non-monotonicity problems in previous work [3–5], and we will not discuss how to construct such an F here. In [3–5], we show the problem to be a network flow problem, give clear directions how to translate the original problem to one solvable by maximum flow techniques, and point to the available algorithms to do so. In the current setting we are able to employ regular monotonicity instead of stochastic monotonicity, in contrast to a setting where voters could assign intensities of preference to their votes, as in [2,5–8].

In the remainder of the paper, we will rather focus on the voting aspect of the problem, the search for an $\succ_\mathcal{R}$ which best adheres to condition (6) for the given profile $\mathcal{R}$. In other words, we aim to determine to which ranking $\succ_\mathcal{R}$ (out of all possible rankings) corresponds a closest monotone F that is closest to $F_\mathcal{R}$. Regrettably, to a single profile $\mathcal{R}$ can correspond multiple optimal rankings $\succ_\mathcal{R}$. If there are indeed multiple optimal rankings, we will output all of them.

A first approach to determine such a ranking $\succ_\mathcal{R}$ would then be to examine each possible ranking and corresponding closest monotone F, and select the optimal one(s). Though examining each possible ranking renders the problem factorial in size, this is not prohibitive in practice due to the limited number of candidates in voting problems. Some electoral methodologies also examine each possible ranking of the candidates, such as a method based on Kendall's Tau distance [9] and the Kemeny–Young method

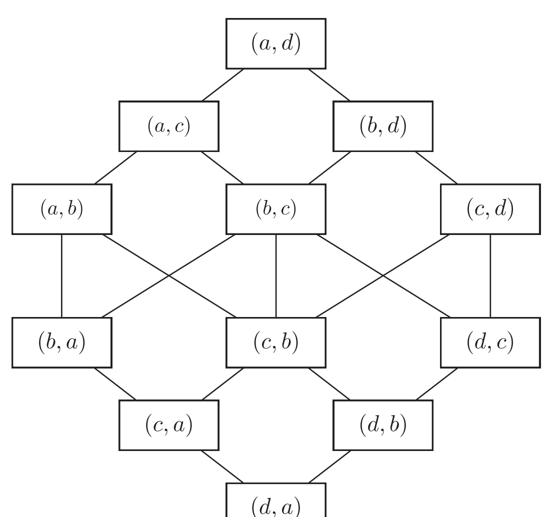

Fig. 1. Graphical representation of $\triangleright_i$ for a ranking $a \succ_i b \succ_i c \succ_i d$, with a couple (x,y) located above a couple (u,v) if $(x,y) \triangleright_i (u,v)$.