On Stability of Aggregation Procedures

Daniel Karabekyan and Vyacheslav Yakuba

Abstract

We consider the problem of aggregation of several rankings of the same objects. For example, different decisions can be ranked by their economic influence, safety, social impact and so on. Or program committee members individually rank the papers. In all these cases we need to come to the final decision taking into account individual rankings. There are some well-known aggregation procedures that come from social choice literature like Plurality, Approval or Borda rules. While it is obvious that each ranking should influence the final decision, a high sensitivity of the rules for small changes is not a good property, for example when those changes are made by mistakes in individual rankings or by attempts of individual members to pursue more preferable result for them. We study 20 aggregation procedures and compare them according to their stability. We found that Threshold rule is the most stable one in most cases. The results can be used in developing decision support systems.

1 INTRODUCTION

Consider a decision support system where one needs to make a choice using multiple criteria that rank several possible alternatives. There is a lot of literature on aggregation procedures and multi-criteria decision making (see a comprehensive overview in Grabisch et al. 2009 and Triantaphyllou 2000). Most of the methods differ in terms of properties and requirements. In general, when all criteria have equal weights in the system and each of them is a ranking one can use some of the aggregation procedures known from social choice literature (Munda 2012). Since all those systems are designed to take each ranking into account as much as possible, one can pose a question: how stable the final solution is if one of the rankings is inaccurate. In this paper, we try to address this question and compare well-known social choice rules from the stability point of view.

The closest topic in the social choice literature is the problem of manipulation in voting when one agent can declare insincere preferences and get better aggregation result. Many papers compare the rules according to the degree of manipulability: share of all manipulable situations (Kelly 1993; F. Aleskerov and Kurbanov 1999; Nitzan 1985; Favardin and Lepelley 2006; Pritchard and Wilson 2007). The main difference in our case is that we do not assume that behind each criterion there is an agent who wants to reach a better aggregation result. In our model ranking can be changed by random mistake and even make the final result even worse. We estimate the effects of those changes and compare rules according to the stability of the final results.

Our contribution. Our notion of stability coincides with the degree of non-sensitivity to a preference change from F. Aleskerov, Karabekyan, et al. 2011. At the same time in that paper, authors concentrate on the indices of degree and freedom of manipulation and do not compare the rules according to the non-sensitivity index. In this paper, we study a broader set of rules, including majority-relation based rules and consider up to 5 alternatives. Our work provides a new way of comparing aggregation procedures.

1.1 COMPUTATION SCHEME

Take some ranking profile j. For each individual criterion i, there are $m! - 1$ of possible incorrect rankings. Let t_{ij} be the number of rankings that do not influence the aggregation results. Assume that all incorrect rankings have the same probability, then $t_{ij}/(m! - 1)$ is the probability of non-influencing the result. Since this mistake can happen for each criterion and we consider the

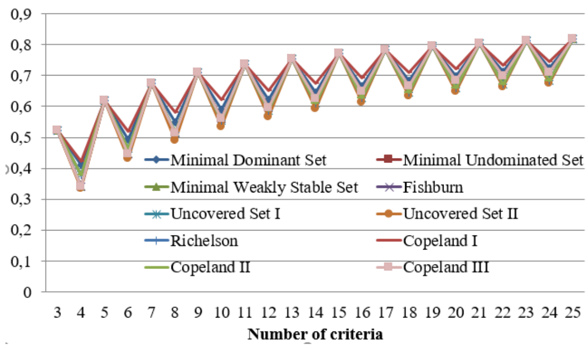

Figure 1: Stability for majority-relation based rules and 3 alternatives

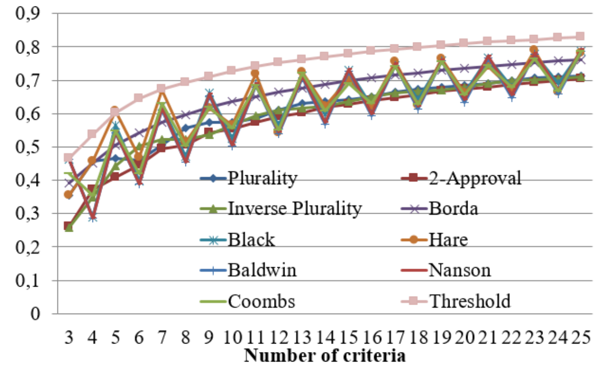

Figure 2: Stability for positional rules and 5 alternatives

Impartial Culture: all possible rankings are equally probable, then the final index is the following

$$I_{st} = \frac{\sum_{j=1}^{(m!)^n} \sum_{i=1}^{n} t_{ij}}{(m!)^n \cdot n \cdot (m! - 1)} \quad (1)$$

As we mentioned earlier it coincide with non-sensitivity index from F. Aleskerov, Karabekyan, et al. 2011. We use Monte Carlo simulations to estimate indices. We generate 1 mln possible ranking for 3, 4 or 5 alternatives and up to 25 criteria and compare rules by their stability. Results of estimation is given in the next section.

Table 1: Most stable aggregation rules

	Alternatives		
Criteria	3	4	5
3	Black, Baldwin, Nanson, all majoritarian rules	Nanson	Threshold
4	Threshold	Threshold	Threshold
5	Hare	Hare	Hare
6	Threshold	Threshold	Threshold
7	Hare	Hare	Threshold
8	Threshold	Threshold	Threshold
9	Black	Threshold	Threshold
10	Threshold	Threshold	Threshold
11	Hare	Hare	Threshold
12	Threshold	Threshold	Threshold
13	Coombs	Threshold	Threshold
14	Threshold	Threshold	Threshold
15	Baldwin, Nanson	Threshold	Threshold
16	Threshold	Threshold	Threshold
17	Hare	Threshold	Threshold
18	Threshold	Threshold	Threshold
19	Coombs	Threshold	Threshold
20	Threshold	Threshold	Threshold
21	Baldwin, Nanson	Threshold	Threshold
22	Threshold	Threshold	Threshold
23	Hare	Threshold	Threshold
24	Threshold	Threshold	Threshold
25	Coombs	Threshold	Threshold

2 DISCUSSION AND FURTHER RESEARCH

At the following figures and tables, one can see how the stability of the rules change when the number of criteria rises. It is obvious that there is a rising trend in the index because the more criteria there are the less is the influence of only one. Stability dependence of even and odd number of criteria is based on the probability of ties. If the true decision is tied, then small mistakes have more influence on the final result. For odd number or criteria majority relation based rules almost always give the same results because majority relation is complete in this case and Condorcet winner always exists.

We find out that Threshold rule introduced in F. T. Aleskerov, Yuzbashev, and Yakuba 2007 is one of the most stable rules for even number of criteria. There are two reasons behind it. First one is that it requires not much information and mistakes that do not affect the worst alternative according to the criteria in many cases will not influence the decision. At the same time, Plurality or Antiplurality rules that also require little information about rankings are not stable. The second reason is that it is highly decisive due to the built-in tie-breaking mechanism. In the case of even number of criteria all other rules (especially majority-relation based rules) in end up with tied decisions more often. For odd number of criteria with except of 9 voters we see Elimination rules (Hare, Baldwin, Coombs, Nanson) among the most stable one. When the number of alternatives goes up, Threshold rule begins to dominate all other rules.

We need to mention the limitations of our research. First of all, our definition of error can be extended to be closer to reality. If we consider ranking as the result of some estimation, it is better to assume that the submitted ranking will differ from the accurate one only in several pairs. In this paper, we consider that small change has the same probability as the complete reversal of ranking. The second extension is to look at what rules are stable for 6 or more alternatives. Together with the change of the probability of error proposed before it will probably make the rules that require less information about lower ranks (like Plurality) more stable. The third direction of further research is to study the case when mistakes in several rankings happen at the same time and connect the results with the literature on preference diversity orderings (Karpov 2017). The fourth direction is to look at the joint distribution of individual rankings since in some cases there can be situations when an alternative that is better by one criterion is more probable to be better by another. One can consider the Impartial Anonymous Culture model here.

References

Aleskerov, F. T., D. A. Yuzbashev, and V. I. Yakuba (2007). "Threshold aggregation of the three-graded rankings". In: *Automation and Remote Control* 68.1, pp. 133–138. ISSN: 1608-3032. DOI: 10.1134/S0005117907010122. URL: https://doi.org/10.1134/S0005117907010122.

Aleskerov, Fuad, Daniel Karabekyan, et al. (2011). "An individual manipulability of positional voting rules". en. In: *SERIEs* 2.4, pp. 431–446. ISSN: 1869-4187, 1869-4195. DOI: 10.1007/s13209-011-0050-y. URL: http://link.springer.com/10.1007/s13209-011-0050-y.

Aleskerov, Fuad and Eldeniz Kurbanov (1999). "Degree of manipulability of social choice procedures". In: *Current trends in economics*. Springer, pp. 13–27.

Favardin, Pierre and Dominique Lepelley (2006). "Some Further Results on the Manipulability of Social Choice Rules". en. In: *Social Choice and Welfare* 26.3, pp. 485–509. ISSN: 0176-1714, 1432-217X. DOI: 10.1007/s00355-006-0106-2. URL: http://link.springer.com/10.1007/s00355-006-0106-2.

Grabisch, Michel et al. (2009). *Aggregation functions*. Vol. 127. Cambridge University Press.

Karpov, Alexander (2017). "Preference Diversity Orderings". In: *Group Decision and Negotiation* 26.4, pp. 753–774. ISSN: 1572-9907. DOI: 10.1007/s10726-017-9532-z. URL: https://doi.org/10.1007/s10726-017-9532-z.

Kelly, Jerry S (1993). "Almost all social choice rules are highly manipulable, but a few aren't". In: *Social choice and Welfare* 10.2, pp. 161–175.

Munda, Giuseppe (2012). "Choosing Aggregation Rules for Composite Indicators". In: *Social Indicators Research* 109.3, pp. 337–354. ISSN: 1573-0921. DOI: 10.1007/s11205-011-9911-9. URL: https://doi.org/10.1007/s11205-011-9911-9.

Nitzan, Shmuel (1985). "The vulnerability of point-voting schemes to preference variation and strategic manipulation". en. In: *Public Choice* 47.2, pp. 349–370. ISSN: 0048-5829, 1573-7101. DOI: 10.1007/BF00127531. URL: http://link.springer.com/10.1007/BF00127531.

Pritchard, Geoffrey and Mark C Wilson (2007). "Exact results on manipulability of positional voting rules". In: *Social Choice and Welfare* 29.3, pp. 487–513.

Triantaphyllou, Evangelos (2000). "Multi-criteria decision making methods". In: *Multi-criteria decision making methods: A comparative study*. Springer, pp. 5–21.