Simulating the Effects of Misperception on the Manipulability of Voting Rules

Johann Mitlöhner

Vienna University of Economics, Austria

mitloehn@wu-wien.ac.at

Daniel Eckert, Christian Klamler

University of Graz, Austria

{daniel.eckert,christian.klamler}@uni-graz.at

Motivation

- Manipulability of voting rules

  (Gibbard-Satterthwaite)

- How to evaluate and measure manipulability (in order to compare voting rules)

- How many profiles are manipulable (degree of manipulability)

  Kelly (1993), Smith (1999)

- Efficiency of manipulation (degree of improvement)

  Smith (1999), Aleskerov/Kurbanov (1999)

Assumptions

In the analysis of manipulability usually 2 assumptions:

- unlimited computational capacity
- perfect information

Robustness of manipulation under uncertainty

- relaxation of perfect information (noisy profiles)
- given small misperceptions of the profile, will manipulation still be advantageous?
- what are the consequences for the evaluation of different voting rules?

Simulation Model

profile p: n strict orders over a set A of m candidates

e.g. $p = ((a \succ b \succ c \succ d), (b \succ c \succ a \succ d), (c \succ a \succ b \succ d))$

social choice correspondence $C(p) \subseteq A$

average rank $r_i(C(p))$ of elements of $C(p)$ in p_i

100.000 random profiles with 5 voters and 4 candidates
voting rules implemented in Python:

Borda (BO), Copeland (CO), Kemeny (KE), Plurality (PL), Antiplurality (AP), Transitive Closure (TC), Maximin (MM), Slater (SL), Nanson (NA), Young (YO), and Dodgson (DO)

Aggregate relations x encoded as binary matrices denoting weak preference i.e. if $x_{i,j} = 1$ and $x_{j,i} = 0$ then $c_i \succ c_j$; if $x_{i,j} = x_{j,i} = 1$ than $c_i = c_j$

[[1,1,1],[1,1,1],[0,0,1]] denotes aggregate ranking $((a = b) \succ c)$

Sample random profile and aggregate rankings

```
pr: a>b>c>d,c>a>d>b,c>a>b>d,b>d>c>a,b>c>d>a
BO: [[1,0,0,1],[1,1,0,1],[1,1,1,1],[0,0,0,1]]    c>b>a>d
CO: [[1,1,1,1],[1,1,1,1],[1,1,1,1],[0,0,0,1]]    a=b=c>d
TC: [[1,1,1,1],[1,1,1,1],[1,1,1,1],[0,0,0,1]]    a=b=c>d
NA: [[1,0,0,1],[1,1,1,1],[1,0,1,1],[0,0,0,1]]    b>c>a>d
MM: [[1,0,0,1],[1,1,1,1],[1,1,1,1],[1,0,0,1]]    b=c>a=d
KE: [[1,0,0,1],[1,1,1,1],[1,0,1,1],[0,0,0,1]]    b>c>a>d
SL: [[1,1,1,1],[0,1,1,1],[0,0,1,1],[0,0,0,1]]    a>b>c>d
YO: [[1,0,0,1],[1,1,1,1],[1,1,1,1],[1,0,0,1]]    b=c>a=d
DO: [[1,0,0,1],[1,1,1,1],[1,1,1,1],[0,0,0,1]]    b=c>a>d
PL: [[1,0,0,1],[1,1,1,1],[1,1,1,1],[0,0,0,1]]    b=c>a>d
AP: [[1,0,0,1],[1,1,0,1],[1,1,1,1],[1,0,0,1]]    c>b>a=d
```

Execution Times

seconds for 1000 random profiles with $n = 9$ voters and $m = 4, 5, 6, 7, 8$ candidates, on 3.2 GHz Intel Pentium D

Rule	$m = 4$	$m = 5$	$m = 6$	$m = 7$	$m = 8$
BO	0.07	0.07	0.08	0.09	0.10
CO	0.09	0.09	0.10	0.11	0.13
PL	0.07	0.07	0.08	0.09	0.10
AP	0.07	0.07	0.08	0.09	0.10
MM	0.08	0.10	0.11	0.13	0.14
NA	0.08	0.09	0.10	0.11	0.13
TC	0.15	0.25	0.43	0.72	1.21
YO	1.21	1.93	2.97	3.71	5.15
KE	0.14	0.52	3.58	31.61	318.46
SL	0.15	0.47	3.03	26.20	253.14
DO	2.32	12.31	51.98	160.56	464.17

Manipulation

true preferences p_i

manipulated preferences p'_i

manipulated profile p'

successful manipulation: $p'_i \neq p_i$ with $C(p') \neq C(p)$

and average rank $r_i(C(p')) < r_i(C(p))$

rank difference $d_i(C(p), C(p')) = r_i(C(p')) - r_i(C(p))$

e.g. $p_i = (a \succ b \succ c),\ C(p) = \{b\},\ C(p') = \{a, b\}$

$d_i = 1.5 - 2 = -0.5$

Misperception

remaining profile p_{-i} perceived as noisy p^e_{-i}

e: pairwise exchanges

e.g. $p_{-i} = ((b \succ c \succ a \succ d), (a \succ c \succ b \succ d))$

with $e = 1$ misperceptions $p^e_{-i} = ((b \succ c \succ a \succ d), (c \succ a \succ b \succ d))$

simulation study: for each true profile p
create noisy profile p^e and observe for $i = 1$

$$\begin{aligned} d_i(C(p^e), C(p'^e)) &< 0 \quad \textit{manipulation} \\ d_i(C(p), C(p')) &< 0 \quad \textit{success} \qquad (1) \\ d_i(C(p), C(p')) &> 0 \quad \textit{failure} \qquad (2) \\ d_i(C(p), C(p')) &= 0 \quad \textit{no effect} \qquad (3) \end{aligned}$$

Expected Changes in Rank Differences

success in noisy profile if $r_i(C(p'^e)) < r(C(p^e))$

$E^M(d) = \frac{1}{|M|} \sum_{p \in M} d_i(C(p^e), C(p'^e))$

success in true profile if $r_i(C(p')) < r(C(p))$

$E^S(d) = \frac{1}{|S|} \sum_{p \in S} d_i(C(p), C(p'))$

failure in true profile if $r_i(C(p')) > r_i(C(p))$

$E^F(d) = \frac{1}{|F|} \sum_{p \in F} d_i(C(p), C(p'))$

expected benefit in true profile

$E(d) = \frac{1}{|M|} \sum_{p \in M} d_i(C(p), C(p'))$

where $M = S + F + O$

Expected Benefit and Punishment Effect

rule	M	EM(d)	S	ES(d)	F	EF(d)	E(d)	PU
AP	26123	-0.6642	21392	-0.6732	1599	0.6995	-0.5085	0.1557
BO	28661	-0.7354	14069	-0.7195	2538	0.6858	-0.2924	0.4430
CO	8941	-0.5221	4259	-0.5409	1462	0.6367	-0.1535	0.3686
DO	10252	-0.5707	4540	-0.5783	1640	0.6283	-0.1556	0.4151
KE	7952	-1.3462	3441	-1.3188	1088	1.2188	-0.4039	0.9423
MM	9225	-0.4527	4761	-0.4507	823	0.7935	-0.1618	0.2909
NA	9477	-0.9920	3518	-0.9506	1371	1.0581	-0.1998	0.7922
PL	16382	-0.7150	12053	-0.7145	1851	0.8071	-0.4345	0.2805
SL	7833	-1.3476	3639	-1.3325	1045	1.2593	-0.4510	0.8966
TC	7342	-0.4164	3821	-0.4209	921	0.9767	-0.0965	0.3199
YO	9052	-0.4526	4686	-0.4496	792	0.7777	-0.1647	0.2879

Interpretation and Conclusions

- Among the scoring rules (BO, PL, AP) BO has the lowest expected benefit of manipulation and the highest punishment effect
- Most rules have higher expected loss in case of failure than expected benefit in case of successful manipulation: risk averse individuals would not manipulate (exception: BO, KE, SL)
- More decisive rules (single winner: KE, SL) exhibit higher punishment effect
- The data also show that the punishment effect for manipulation with misperception is not a rare exceptional case. It occurs frequently enough to provide an additional dimension for the evaluation and comparison of voting rules.

Summary and Future Work

In simulations with a limited range of parameters we have explored

- the extent to which manipulators can lose rather than gain from manipulation in a setting with misperception and
- the susceptibility of various rank aggregation rules to these effects.

Future work will

- test the validity of these results for a wider range of parameters and
- expand the range of applications of the software package developed for the simulations.

`http://prefrule.sf.net`

SE

rule	SES(d)	SEF(d)	SEM(d)	SE(d)
AP	0.0037	0.0216	0.0033	0.0056
BO	0.0042	0.0151	0.0030	0.0058
CO	0.0069	0.0136	0.0042	0.0097
DO	0.0075	0.0142	0.0047	0.0093
KE	0.0178	0.0256	0.0120	0.0213
MM	0.0072	0.0198	0.0047	0.0085
NA	0.0196	0.0246	0.0118	0.0158
PL	0.0044	0.0138	0.0038	0.0086
SL	0.0174	0.0282	0.0122	0.0221
TC	0.0033	0.0178	0.0026	0.0107
YO	0.0072	0.0206	0.0047	0.0085