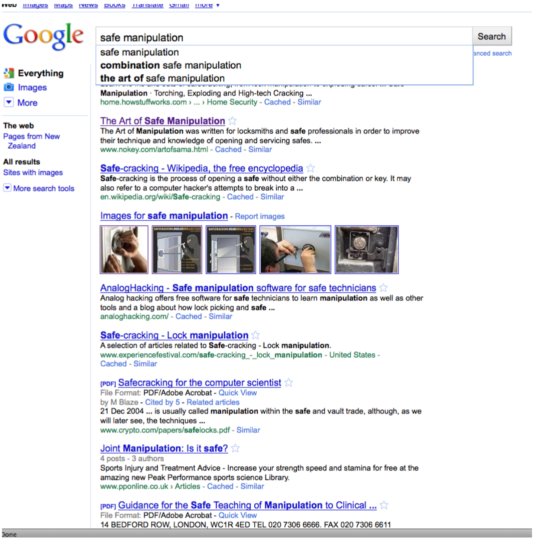
The probability of safe manipulation

Mark C. Wilson
www.cs.auckland.ac.nz/~mcw/blog/
(joint with Reyhaneh Reyhani)

Department of Computer Science
University of Auckland

COMSOC, Düsseldorf, 2010-09-16

What Google thinks this talk is about



The University of Auckland
NEW ZEALAND

Basic setup

- A set $\mathcal{C}$ of alternatives (candidates) of size m, and a set $\mathcal{V}$ of voters, of size n.

Basic setup

- A set $\mathcal{C}$ of alternatives (candidates) of size m, and a set $\mathcal{V}$ of voters, of size n.
- Each voter v has a type (sincere preference) and submits an expressed preference. These are permutations L_v of the candidates.

Basic setup

- A set $\mathcal{C}$ of alternatives (candidates) of size m, and a set $\mathcal{V}$ of voters, of size n.
- Each voter v has a type (sincere preference) and submits an expressed preference. These are permutations L_v of the candidates.
- A profile is a function $\mathcal{V} \to \mathcal{T}$. A voting situation is a multiset from $\mathcal{T}$ with total weight n.

Basic setup

- A set $\mathcal{C}$ of alternatives (candidates) of size m, and a set $\mathcal{V}$ of voters, of size n.
- Each voter v has a type (sincere preference) and submits an expressed preference. These are permutations L_v of the candidates.
- A profile is a function $\mathcal{V} \to \mathcal{T}$. A voting situation is a multiset from $\mathcal{T}$ with total weight n.
- The positional scoring rule determined by a vector w with $w_1 \geq w_2 \geq \cdots \geq w_{m-1} \geq w_m$ assigns the usual score

$$|c| := \sum_{t \in \mathcal{T}} |\{v \in \mathcal{V} \mid L_v = t\}| w_{L_v^{-1}(c)}.$$

The University of Auckland NEW ZEALAND

Basic setup

- A set $\mathcal{C}$ of alternatives (candidates) of size m, and a set $\mathcal{V}$ of voters, of size n.
- Each voter v has a type (sincere preference) and submits an expressed preference. These are permutations L_v of the candidates.
- A profile is a function $\mathcal{V} \to \mathcal{T}$. A voting situation is a multiset from $\mathcal{T}$ with total weight n.
- The positional scoring rule determined by a vector w with $w_1 \geq w_2 \geq \cdots \geq w_{m-1} \geq w_m$ assigns the usual score

$$|c| := \sum_{t \in \mathcal{T}} |\{v \in \mathcal{V} \mid L_v = t\}| w_{L_v^{-1}(c)}.$$

- In this talk tiebreaking is mostly not relevant, so we ignore it completely.

Manipulation

- Standard social choice definition: a voter expresses an insincere preference to achieve a better outcome than otherwise, assuming other voters vote sincerely. This is individual manipulation.

Manipulation

- Standard social choice definition: a voter expresses an insincere preference to achieve a better outcome than otherwise, assuming other voters vote sincerely. This is individual manipulation.
- Coalitional manipulation occurs when a subset X of $\mathcal{V}$ all simultaneously adopt the above strategy. Their expressed preferences need not be the same, nor their sincere preferences. However all must (weakly) prefer the new outcome to the sincere one.

Manipulation

- Standard social choice definition: a voter expresses an insincere preference to achieve a better outcome than otherwise, assuming other voters vote sincerely. This is individual manipulation.
- Coalitional manipulation occurs when a subset X of $\mathcal{V}$ all simultaneously adopt the above strategy. Their expressed preferences need not be the same, nor their sincere preferences. However all must (weakly) prefer the new outcome to the sincere one.
- There is no claim that such strategic voting will take place, just that there is incentive to consider it.

Difficulties with coalitional manipulation

- How do coalition members identify each other?

Difficulties with coalitional manipulation

- How do coalition members identify each other?
- How do coalition members communicate?

Difficulties with coalitional manipulation

- How do coalition members identify each other?
- How do coalition members communicate?
- How do coalition members compute their joint strategy?

Difficulties with coalitional manipulation

- How do coalition members identify each other?
- How do coalition members communicate?
- How do coalition members compute their joint strategy?
- How do coalition members enforce the strategy?

The concept of safe manipulation

- A voter of type t (the leader) announces that (s)he will in fact express the preference t'.

The concept of safe manipulation

- A voter of type t (the leader) announces that (s)he will in fact express the preference t'.
- We assume that only voters of type t hear this message, and other voters vote sincerely. The type t voters can either vote as t or t'. Let x denote the number who switch to t'.

The University of Auckland NEW ZEALAND

The concept of safe manipulation

- A voter of type t (the leader) announces that (s)he will in fact express the preference t'.
- We assume that only voters of type t hear this message, and other voters vote sincerely. The type t voters can either vote as t or t'. Let x denote the number who switch to t'.
- The announced vote is safe if for all x, the outcome is never worse for these voters. In particular this applies to the maximal manipulation, where all voters of type t switch. Note that a voter who ranks the sincere winner lowest can never vote unsafely.

The concept of safe manipulation

- A voter of type t (the leader) announces that (s)he will in fact express the preference t'.
- We assume that only voters of type t hear this message, and other voters vote sincerely. The type t voters can either vote as t or t'. Let x denote the number who switch to t'.
- The announced vote is safe if for all x, the outcome is never worse for these voters. In particular this applies to the maximal manipulation, where all voters of type t switch. Note that a voter who ranks the sincere winner lowest can never vote unsafely.
- If in addition there is some x for which the outcome is better for these voters, the profile is safely manipulable by type t in direction t'.

Safe manipulation nonexample

- Let $m = 5$ and use $w = (55, 39, 33, 21, 0)$. Suppose that there are 3 voters of each possible type, and 1 extra voter of type 12345. The sincere winner is alternative 1.

Safe manipulation nonexample

- Let $m = 5$ and use $w = (55, 39, 33, 21, 0)$. Suppose that there are 3 voters of each possible type, and 1 extra voter of type 12345. The sincere winner is alternative 1.
- If 1 type 53124 voter votes instead as 35241, alternative 2 wins; if 2 switch, alternative 3 wins; if 3 switch, alternative 4 wins.

Safe manipulation nonexample

- Let $m = 5$ and use $w = (55, 39, 33, 21, 0)$. Suppose that there are 3 voters of each possible type, and 1 extra voter of type 12345. The sincere winner is alternative 1.
- If 1 type 53124 voter votes instead as 35241, alternative 2 wins; if 2 switch, alternative 3 wins; if 3 switch, alternative 4 wins.
- Thus such voters can both undershoot and overshoot in the same profile.

Previous work

- Slinko and White showed that the analogue of the Gibbard-Satterthwaite theorem holds for safe manipulation. They asked about the probability that safe manipulation would succeed.

Previous work

- Slinko and White showed that the analogue of the Gibbard-Satterthwaite theorem holds for safe manipulation. They asked about the probability that safe manipulation would succeed.
- Hazon and Elkind studied the complexity of safe manipulation (COMSOC 2010, Tuesday). Their main relevant results:

Previous work

- Slinko and White showed that the analogue of the Gibbard-Satterthwaite theorem holds for safe manipulation. They asked about the probability that safe manipulation would succeed.
- Hazon and Elkind studied the complexity of safe manipulation (COMSOC 2010, Tuesday). Their main relevant results:
  - The results are strongly determined by the complexity of the tiebreaking algorithm.

Previous work

- Slinko and White showed that the analogue of the Gibbard-Satterthwaite theorem holds for safe manipulation. They asked about the probability that safe manipulation would succeed.
- Hazon and Elkind studied the complexity of safe manipulation (COMSOC 2010, Tuesday). Their main relevant results:
  - The results are strongly determined by the complexity of the tiebreaking algorithm.
  - (IsSafe) Given t, t', and an anonymous rule, it is decidable in polynomial time whether safe manipulation is possible.

Previous work

- Slinko and White showed that the analogue of the Gibbard-Satterthwaite theorem holds for safe manipulation. They asked about the probability that safe manipulation would succeed.
- Hazon and Elkind studied the complexity of safe manipulation (COMSOC 2010, Tuesday). Their main relevant results:
  - The results are strongly determined by the complexity of the tiebreaking algorithm.
  - (IsSafe) Given t, t', and an anonymous rule, it is decidable in polynomial time whether safe manipulation is possible.
  - (ExistsSafe) Given t, for a few common rules it is decidable in polynomial time whether safe manipulation is possible. Otherwise the answer is unknown.

Our goals for scoring rules

- Give efficient algorithms for solving the IsSafe and ExistsSafe problems.

Our goals for scoring rules

- Give efficient algorithms for solving the IsSafe and ExistsSafe problems.
- Characterize those situations that are safely manipulable.

Our goals for scoring rules

- Give efficient algorithms for solving the IsSafe and ExistsSafe problems.
- Characterize those situations that are safely manipulable.
- Compute the (exact limiting, as $n \to \infty$) probability that a voting situation is safely manipulable, under the uniform distribution (IAC). The limiting probability of a tie is zero, so we can ignore tiebreaking.

Our goals for scoring rules

- Give efficient algorithms for solving the IsSafe and ExistsSafe problems.
- Characterize those situations that are safely manipulable.
- Compute the (exact limiting, as $n \to \infty$) probability that a voting situation is safely manipulable, under the uniform distribution (IAC). The limiting probability of a tie is zero, so we can ignore tiebreaking.
- Let $S_{t,t'}$ denote the set of situations safely manipulable by switching from t to t'. We seek the size of the union

$$S := \bigcup_{t \in \mathcal{T}} S_t := \bigcup_{\substack{t \in \mathcal{T} \\ t \neq t' \in \mathcal{T}}} S_{t,t'}.$$

The University of Auckland
NEW ZEALAND

Basic observations for positional scoring rules

- Let a be the sincere winner. Call candidates preferred over a by t good and those ranked below a bad. Manipulation is safe iff bad candidate never wins for any value of x, good candidate wins for some x.

Basic observations for positional scoring rules

- Let a be the sincere winner. Call candidates preferred over a by t good and those ranked below a bad. Manipulation is safe iff bad candidate never wins for any value of x, good candidate wins for some x.
- Let $|c|(x)$ denote the score of c when x voters of type t switch to t'. This extends to real values of x in the obvious way. The graphs $x \mapsto |c|(x)$ are straight lines (the score lines).

Algorithm for positional scoring rules, I

- Fix t and t' and let $0 \leq x \leq |\mathcal{V}_t|$. Define

$$G(x) = \max\{|c|(x) \mid c \text{ is good }\}$$
$$B(x) = \max\{|c|(x) \mid c \text{ is bad }\}$$
$$U(x) = |c|(x), \text{where } c \text{ is the sincere winner.}$$

The University of Auckland
NEW ZEALAND

Algorithm for positional scoring rules, I

- Fix t and t' and let $0 \leq x \leq |\mathcal{V}_t|$. Define

$$G(x) = \max\{|c|(x) \mid c \text{ is good } \}$$
$$B(x) = \max\{|c|(x) \mid c \text{ is bad } \}$$
$$U(x) = |c|(x), \text{where } c \text{ is the sincere winner.}$$

- Compute the ordered list $I := \{i_1, i_2, \ldots, i_N := |\mathcal{V}_t|\}$ of intersections of the score lines. Initialize $k := 1$ and then loop through values of k:

Algorithm for positional scoring rules, I

- Fix t and t' and let $0 \leq x \leq |\mathcal{V}_t|$. Define

$$G(x) = \max\{|c|(x) \mid c \text{ is good }\}$$
$$B(x) = \max\{|c|(x) \mid c \text{ is bad }\}$$
$$U(x) = |c|(x), \text{where } c \text{ is the sincere winner.}$$

- Compute the ordered list $I := \{i_1, i_2, \ldots, i_N := |\mathcal{V}_t|\}$ of intersections of the score lines. Initialize $k := 1$ and then loop through values of k:
  - let $q_k := \lceil i_k \rceil$; if $q_k \geq i_{k+1}$ then go to start of loop;

Algorithm for positional scoring rules, I

- Fix t and t' and let $0 \leq x \leq |\mathcal{V}_t|$. Define

$$
\begin{aligned}
G(x) &= \max\{|c|(x) \mid c \text{ is good } \} \\
B(x) &= \max\{|c|(x) \mid c \text{ is bad } \} \\
U(x) &= |c|(x), \text{where } c \text{ is the sincere winner.}
\end{aligned}
$$

- Compute the ordered list $I := \{i_1, i_2, \ldots, i_N := |\mathcal{V}_t|\}$ of intersections of the score lines. Initialize $k := 1$ and then loop through values of k:
  - let $q_k := \lceil i_k \rceil$; if $q_k \geq i_{k+1}$ then go to start of loop;
  - check the inequalities: $B(q_k) > \max\{G(q_k), U(q_k)\}$ and $G(q_k) > U(q_k)$. If first inequality holds, return SAFE $=$ false; if second holds, return MANIP $=$ true.

The University of Auckland
NEW ZEALAND

Algorithm for positional scoring rules, I

- Fix t and t' and let $0 \leq x \leq |\mathcal{V}_t|$. Define

$$G(x) = \max\{|c|(x) \mid c \text{ is good }\}$$
$$B(x) = \max\{|c|(x) \mid c \text{ is bad }\}$$
$$U(x) = |c|(x), \text{where } c \text{ is the sincere winner.}$$

- Compute the ordered list $I := \{i_1, i_2, \ldots, i_N := |\mathcal{V}_t|\}$ of intersections of the score lines. Initialize $k := 1$ and then loop through values of k:
  - let $q_k := \lceil i_k \rceil$; if $q_k \geq i_{k+1}$ then go to start of loop;
  - check the inequalities: $B(q_k) > \max\{G(q_k), U(q_k)\}$ and $G(q_k) > U(q_k)$. If first inequality holds, return SAFE = false; if second holds, return MANIP = true.
- This determines whether a given situation belongs to $S_{t,t'}$

Algorithm for positional scoring rules, II

- I has size $O(m^2)$, and we simulate the voting rule once for each element of I. Each simulation requires $O(m)$ comparisons and m score updates each of which requires $O(1)$ arithmetic operations on numbers of size n.

Algorithm for positional scoring rules, II

- I has size $O(m^2)$, and we simulate the voting rule once for each element of I. Each simulation requires $O(m)$ comparisons and m score updates each of which requires $O(1)$ arithmetic operations on numbers of size n.
- The algorithm simplifies greatly when $m = 3$: safe manipulation is possible if and only if the maximal manipulation elects a good candidate.

The University of Auckland
NEW ZEALAND

Algorithm for positional scoring rules, II

- I has size $O(m^2)$, and we simulate the voting rule once for each element of I. Each simulation requires $O(m)$ comparisons and m score updates each of which requires $O(1)$ arithmetic operations on numbers of size n.
- The algorithm simplifies greatly when $m = 3$: safe manipulation is possible if and only if the maximal manipulation elects a good candidate.
- We now have a characterization of manipulable situations by linear (in)equalities.

Algorithm for positional scoring rules, III

- To compute S, we have some simplifications:

Algorithm for positional scoring rules, III

- To compute S, we have some simplifications:
  - S_t is empty if there are no good candidates (also if the top element of t has the lowest score, and the next element of t is the sincere winner);

Algorithm for positional scoring rules, III

- To compute S, we have some simplifications:
  - S_t is empty if there are no good candidates (also if the top element of t has the lowest score, and the next element of t is the sincere winner);
  - We need only consider t' for which all good candidates are ranked ahead of all bad ones and the sincere winner.

Algorithm for positional scoring rules, III

- To compute S, we have some simplifications:
  - S_t is empty if there are no good candidates (also if the top element of t has the lowest score, and the next element of t is the sincere winner);
  - We need only consider t' for which all good candidates are ranked ahead of all bad ones and the sincere winner.
  - We then use inclusion-exclusion.

Algorithm for positional scoring rules, III

- To compute S, we have some simplifications:
  - S_t is empty if there are no good candidates (also if the top element of t has the lowest score, and the next element of t is the sincere winner);
  - We need only consider t' for which all good candidates are ranked ahead of all bad ones and the sincere winner.
  - We then use inclusion-exclusion.
- This is very probably super-exponential in m, but polynomial in n.

IAC computations via polytopes

- The scores are all linear functions of the variables x_t, where x_t denotes the number of voters of type t.

IAC computations via polytopes

- The scores are all linear functions of the variables x_t, where x_t denotes the number of voters of type t.
- The inequalities above define a polytope $n\mathbb{P}$ with dimension $m!$, lying in the simplex $n\mathbb{S} := \{x \mid \sum_t x_t = n, \forall t\, x_t \geq 0\}$. The intersection of two S_t corresponds to the polytope with the union of constraints.

The University of Auckland
NEW ZEALAND

IAC computations via polytopes

- The scores are all linear functions of the variables x_t, where x_t denotes the number of voters of type t.
- The inequalities above define a polytope $n\mathbb{P}$ with dimension $m!$, lying in the simplex $n\mathbb{S} := \{x \mid \sum_t x_t = n, \forall t\, x_t \geq 0\}$. The intersection of two S_t corresponds to the polytope with the union of constraints.
- Under IAC, the probability distribution is uniform on $\mathbb{S}$, so probabilities reduce to counting lattice points in the polytope.

IAC computations via polytopes

- The scores are all linear functions of the variables x_t, where x_t denotes the number of voters of type t.
- The inequalities above define a polytope $n\mathbb{P}$ with dimension $m!$, lying in the simplex $n\mathbb{S} := \{x \mid \sum_t x_t = n, \forall t\, x_t \geq 0\}$. The intersection of two S_t corresponds to the polytope with the union of constraints.
- Under IAC, the probability distribution is uniform on $\mathbb{S}$, so probabilities reduce to counting lattice points in the polytope.
- The asymptotic leading term of the probability equals the volume of the normalized polytope $\mathbb{P}$ divided by that for $\mathbb{S}$. Such volumes can be computed by publicly available software implementing standard algorithms.

Linear system example: Borda, $m = 3$

- Suppose that the sincere election result is $|a| > |b| \geq |c|$, and we take $t = cba, t' = bca$. Order the types $abc, acb, bac, bca, cab, cba$ and let n_i be the size of $\mathcal{V}_i$.

Linear system example: Borda, $m = 3$

- Suppose that the sincere election result is $|a| > |b| \geq |c|$, and we take $t = cba, t' = bca$. Order the types $abc, acb, bac, bca, cab, cba$ and let n_i be the size of $\mathcal{V}_i$.
- Let $|a|'$ denote a's score after a strategic attempt as above, etc. Then the attempt is successful if and only if $|b|' \geq |a|', |c|'$. We can express $|a|'$, etc, as a linear combination of the n_i. This yields $n_i \geq 0$, $\sum_i n_i = n$, and

$$\begin{aligned}
0 &\leq n_1 + n_2 - n_3 - n_4 \\
0 &\leq n_3 + n_4 - n_5 - n_6 \\
0 &\leq -n_1 - n_2 + n_3 + n_4 + n_6 \\
0 &\leq -n_1 - n_2 + 2n_3 + 2n_4 - n_5 + 2n_2.
\end{aligned}$$

Numerical results for $m = 3$

Table: Asymptotic probability under IAC of a situation being (safely) manipulable.

scoring rule	P(manip)	P(safely)	P (safely | manip)
Plurality	0.292	0.292	1.00
(3,1,0)	0.422	0.322	0.76
Borda	0.502	0.347	0.69
(3,2,0)	0.535	0.330	0.62
(10,9,0)	0.533	0.264	0.49
Antiplurality	0.525	0.222	0.42

Discussion of results

- The ordering of rules according to their asymptotic susceptibility to manipulation is different when we restrict to safe manipulation.

Discussion of results

- The ordering of rules according to their asymptotic susceptibility to manipulation is different when we restrict to safe manipulation.
- The asymptotic conditional probability of being safely manipulable given manipulable decreases as the weight given to the second ranked alternative increases.

Extensions

- It seems natural to consider the uniform distribution on profiles (IC). However we don't expect this to be interesting for positional scoring rules, at least for large n. Reason: with high probability the differences in candidate scores are of order $\sqrt{n}$ but the number of voters of each type is of order n. Thus some types of votes will be safe almost always, other types almost never.

The University of Auckland
NEW ZEALAND

Extensions

- It seems natural to consider the uniform distribution on profiles (IC). However we don't expect this to be interesting for positional scoring rules, at least for large n. Reason: with high probability the differences in candidate scores are of order $\sqrt{n}$ but the number of voters of each type is of order n. Thus some types of votes will be safe almost always, other types almost never.
- Is there a polynomial time algorithm for ExistsSafe, for a general positional scoring rule? We know there is one for easy rules like plurality and antiplurality. What about Borda? (Recent: Egor Ianovski appears to have solved this).

Extensions

- It seems natural to consider the uniform distribution on profiles (IC). However we don't expect this to be interesting for positional scoring rules, at least for large n. Reason: with high probability the differences in candidate scores are of order $\sqrt{n}$ but the number of voters of each type is of order n. Thus some types of votes will be safe almost always, other types almost never.
- Is there a polynomial time algorithm for ExistsSafe, for a general positional scoring rule? We know there is one for easy rules like plurality and antiplurality. What about Borda? (Recent: Egor Ianovski appears to have solved this).
- What happens when we extend to coalitional manipulation, or some intermediate model?

Game interpretation

- Let $\mathcal{M}$ denote the potential manipulators in a voting situation. The voting rule defines a game form and the given profile an ordinal game with set of players $\mathcal{M}$. Assume that all players have complete information.

Game interpretation

- Let $\mathcal{M}$ denote the potential manipulators in a voting situation. The voting rule defines a game form and the given profile an ordinal game with set of players $\mathcal{M}$. Assume that all players have complete information.
- For ordinary manipulation, $\mathcal{M} = \mathcal{V}$. A profile is individually (coalitionally) manipulable if and only if it is not a Nash (strong Nash) equilibrium.

Game interpretation

- Let $\mathcal{M}$ denote the potential manipulators in a voting situation. The voting rule defines a game form and the given profile an ordinal game with set of players $\mathcal{M}$. Assume that all players have complete information.
- For ordinary manipulation, $\mathcal{M} = \mathcal{V}$. A profile is individually (coalitionally) manipulable if and only if it is not a Nash (strong Nash) equilibrium.
- For safe manipulation, $\mathcal{M} = \mathcal{V}_t$ for some fixed t. Suppose that t and t' are specified. The players in $\mathcal{M}$ have a unique dominant strategy in a given profile ("all switch to t'") if and only if the profile is safely manipulable.

Game interpretation

- Let $\mathcal{M}$ denote the potential manipulators in a voting situation. The voting rule defines a game form and the given profile an ordinal game with set of players $\mathcal{M}$. Assume that all players have complete information.
- For ordinary manipulation, $\mathcal{M} = \mathcal{V}$. A profile is individually (coalitionally) manipulable if and only if it is not a Nash (strong Nash) equilibrium.
- For safe manipulation, $\mathcal{M} = \mathcal{V}_t$ for some fixed t. Suppose that t and t' are specified. The players in $\mathcal{M}$ have a unique dominant strategy in a given profile ("all switch to t'") if and only if the profile is safely manipulable.
- What happens in other cases? What do symmetric (mixed) Nash equilibria look like? What if we only want safety with high probability?