Strategic Voting: Safely Manipulating under Uncertainty

Introduction

Our voting system is imperfect. This is widely appreciated in today's day and age, but it is not anthropic corruption alone that skews the purity of electing a leader. Voting methods (the ways we choose our leaders) are inherently susceptible to manipulation, provided they are constrained by a few reasonable assumptions. The Gibbard-Satterthwaite theorem [2] tells us that *any* reasonable voting method is susceptible to *strategic voting*. Reasonable meaning that the method is: 1) resolute (i.e., it selects a single winner), 2) unanimous (i.e., if all voters choose Bernie Sanders over Trump, then Trump cannot win) and finally 3) nondictorial (i.e., there is no single voter that solely dictates the outcome of the vote). Needless to say, these are restrictions that any fair election should adhere to. To vote strategically is to manipulate an election by submitting insincere preferences. This means that voters are not guaranteed their best result by being honest. What a novel thought: situations that decide power can, and often do, favor dishonesty. Even in a formal context, you cannot escape the true nature of politics.

Literature Review

Existence of manipulation is not the whole story; there is a large body of work that addresses the computational complexity of finding manipulations. In most cases, the manipulator is omniscient (i.e., is given full knowledge of everyone's preferences and the method for selecting a winner). The complexity of these situations, as well as algorithms for determining specific manipulations are well understood in many situations ([5],[4],[11],[10],[1]). However, an omniscient voter is far from realistic. It is natural to assume that a manipulator is restricted by a level of uncertainty. In other words, the manipulator's knowledge is consistent with multiple possible systems. For formal analysis, it is customary to assume that the manipulator is sure of his own preferences. Thus, the uncertainty must stem from a lack of knowledge about either the voting rule or the preferences of the other voters. The former has recently been analyzed by Holliday and Pacuit in [8]. They analyze three different manipulation criterion: sure, safe, and expected manipulations. These criteria capture different levels of confidence for the strategic voter. A *sure* manipulation is one where you are certain that submitting an insincere preference

will produce an improved result, while a *safe* manipulation is one where you are only certain that your manipulation won't lead to a worse result. An *expected* manipulation occurs when your manipulation is more likely to be beneficial than not. Although they produced results in all three regimes, the most applicable to this work are their results on safe manipulation. Providing multiple possible voting methods is not sufficient to eliminate safe manipulation; however, it can reduce the incentive to manipulate. This does show that introducing even a minimal uncertainty (2 possible voting methods) can be used as a tool to incentivize honest voting. Although this work will focus on situations where you restrict the manipulator's knowledge about the preferences of other voters, there seems to be a definite motivation for future work that melds these two regimes.

One of the first investigations of strategic voting under uncertainty of voters' preferences was carried out by Conitzer, Walsh and Xia [3]. Here they develop a framework for interpreting the uncertainty as a set of possible profiles that extend your current state of knowledge. This leads to a natural epistemic interpretation of these situations as developed in [6]. Conitzer et al. were also interested in the conditions that would eliminate all safe manipulation. From a perspective of civility, their work motivates the search for a minimal set of information that will guarantee the elimination of safe manipulation. From the manipulator's perspective, it begs the question of how much information one needs to ensure that a potentially advantageous manipulation remains safe. They first show that in the extreme case that the manipulator has no information about other voter's preferences, most voting methods (with a sufficient ratio of voters to candidates) will be immune to a dominating manipulation.[1] Especially in the case of positional scoring methods, this result suggests that there is a critical point for each voting method at which the uncertainty set will eliminate safe manipulation. Conitzer et al. show that for the case of the Borda method, which will be defined formally in the next section, that slightly restricting the manipulator's information forces the computation for finding a safe manipulation to be an NP-hard problem. Naturally, these results motivate the study of specific uncertainty sets that are prevalent in everyday life. This investigation was developed by Reijngoud and Endriss in [9] and furthered by Endriss et al. in [7]. The latter analyzes 3 different types of uncertainty sets; the first assumes you know nothing aside from who wins the election. The second, a more informative set, considers that you know the score for each candidate. The last uncertainty set assumes the manipulator knows if the majority of the voting population prefers candidate A over B for each pairwise match up of candidates. The following sections will develop the notion of strategic voting under uncertainty, with an identical goal in mind: characterizing particular uncertainty sets. The notation and terminology used will closely resemble that of [8] and will focus in on the voting methods of Borda, Plurality and k-approval.

Preliminaries

Let V be a nonempty, finite set of n *voters* denoted by $\{1, \cdots, n\}$. Let C be the set of m *candidates* $\{c_1, \cdots, c_m\}$. A ballot then is a full set of preferences submitted by a given

[1] for n voters and m candidates you need $n \geq 6(m-2)$ to eliminate safe manipulations in positional scoring rules [3]

voter; this is nothing more than a linear order on C. Thus, consider the set $\mathcal{L}(C)$ which will denote the set of a all linear orders on C. For $i \in V$, we let $P_i \in \mathcal{L}(C)$ represent the truthful preferences of the i^{th} voter. Thus P_i allows you to compare the ranking of two candidates according to voter i; if i prefers c_j to $c_{j'}$ then $c_j P_i c_{j'}$. This total collection of binary rankings is often referred to as i's ballot. This paper will only consider a single manipulator, $1 \in V$, for simplicity. This will be the only voter submitting an insincere ballot which will be denoted by P_1^*. By considering the preferences of each candidate, you can define a *profile* for the population of voters, given by $\mathbb{P} = \{P_i : i \in V\}$. Throughout this work $\mathbb{P}$ will denote the generic *truthful profile* populated by genuine preferences. A *manipulation* will be represented as a function on a truthful profile that replaces P_1 with P_1^* and applies the identity to $i \in V - \{1\}$. The *manipulated profile* is the image of this function, denoted as $\mathbb{P}_{[P_i \to P_i^*]}$. The outcome of an election is determined by a *voting method*; this will be a function $f : \mathcal{L}(C)^n \to C$ that picks out a single winner.

The voting methods of interest in this paper fall under what are commonly referred to as positional scoring rules. These assign a numerical score to each candidate based on their rank in each P_i; the candidate with the maximum score is selected as the winner. For purposes of simplicity assume that scoring ties are broken lexicographically, so in favor of the candidate with the lowest index as done in [9]. Define a *scoring vector* to be $\langle s_1, s_2, \cdots s_m \rangle$ where for $j = 1, \cdots, m-1$, $s_j \geq s_{j+1}$. Then the the score of $x \in C$ for profile $\mathbb{P}$ is given by $\text{score}(\mathbb{P}, x) = \Sigma_{i=1}^{n} \text{score}(P_i, x)$, where $\text{score}(P_i, x) = s_r$ and r is the position that x occurs in the linear order P_i. These voting methods can be defined as follows: $\forall \mathbb{P} \in \mathcal{L}(C)^n, f(\mathbb{P}) = \max_{x \in C} \text{score}(\mathbb{P}, x)$. This assumes that this maximum intrinsically accounts for the lexicographic tie breaking rule. Thus you can fully generate a positional scoring rule, given their scoring vector. This yields the following definitions:

- *Plurality*: the rule given by $\langle 1, 0 \cdots, 0 \rangle$;
- *k-approval*: the rule given by $\langle 1, 1 \cdots, 0 \rangle$, where you have k 1's followed by zeros;
- *Borda*: the rule given by $\langle m-1, m-2 \cdots, 0 \rangle$.

Uncertainty Sets and Safe Manipulation

As noted prior, this paper aims to investigate profiles in which we know that a manipulation exists. This allows one to beg the question of how much information the manipulator needs to ensure that their manipulation is safe. *Safe*, in this context, means that insincere preferences will fair just as well, or better, than truthful preferences. It has be shown that omniscience is not always needed to preserve the safety of a manipulation [9]. Thus it is natural to restrict the information available. This is done by defining an uncertainty set as follows: let an *uncertainty set* be a subset $U \subseteq \mathcal{L}(C)^n$ that collects the profiles that agree with the information provided to the manipulator.[2] For example, consider the set generated by $\mathbb{P}$: $\{\mathbb{P}' : P_1 \simeq P_1'\}$. This is an uncertainty set for a manipulator who only

[2]This method of defining an uncertainty set by extending an incomplete knowledge state into a set of possible election scenarios was pioneered by Conitzer in [3]

knows their own ballot; the lack of information is captured by the collection of possible profiles $\mathbb{P}'$ that extend P_1. This will serve as the largest uncertainty set as we will always assume that the manipulator is certain/aware of their true preferences. The most interesting cases of uncertainty for positional scoring rules occur when the manipulator has at least the knowledge of their own preferences and the outcome of the election. This can be generated by an *uncertainty function*, U, such that $U(\mathbb{P}) = \{\mathbb{P}' : f(\mathbb{P}) = f(\mathbb{P}') \text{ and } \mathbb{P}'_1 \simeq \mathbb{P}_1\}$. The image of U allows the manipulator to compare the genuine winner to other candidates that might be selected. This gives a concrete way to assess whether or not an advantageous manipulation remains safe when considered in all profiles of an uncertainty set.

The goal of this paper is to analyze whether or not some realistic examples of uncertainty sets eliminate *safe manipulation*. Let $\preceq$ be the weak order given by the manipulator's true preferences P_1 and the identity. And $\prec$ be the strict order determined by P_1. Let the existence of a safe manipulation by a single manipulator be defined as a property of a voting method f and an uncertainty set $U(\mathbb{P})$. Then to say that $\langle f, U(\mathbb{P})\rangle$ is susceptible to a safe manipulation means that:

$$\exists P_1^* \in \mathcal{L}(C) \text{ such that } \forall \mathbb{P}' \in U'(\mathbb{P}) :$$
$$f(\mathbb{P}) \preceq f(\mathbb{P}'_{[P_1 \to P_1^*]}) \text{ and } \exists \mathbb{P}^* \text{ such that } f(\mathbb{P}) \prec f(\mathbb{P}^*_{[P_1 \to P_1^*]}).$$

This investigation is interested in situations where safe manipulation breaks down. A necessary prerequisite is a profile $\mathbb{P}$ that does in fact witness manipulation by $1 \in V$ via P_1^*. This is equivalent to saying that $\langle f, \mathbb{P}\rangle$ is susceptible to safe manipulation, where the manipulator has full information about the election. However, by considering the multitude of profiles in an uncertainty set, it is possible that P_1^* violates safety. Safety of P_1^* is broken for uncertainty set $U(\mathbb{P})$ when:

$$\exists \mathbb{P}' \in U(\mathbb{P}) : f(\mathbb{P}_{[P_i \to P_1^*]}) \prec f(\mathbb{P}) \tag{1}$$

This means there is a profile in your uncertainty set that returns a strictly less preferable candidate. We say that a manipulation P_1^* of $\mathbb{P}$ is *not safe* for $\langle f, U(\mathbb{P})\rangle$ if (1) holds.

Hopefully this will motivate the search for the following set:

$$U_{\min}(\mathbb{P}) = \bigcap \{U(\mathbb{P}) : \forall P_1^*, P_1^* \text{ is } \textit{not safe} \text{ for } \langle f, U(\mathbb{P})\rangle\}$$

This represents the maximal amount of knowledge a manipulator can have while still rendering each possible manipulation unsafe. Ensuring this knowledge threshold for each voter would reserve insincere ballots for gamblers.[3]

Results

Consider the following uncertainty sets, generated by a true profile $\mathbb{P}$ that witnesses at least one manipulation by voter $1 \in V$:

[3]Those who want to take a chance on a manipulation cannot be stopped; however, in this scenario honesty becomes the safe option which seems ideal.

$U_{\text{Omni}}(\mathbb{P}) = \{\mathbb{P}\}$
$U(\mathbb{P}) = \{\mathbb{P}' : f(\mathbb{P}) = f(\mathbb{P}_i) \text{ and } \mathbb{P}'_1 \simeq \mathbb{P}_1\}$
$U'(\mathbb{P}) = U \cap \{\mathbb{P}' : \forall i \in V, \max(P_i) = \max(P'_i)\}$
$U''(\mathbb{P}) = U \cap \{\mathbb{P}' : \mathbb{P}' = \mathbb{P}_{[P_a \to P'_a]}, P'_a \in \mathcal{L}(C)\}$
$U_{\text{Ig}}(\mathbb{P}) = \{\mathbb{P}' : \mathbb{P}'_1 \simeq \mathbb{P}_1\}$

Let these sets be referred to as the "omniscient", "winning", "first choice", "missing person", and "ignorant" priors, respectively. Arguably, these representations of initial information states are viable in the real world. The most extreme cases are captured by the ignorant and omniscient priors. The next two corollaries characterize safe manipulation for these cases.[4]

Corollary 1. *For $m \geq 3$, there exists a manipulation P_1^* that is safe for $\langle f, U_{Omni}(\mathbb{P})\rangle$ where f is any of the following methods: Borda, k-approval, or Plurality.*

Proof. See Theorem 1 of [9] for details. ■

Corollary 2. *When $n \geq 2m - 2$, there are no safe manipulations for $\langle f, U_{Ig}(\mathbb{P})\rangle$ where f is any of the following methods: Borda, k-approval, or Plurality.*

Proof. See Theorem 5 of [9] for details. ■

These results indicate that for the three voting methods under investigation, there must be some critical prior that marks the change from *safe* to *unsafe* for any given manipulation. The winning prior has also been well studied, yielding the following result:

Corollary 3. *For $m \geq 3$ and $n \geq 4$ there exists a manipulation P_1^* that is safe for $\langle f, (\mathbb{P})\rangle$ where f is any of the following methods: Borda or Plurality.*

Proof. See Theorem 3 of [9] for details. ■

Corollary 4. *If $m \geq 3$,$n \geq 4$, and $k \leq m - 2$, then there exists a manipulation P_1^* that is safe for $\langle f, U(\mathbb{P})\rangle$ where f is k-approval.*

Proof. See Theorem 2 of [7] for details. ■

Lemma 1. *For an uncertainty function $\tilde{U}$ such that $\tilde{U}(\mathbb{P}) \subseteq U(\mathbb{P})$, if there exists a manipulation P_1^* that is safe for $\langle f, U(\mathbb{P})\rangle$ then P_1^* is also safe for $\langle f, \tilde{U}(\mathbb{P})\rangle$.*

Proof. This reduces to showing that 1) $\forall \mathbb{P}' \in U'(\mathbb{P}) : f(\mathbb{P}) \preceq f(\mathbb{P}'_{[P_1 \to P_1^*]})$ and 2) $\exists \mathbb{P}'' \in \tilde{U}(\mathbb{P})$ such that $f(\mathbb{P}) \prec f(\mathbb{P}''_{[P_1 \to P_1^*]})$. Since $\tilde{U}(\mathbb{P}) \subseteq U(\mathbb{P})$, $\mathbb{P}' \in \tilde{U}(\mathbb{P}) \implies \mathbb{P}' \in U(\mathbb{P})$. Since P_1^* is safe for $\langle f, U(\mathbb{P})\rangle$, 1) follows immediately. Furthermore, the true profile must be consistent with any uncertainty set. Thus $\mathbb{P} \in \tilde{U}(\mathbb{P})$, and by assumption the true profile $\mathbb{P}$ witnesses a manipulation by $1 \in V$. Thus $f(\mathbb{P}) \prec f(\mathbb{P}_{[P_1 \to P_1^*]})$. This satisfies 2) and completes the proof that P_1^* must also be safe for $\langle f, \tilde{U}(\mathbb{P})\rangle$. ■

Theorem 1. *When $m \geq 3$ and $n \geq 4$, there exists a manipulation P_1^* that is safe for $\langle f, U(\mathbb{P})\rangle$ and $\langle f, U''(\mathbb{P})\rangle$, where f is Borda or Plurality.*

[4]These are corollaries derived from the results of Reijgnoud and Endriss.

Proof. Given that $m \geq 3$ and $n \geq 4$, Corollary 3 implies that $\exists P_1^* \in \mathcal{L}(C)$ that is safe for $\langle f, U(\mathbb{P})\rangle$. Given that $U'(\mathbb{P}) \subseteq U(\mathbb{P})$ and $U''(\mathbb{P}) \subseteq U(\mathbb{P})$, Lemma 1 implies that P_1^* is safe for both $\langle f, U(\mathbb{P})\rangle$ and $\langle f, U''(\mathbb{P})\rangle$. ■

Theorem 2. *If $m \geq 3$,$n \geq 4$, and $k \leq m-2$, then there exists a manipulation P_1^* that is safe for $\langle f, U(\mathbb{P})\rangle$ and $\langle f, U''(\mathbb{P})\rangle$ where f is k-approval.*

Proof. Fix f as the k-approval voting method. Given the restrictions on (n, m, k), Corollary 4 implies that $\exists P_1^* \in \mathcal{L}(C)$ that is safe for $\langle f, U(\mathbb{P})\rangle$. Given that $U'(\mathbb{P}) \subseteq U(\mathbb{P})$ and $U''(\mathbb{P}) \subseteq U(\mathbb{P})$, Lemma 1 implies that P_1^* is safe for both $\langle f, U(\mathbb{P})\rangle$ and $\langle f, U''(\mathbb{P})\rangle$. ■

Conclusion

These results show that both the "first choice" and the "missing person" uncertainty sets are in fact susceptible to safe manipulation. This process can be applied to a number of other intuitive uncertainty sets that extend the "winning" prior. This could be very valuable information for both manipulators and committees looking to eliminate safe manipulation. As shown in [3], actually finding this safe manipulation is an NP-hard problem. Most manipulator's will not be able to solve this sufficiently fast to make the right decision. However, is there another way to gather useful information from this system from the standpoint of the manipulator? In larger elections, with many more than 3 candidates and 4 voters, it is possible to come up with a number of different manipulations. Thus it seems natural to inquire about which of these manipulations are safe under particular uncertainty sets. Furthermore, it has been shown that the elimination of safe manipulation is possible for any voting method if you restrict the manipulator's knowledge enough. Thus, it stands to reason (by an adaptation of the Intermediate Value Theorem) that there is a minimum uncertainty set that will eliminate safe manipulation. As seen from this work in conjunction with [8], there is a definite motivation for combining uncertainty about voters preferences and the voting method to find this set for any election scenario.

References

[1] J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989.

[2] Jean-pierre Beno. The Gibbard–Satterthwaite Theorem: a Simple Proof. *Economics Letters*, 69:319–322, 2000.

[3] Vincent Conitzer, Toby Walsh, and Lirong Xia. Dominating manipulations in voting with partial information. *Proceedings of the National Conference on Artificial Intelligence*, 1:638–643, 2011.

[4] Palash Dey. Resolving the Complexity of Some Fundamental Problems in Computational Social Choice. 012(August), 2016.

[5] Palash Dey, Neeldhara Misra, and Y. Narahari. Complexity of manipulation with partial information in voting. *Theoretical Computer Science*, 726:78–99, 2018.

[6] Hans Van Ditmarsch. Strategic voting and the logic of knowledge. (Extended Abstract). *Aamas*, pages 196–205, 2012.

[7] Ulle Endriss, Svetlana Obraztsova, Maria Polukarov, and Jeffrey S. Rosenschein. Strategic voting with incomplete information. *IJCAI International Joint Conference on Artificial Intelligence*, 2016-Janua:236–242, 2016.

[8] Wesley H. Holliday and Eric Pacuit. Strategic voting under uncertainty about the voting method. *Electronic Proceedings in Theoretical Computer Science, EPTCS*, 297:252–272, 2019.

[9] Annemieke Reijngoud and Ulle Endriss. Voter response to iterated poll information. *11th International Conference on Autonomous Agents and Multiagent Systems 2012, AAMAS 2012: Innovative Applications Track*, 1(Illc):376–383, 2012.

[10] Reyhaneh Reyhani. Strategic Manipulation in Voting Systems. (January 2013), 2013.

[11] Lirong (Duke University) Xia. Computational Voting Theory : Computational Voting Theory: Game-Theoretic and Combinatorial Aspects. pages 1–316, 2011.