Complexity of Distances in Elections

Doctoral Consortium

Tobias Hogrebe
Heinrich-Heine-Universität Düsseldorf
Düsseldorf, Germany
hogrebe@cs.uni-duesseldorf.de

ABSTRACT

Our research investigates the influence of distances on elections and the associated problems such as bribery, manipulation, and control. Distances can be used in elections in many different ways. In the case of interference problems, distances can be used to limit the changes or to calculate costs. At the same time, distances can themselves be used to design fair elections. Our goals here are to define appropriate frameworks and to determine the complexity of the problems with respect to approximation and parameterization.

KEYWORDS

Elections; Distances; Scoring Systems; Bribery; Manipulation

ACM Reference Format:

Tobias Hogrebe. 2019. Complexity of Distances in Elections. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019*, IFAAMAS, 3 pages.

1 INTRODUCTION

The research presented in the following mainly focuses on the area of *computational social choice* (COMSOC). Social choice is an interdisciplinary field of sociology, economy, mathematics and many more that deals with elections, their design and properties. Computational social choice deals in particular with the complexity of the evaluation of elections and related problems. Fundamental to this field were the following two publications. In 1973, Gibbard [10] showed that no rational election rule can be resistant to manipulation. In 1989, Bartholdi, Tovey, and Trick [1] started the investigation of election problems with regard to their complexity studying the complexity of the manipulation problem.

In order to analyze elections from a mathematical point of view, we first introduce the formal foundations of elections. An election is given by a pair (C, V) consisting of a set of *candidates* C and a list of votes $V = (v_1, \ldots, v_n)$ over C, one for each *voter* $i \in \{1, \ldots, n\}$. The list of votes is also called the *profile*. Here we focus on the well-established representation of the votes as complete *linear orders* over the candidates. This means that each voter presents a complete and strict ranking of the candidates from position 1, most preferred, to position m, least preferred. The set of all complete linear orders over C is denoted by $\mathcal{L}(C)$. Finally, an *election rule* maps an election (C, V) to a subset of C, namely the winners of the election. The most prominent election rules are the scoring rules. A *scoring rule* for m candidates is defined by a scoring vector $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{R}^m_{\geq 0}$ with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_m$, where α_j denotes the number of points a candidate receives for being placed on position j by one of the voters. Those candidates with the maximum number of points are the winners of the election. A scoring rule for a variable number of candidates is given by an efficiently evaluable function determining a scoring vector for each number of candidates. A natural subclass of scoring rules are the *pure scoring rules*, where the scoring vector for $m + 1$ candidates can be generated by inserting an appropriate value into the scoring vector for m candidates. Prominent members of this class are for example Borda with $(m - 1, m - 2, \ldots, 0)$ and Plurality with $(1, 0, \ldots, 0)$.

We study elections in particular with respect to distances. A *distance* d on a space A is a mapping $d : A \times A \to \mathbb{R}$ that fulfills the following properties for all $a, b, c \in A$: (i) $d(a, b) \geq 0$ (non-negativity), (ii) $d(a, b) = 0$ if and only if $a = b$ (identity of indiscernibles), (iii) $d(a, b) = d(b, a)$ (symmetry), and (iv) $d(a, b) + d(b, c) \geq d(a, c)$ (triangle inequality). We mainly focus on distances between vectors and distances between votes. For a fixed set of candidates the latter one is a distance of the form $d : \mathcal{L}(C) \times \mathcal{L}(C) \to \mathbb{R}_{\geq 0}$. To study the complexity of problems involving such distances we consider families of distances that contain one distance function for each possible number of candidates. The most prominent distances on rankings are the swap distance introduced by Sir Maurice George Kendall [12] in 1938 and the footrule distance introduced by Charles Spearman [16] in 1906. Given two votes v and v' from $\mathcal{L}(C)$, the *swap distance* $\mathrm{swap}(v, v')$ is defined as $\mathrm{sw}(v, v') = |\{(x, y) \in C \times C \mid x >_v y \text{ and } y >_{v'} x\}|$. Instead of counting swaps, the footrule distance counts the positions that a candidate needs to be shifted by to obtain the target ranking. The *footrule distance* $\mathrm{fr}(v, v')$ is defined as: $\mathrm{fr}(v, v') = \sum_{y \in C} |\mathrm{pos}(v, y) - \mathrm{pos}(v', y)|$ with $\mathrm{pos}(v, c)$ denoting the position of candidate c in vote v. Both distances can be extended by a weight function which assigns a weight to each pair of candidates.

Combining distances and elections is by no means a novel approach. Charles Dodgson [7] already implicitly used the unweighted swap distance in the definition of his election rule in 1876. In 1959, the unweighted swap distance has been used by John Kemeny [11] in the definition of the Kemeny method. In 2009, Elkind, Faliszewski, and Slinko [8] used distances to characterize election rules. In addition to the design of election rules, distances can also be used in the definition of interference problems like bribery. The *bribery* problem asks whether an agent can change the outcome of an election given a specific budget if the cost of the bribery is determined by a certain cost function. In the case of swap bribery, introduced by Elkind, Faliszewski, and Slinko [9] in 2009, the cost function is given by the weighted swap distance. In 2016, Yang, Shrestha and Guo [17] examined the influence of unweighted distances on the bribery problem, looking at a more local variation where a limited

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada. © 2019 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

number of votes is allowed to be changed within a fixed uniform distance limit.

Among other things, we study the *distance bribery* introduced by Baumeister, Hogrebe, and Rey [3]. The problem generalizes the swap bribery by allowing the cost of the bribery to be determined by arbitrary distances on linear orders and different types of weightings. In the constructive case, we are interested in making a certain candidate the winner of the election. In the destructive case, we try to prevent a certain candidate from winning. The distances we consider are the weighted swap and footrule distance, including the element-weighted variants inspired by the work of Kumar and Vassilvitskii [13] in which the pairwise weights are the product of candidate weights, and the prominent unweighted variants. The swap and footrule distance, despite their diverse nature, are known to be highly related. This connection has been studied for example by Diaconis [6] and Kumar and Vassilvitskii [13]. We are interested in finding this type of connections for the variants we used.

The concept of distance bribery is representative for many other motivations: The robustness of the outcome of an election considered by Shiryaev, Yu, and Elkind [15], where the robustness was measured using the unweighted swap distance, whereby the basic problem was the destructive swap bribery. The optimal manipulation problem introduced by Obraztsova and Elkind [14], is a special case of the distance bribery problem, where a single voter is interested into manipulating the election by changing her vote while minimizing the distance to her true ranking to keep the manipulation as unobtrusive as possible.

2 RESULTS

Merging swap bribery, optimal manipulation, and robustness of elections in the framework of constructive and destructive distance bribery revealed many question marks in the complexity of the respective problems. For example, the complexity of the standard swap bribery has never been investigated before in the destructive case. For this problem we establish the complexity for a large number of pure scoring rules including the well-known ones: For highly differentiating scoring rules such as Borda, the problem is intractable, while it is easy for more static scoring rules like k-approval. Note that these differences in complexity are very different from the complexity of the constructive variant, which is intractable for all pure scoring rules except plurality and veto. Interestingly, the complexity changes for some cases when switching from swap to footrule distance. For the scoring rule characterized by $(2, 1, \ldots, 1, 0)$, we show that the problem is easy for the swap distance, while it is intractable for the footrule distance. For the constructive case we present a dichotomy result for the element-weighted distances and pure scoring rules, which shows that the complexity of the element-weighted variant matches that of the fully weighted variant, which was determined by Elkind, Faliszewski, and Slinko [9] based on the results of Betzler and Dorn [5] and Baumeister and Rothe [4]. In addition, we show that in the constructive case the problems are also intractable for the unweighted distances and certain pure scoring rules.

Combining all these results, one can derive basic insights about the complexity of the problems in terms of pure scoring rules. For example, it follows that in the destructive case, hardness usually arises from the change of only one vote. This leads to the fact that both the P and the NP-hardness results can be transferred to the optimal manipulation problem. However, for many rules, for which the destructive variant is easy, such as 2-approval, the hardness of the constructive bribery can be shown through reductions involving the corruption of multiple votes.

3 FUTURE WORK

As for distance bribery, our examination revealed many future research possibilities. In addition to the cases that are still open, it would be interesting to see generally what features of distances and election rules are crucial to the complexity of the problem in the constructive and destructive case. A further step would be the investigation of the complexity regarding approximation and parameterization, for example, with respect to the weight functions. The above insight regarding the number of bribed voters needed to show hardness also indicates that parameterization by the impact of the bribery, like the number of bribed voters, is a deciding factor that should be further investigated.

Apart from bribery, we are interested in the importance of distances for committee election rules. Distances can be used to determine how satisfied a voter is with the outcome of the election, the committee. In this case, it is particularly interesting that distances can not only be used to evaluate the quality of already existing election rules but can also be implemented in different ways as election rules themselves. The combination of distances and implementations decide whether the resulting rules fulfill properties such as the justified representation.

Returning to the election problems, distances on profiles rather than single votes offer the possibility to unify the various election problems such as manipulation, possible winner, bribery, and control in a framework in which each of the problems is represented by specific distances. This allows, for example, combinations of the problems to be considered.

An interesting problem that has been studied only marginally in the literature is the manipulative design of election rules (see Baumeister and Hogrebe [2]). The problem is to check whether a set of election rules, such as the scoring rules, contains one election rule that will guarantee a specific outcome. The selected election rule must guarantee the outcome for a list of profiles that are most likely to apply according to predictions or maximize chances for the outcome for a given distribution of profiles. Especially for scoring rules, the problem is very interesting, as there are many different natural restrictions that the scoring vectors should meet to be accepted. If someone wants to change an existing scoring rule, it would also be reasonable to limit the distance between the original and the new system. We assume that the various restrictions and possibilities of parameterization lead to a diverse complexity.

In the future, we also want to further explore the connection between election rules and the determination of winners in sports. This connection seems to be enormously important for the influence of manipulation, bribery, and design in sports.

ACKNOWLEDGMENTS

This work is supported by the DFG-grant BA6270/1-1.

REFERENCES

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

[2] D. Baumeister and T. Hogrebe. 2019. Manipulative Design of Scoring Systems (Extended Abstract). In *Proceedings of the 18th Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, To appear.

[3] D. Baumeister, T. Hogrebe, and L. Rey. 2019. Generalized Distance Bribery. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. To appear.

[4] D. Baumeister and J. Rothe. 2012. Taking the final step to a full dichotomy of the possible winner problem in pure scoring rules. *Information Processing Letters* 112, 5 (2012), 186–190.

[5] N. Betzler and B. Dorn. 2010. Towards a Dichotomy of Finding Possible Winners in Elections Based on Scoring Rules. *Journal of Computer and System Sciences* 76, 8 (2010), 812–836.

[6] P. Diaconis and R. Graham. 1977. Spearman's footrule as a measure of disarray. *Journal of the Royal Statistical Society. Series B (Methodological)* (1977), 262–268.

[7] C. Dodgson. 1876. A method of taking votes on more than two issues. *The Theory of Committees and Elections* (1876).

[8] E. Elkind, P. Faliszewski, and A. Slinko. 2009. On distance rationalizability of some voting rules. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*. ACM, 108–117.

[9] E. Elkind, P. Faliszewski, and A. Slinko. 2009. Swap Bribery. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory*. Springer-Verlag *Lecture Notes in Computer Science #5814*, 299–310.

[10] A. Gibbard. 1973. Manipulation of voting schemes: a general result. *Econometrica: Journal of the Econometric Society* (1973), 587–601.

[11] J. Kemeny. 1959. Mathematics without numbers. *Daedalus* 88, 4 (1959), 577–591.

[12] M. Kendall. 1938. A new measure of rank correlation. *Biometrika* 30, 1/2 (1938), 81–93.

[13] R. Kumar and S. Vassilvitskii. 2010. Generalized distances between rankings. In *Proceedings of the 19th International Conference on World Wide Web*. ACM Press, 571–580.

[14] S. Obraztsova and E. Elkind. 2012. Optimal Manipulation of Voting Rules. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 619–626.

[15] D. Shiryaev, L. Yu, and E. Elkind. 2013. On elections with robust winners. In *Proceedings of the 13th International Joint Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 415–422.

[16] C. Spearman. 1906. 'Footrule' for measuring correlation. *British Journal of Psychology* 2, 1 (1906), 89–108.

[17] Y. Yang, Y. Shrestha, and J. Guo. 2016. How Hard Is Bribery with Distance Restrictions. In *Proceedings of the 22nd European Conference on Artificial Intelligence*. IOS Press, 363–371.