On the Complexity of Candidates-Embedded Multiwinner Voting under the Hausdorff Function

Yongjie Yang
Chair of Economic Theory, Saarland University
Saarbrücken, Germany
yyongjiecs@gmail.com

ABSTRACT

We study candidates-embedded approval-based multiwinner voting. In this model, we are given a metric f on the set of candidates, and voters are free to approve or disapprove any candidates. The task is to select a k-committee that either minimizes the sum of distances from the committee to all votes (utilitarian rules) or minimizes the maximum distance from the committee to any vote (egalitarian rules). The distance from a committee to a vote is measured by certain set-to-set functions derived from f. Previous works have considered the min, the max, and the sum functions. This paper examines the Hausdorff function. We show that in general computing winners under the Hausdorff function is hard, but we also derive several polynomial-time algorithms for certain special cases.

KEYWORDS

multiwinner voting; Hausdorff distance; NP-hardness, parameterized complexity

ACM Reference Format:
Yongjie Yang. 2024. On the Complexity of Candidates-Embedded Multiwinner Voting under the Hausdorff Function. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

Approval-based multiwinner voting (ABMV) has experienced a resurgence in popularity over the past decade due to its wide range of applications across various AI subareas [3, 7, 13, 22, 26], despite its historical roots dating back more than two centuries [9, 21, 24]. Canonical ABMV assumes candidates to be indistinguishable. However, in numerous applications, candidates exhibit correlations, inspiring researchers to explore more versatile ABMV models that account for various types of relations among candidates (see, e.g., [16, 27, 28]). Particularly, Yang [27] proposed a model wherein candidates possess distance relationships. This model finds relevance in scenarios such as collective selection of travel destinations or vaccination locations. Depending on the specific scenario, the distance between two candidates can represent physical distance, similarity degree, communication cost, and more. In this model, Yang [27] studied rules which select a winning k-committee w minimizing either the sum of the distance from w to all votes (*utilitarian rules*), or the maximum distance from w to every vote (*egalitarian rules*). These distances are defined by certain set (set-to-set) functions derived from a metric over candidates. Yang [27] studied the min function, the max function, and the sum function, obtaining comprehensive complexity results for the winner determination problems in both the general cases and special cases where the candidates are embedded in discrete metrics, star-metrics, or path-metrics.

Yang's model is quite general in the sense that it encapsulates many previously studied problems. For instance, the widely-studied approval voting (AV) rule and the Chamberlin-Courant approval voting rule are respectively the utilitarian rule under the sum function and the utilitarian rule under the min function, when candidates are embedded in a discrete metric. The model is also closely related to the classic clustering problems k-Median, k-Supplier, and k-Center. We refer to [27] and references therein for more detailed discussion.

This paper takes a further step toward a better understanding of this model by exploring the complexity of the same problems but with respect to the Hausdorff distance (function), arguably one of the most appealing concepts to measure similarity between sets. Since its inception [14, 23], Hausdorff function has been pervasive in almost all fields involving distance measurements between sets [1, 2, 8, 15, 17, 25]. Notably, it has attracted considerable interest from the social choice community [4, 6, 12].

Among the set functions mentioned above, the Hausdorff function and the min function are both metrics, while the other two are not, making the max and the sum functions inappropriate in certain circumstances. A downside of the min function is that it is computationally hard to compute winners for both utilitarian rules and egalitarian rules, even in very extreme cases [27].

Our investigation yields plenty of complexity results. In particular, while we have hardness results for the Hausdorff function in general, we also derive numerous polynomial-time algorithms for several special metrics. Our overarching conclusion is that the Hausdorff function significantly outperforms the min function in terms of the complexity of computing winners.

Our work aligns with the literature on approval-based multiwinner voting, especially with papers that investigate the complexity of computing winning committees. We refer to the survey by Lackner and Skowron [18] for a comprehensive overview.

2 PRELIMINARIES

For a graph G and two vertices c and c' in G, we use $d_G(c, c')$ to denote the *distance* between c and c' in G, i.e., the length of a shortest path between c and c' in G. A metric f on a set C is a *shortest-path metric* (SPM) if there is a graph G with C being the vertex set such that for all $c, c' \in C$, it holds $f(c, c') = d_G(c, c')$. We

This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

Table 1: The W[1]/W[2]-hardness results are with respect to the size of the winning committee. Results marked by ◇ mean that they hold even when the input election is both candidates interval and voters interval.

		shortest-path	discrete	star	path
HD	UWD	**W[1]-h**$^{\diamond}$	**P**	**W[1]-h**	**NP-h**
	EWD	**W[2]-h**$^{\diamond}$	**P**	**P**	**P**
min [5]	UWD	W[2]-h	W[2]-h	W[2]-h	W[2]-h
	EWD	W[2]-h	W[2]-h	W[2]-h	W[2]-h
max [27]	UWD	W[1]-h	W[1]-h	W[1]-h	W[1]-h
	EWD	P	P	P	P
sum [27]	UWD	P	P	P	P
	EWD	W[2]-h	W[2]-h	W[2]-h	NP-h

call G a *graph-witness* of f. For simplicity, we call a SPM with a complete/star/path graph-witness a *complete/star/path-metric*. A complete-metric is also called a *discrete metric* in the literature. Star-metrics are pertinent to scenarios where a limited number of expensive resources need to be deployed on numerous computers connected in a star-like network. The decisions on which computers receive the resources are collectively made by a group of experts or determined by a set of criteria, where each expert or each criterion favors a subset of locations. Path-metrics are relevant for scenarios where a government decides to build a new railway line through certain cities. The locations for setting up the stations are determined by a set of criteria or collectively decided upon by a group of professionals.

In candidates-embedded voting, there is a set C of *candidates*, a multiset V of *votes* submitted by a set of voters, and a *metric* $f : C \times C \to \mathbb{R}$. Each vote over C is defined as a nonempty subset of C consisting of all candidates approved by the corresponding voter. The pair (C, V) is called an *election*. A subset of k candidates is called a *k-committee*. A *k-winners selection rule* (k-WSR) maps each triple (C, V, f) to a class of k-committees, the *winning k-committees*. We study rules selecting k-committees as close as possible to the votes, where the closeness is measured by certain *set functions* $g_f : 2^C \times 2^C \to \mathbb{R}$. These rules can be categorized into two classes: the *utilitarian rules* and the *egalitarian rules*.

- Under a *g_f-utilitarian k-WSR*, optimal committees are k-committees $w \subseteq C$ that minimize $\sum_{v \in V} g_f(w, v)$ among all k-committees of C.
- Under a *g_f-egalitarian k-WSR*, optimal committees are k-committees $w \subseteq C$ that minimize $\max_{v \in V} g_f(w, v)$ among all k-committees of C.

Under the max set function, the distance from a committee w to a vote v is $g_f^{\max}(w, v) = \max_{c \in w} (\min_{c' \in v} f(c, c'))$. The Hausdorff distance (HD) between v and w with respect to f is[1]

$$g_f^{\mathrm{HD}}(v, w) = \max\{g_f^{\max}(v, w), g_f^{\max}(w, v)\}.$$

[1]The original definition of Hausdorff distance applies to a much more general setting. For a more elaborate discussion, we refer to [19].

It is obvious that $g_f^{\mathrm{HD}}(v, w) = g_f^{\mathrm{HD}}(w, v)$. We call $\sum_{v \in V} g_f^{\mathrm{HD}}(v, w)$ (respectively, $\max_{v \in V} g_f^{\mathrm{HD}}(v, w)$) the *utilitarian-HD* (respectively, *egalitarian-HD*) distance between w and V with respect to f.

An election is candidate interval (CI) if there is a linear order of the candidates such that every vote approves only consecutive candidates in the order. An election is voter interval (VI) if there is a linear order of the votes such that for every candidate the votes approving the candidate are consecutive in the order. These two specific domains have received considerable attention. Particularly, many voting problems become polynomial-time solvable when restricted to one of the two domains (see, e.g., [10, 11, 20]).

In the paper, we study the following problems.

Utilitarian/Egalitarian-HD Winners Determination (UWD-HD/EWD-HD)

Input: A set C of candidates, a multiset V of votes over C, a metric $f : C \times C \to \mathbb{R}$, a nonnegative integer $k \leq |C|$, and a nonnegative number s.

Question: Is there a k-committee $w \subseteq C$ such that the uti/egalitarian-HD distance between w and V with respect to f is at most s?

If we replace HD in the above definition by other set functions X, we obtain the UWD-X and the EWD-X problems.

3 CONCLUSION

We studied the complexity of the winner determination problems under candidates-embedded approval-based multiwinner voting with respect to the HD. Our exploration leads to a comprehensive understanding of the (parameterized) complexity of the problems. Our results reveal that the HD exhibits distinct behavior in terms of the complexity of these problems, compared to the min, the max, and the sum functions. For instance, for the latter three functions, the complexities of UWD and EWD stay the same when restricted to star-metrics and restricted to discrete metrics [27]. In contrast, UWD-HD is hard to solve when restricted to path-metrics but becomes polynomial-time solvable when restricted to discrete metrics. Our results indicate that, both being metrics, the HD function outperforms the min function concerning the complexity of computing winners. For a more fine-grained comparison, we refer to Table 1.

We point out that all reductions established in the paper can be carried out in polynomial time. As problems studied in the paper are in NP, if a problem is shown to be W[1]-hard or W[2]-hard, it is also NP-complete.

For future research, it would be intriguing to study meaningful parameters leading to fixed-parameter tractability results. Two natural parameters are the number of candidates m and the number of votes n. Both UWD-HD and EWD-HD are easily seen to be FPT with respect to m. It is worth mentioning that one of our reductions implies that EWD-HD is W[2]-hard with respect to k even when $n = 2$ and $s = 1$. However, whether UWD-max and EWD-sum are FPT with respect to n remained open. In addition, as our research is only theoretic-based, conducting an experimental work is another important avenue for future research.

REFERENCES

[1] P. K. Agarwal, S. Har-Peled, M. Sharir, and Y. Wang. 2010. Hausdorff Distance under Translation for Points and Balls. *ACM Trans. Algorithms* 6, 4 (2010), 71:1–71:26. https://doi.org/10.1145/1824777.1824791

[2] S. G. Aksoy, K. E. Nowak, E. Purvine, and S. J. Young. 2019. Relative Hausdorff Distance for Network Analysis. *Appl. Netw. Sci.* 4, 1 (2019), 80:1–80:25. https://doi.org/10.1007/s41109-019-0198-0

[3] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. 2018. On the Complexity of Extended and Proportional Justified Representation. In *AAAI*. 902–909. https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17279

[4] J. S. Banks and J. Duggan. 2008. A Dynamic Model of Democratic Elections in Multidimensional Policy Spaces. *Q. J. Polit. Sci.* 3, 3 (2008), 269–299.

[5] N. Betzler, A. Slinko, and J. Uhlmann. 2013. On the Computation of Fully Proportional Representation. *J. Artif. Intell. Res.* 47 (2013), 475–519. https://doi.org/10.1613/jair.3896

[6] F. Brandl and F. Brandt. 2019. Arrovian Aggregation of Convex Preferences. *Econometrica* 88, 2 (2019), 799–844.

[7] M. Brill, P. Gölz, D. Peters, U. Schmidt-Kraepelin, and K. Wilker. 2020. Approval-Based Apportionment. In *AAAI*. 1854–1861.

[8] B. B. Chaudhuri and A. Rosenfeld. 1999. A Modified Hausdorff Distance Between Fuzzy Sets. *Inf. Sci.* 118, 1-4 (1999), 159–171. https://doi.org/10.1016/S0020-0255(99)00037-7

[9] V. D'Hondt. 1878. *Question électorale: La représentation proportionnelle des partis.* Bruxelles.

[10] E. Elkind and M. Lackner. 2015. Structure in Dichotomous Preferences. In *IJCAI*. 2019–2025. http://ijcai.org/Abstract/15/286

[11] E. Elkind, M. Lackner, and D. Peters. 2022. Preference Restrictions in Computational Social Choice: A Survey. *CoRR* abs/2205.09092 (2022). https://doi.org/10.48550/arXiv.2205.09092 arXiv:2205.09092

[12] T. Erber and M. J. Frank. 2019. Arrow, Hausdorff, and Ambiguities in the Choice of Preferred States in Complex System. *J. Interdiscip. Math.* 22, 2 (2019), 129–137.

[13] G. Gawron and P. Faliszewski. 2019. Robustness of Approval-Based Multiwinner Voting Rules. In *ADT*. 17–31. https://doi.org/10.1007/978-3-030-31489-7_2

[14] F. Hausdorff. 2014. *Grundzuege der Mengenlehre.* Viet, Leipzig.

[15] D. P. Huttenlocher, G. A. Klanderman, and W. Rucklidge. 1993. Comparing Images Using the Hausdorff Distance. *IEEE Trans. Pattern Anal. Mach. Intell.* 15, 9 (1993), 850–863. https://doi.org/10.1109/34.232073

[16] R. Izsak, N. Talmon, and G. J. Woeginger. 2018. Committee Selection with Intraclass and Interclass Synergies. In *AAAI*. 1071–1078. https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17312

[17] O. Jesorsky, K. J. Kirchberg, and R. Frischholz. 2001. Robust Face Detection Using the Hausdorff Distance. In *AVBPA*. 90–95. https://doi.org/10.1007/3-540-45344-X_14

[18] M. Lackner and P. Skowron. 2023. *Multi-Winner Voting with Approval Preferences.* Springer. https://doi.org/10.1007/978-3-031-09016-5

[19] F. Mémoli. 2012. Some Properties of Gromov-Hausdorff Distances. *Discret. Comput. Geom.* 48, 2 (2012), 416–440. https://doi.org/10.1007/s00454-012-9406-8

[20] D. Peters. 2018. Single-Peakedness and Total Unimodularity: New Polynomial-Time Algorithms for Multi-winner Elections. In *AAAI*. 1169–1176.

[21] E. Phragmén. 1894. Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis. *Öfversigt af Kongliga Vetenskaps Akademiens Förhandlingar* 51, 3 (1894), 133–137.

[22] G. Pierczynski and P. Skowron. 2019. Approval-Based Elections and Distortion of Voting Rules. In *IJCAI*. 543–549. https://doi.org/10.24963/ijcai.2019/77

[23] D. Pompeiu. 1905. Sur la Continuité des Fonctions de Variables Complexes (Thèse). *Ann. Fac. Sci. de Toulouse* 7 (1905), 264–315.

[24] T. N. Thiele. 1895. Om Flerfoldsvalg. *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger* (1895), 415–441.

[25] J.-Q. Wang, J.-T. Wu, J. Wang, H.-Y. Zhang, and X-H. Chen. 2016. Multi-Criteria Decision-Making Methods Based on the Hausdorff Distance of Hesitant Fuzzy Linguistic Numbers. *Soft Comput.* 20, 4 (2016), 1621–1633. https://doi.org/10.1007/s00500-015-1609-5

[26] Y. Yang. 2019. Complexity of Manipulating and Controlling Approval-Based Multiwinner Voting. In *IJCAI*. 637–643. https://doi.org/10.24963/ijcai.2019/90

[27] Y. Yang. 2022. On the Complexity of Calculating Approval-Based Winners in Candidates-Embedded Metrics. In *IJCAI*. 585–591. https://doi.org/10.24963/ijcai.2022/83

[28] Y. Yang and J. Wang. 2018. Multiwinner Voting with Restricted Admissible Sets: Complexity and Strategyproofness. In *IJCAI*. 576–582. https://doi.org/10.24963/ijcai.2018/80