Technical Report: Satisfying Constrained Multiple Selection Queries Using IRV

ACM Reference format:
. 2024. Technical Report: Satisfying Constrained Multiple Selection Queries Using IRV. In *Proceedings of SIGMOD 2024, Chile, Santiago,* 3 pages. https://doi.org/10.1145/1122445.1122456

0.1 Hardness Results

THEOREM 0.1. **MqIRV** *is NP-Complete, even when $\ell = 2$.*

Consider an election in which m voters need to elect $k = 1$ candidate out of n candidates. In the election, each voter casts his/her ballot for two candidate in ranked order. The final candidate is determined using the IRV process. For a given instance of the election, we define the margin as the number of ballot changes required to ensure that a specific candidate wins.

We prove that computing the margin is NP-Complete. Our proof is inspired by the NP-Hardness proof of [1]. It is straightforward that the problem is in NP since the outcome of an IRV election can be computed in polynomial time. The NP-hardness is proved by reduction from the 3-Exact Cover problem (3XC). In this problem, we are given a universal set $\{e_1, \ldots, e_{3n}\}$, and $m \geq n$ subsets $S_1, \ldots, S_m$ of size 3 each. We need to determine whether there are n sets whose union is the universal set.

Suppose that we are given an instance of the 3XC problem. We show how to define a related IRV margin problem and then prove that the IRV has a margin n if and only if the answer to the respective 3XC problem is Yes.

The IRV problem has $2m+3n+2$ candidates $b_1, \ldots, b_m, c_1, \ldots, c_m$, $e_1, \ldots, e_{3n}$ and u, v. We must ensure that u wins the election. There are $6m + m^2 + m(m + 5) + 3n(2m + 10) + 2m + 11 + 2m + 13 = 2m^2 + 6mn + 15m + 30n + 24$ ballots as follows:

- For every $i \in [1..m]$, let $S_i = \{e_x, e_y, e_z\}$. There are 6 ballots that we call "cover ballots". These ballots are two of each $[b_i, e_x]$, $[b_i, e_y]$, and $[b_i, e_z]$
- For every $i \in [1..m]$ there are m ballots $[b_i, c_i]$
- For every $i \in [1..m]$ there are $m + 5$ ballots $[c_i, b_i]$
- For every $i \in [1..3n]$ there are $2m + 10$ ballots $[e_i, v]$
- There are $2m + 11$ ballots $[v, u]$
- There are $2m + 13$ ballots $[u, v]$

Suppose that the 3XC instance has an exact cover. Let the indices of the sets in the cover be $j_1 \ldots, j_n$. We change n ballots as follows. For every $i \in [1..n]$ change a ballot $[b_{j_i}, c_{j_i}]$ to $[c_{j_i}, b_{j_i}]$.

We successively eliminated all candidates who got the least number of votes, which is initially $m+5$. There are m candidates with this number of votes: $m - n$ candidates c_x, for $x \in [1..m] \setminus \{j_1 \ldots, j_n\}$, and n candidates b_x, for $x \in \{j_1 \ldots, j_n\}$. As a result of eliminating the $m - n$ candidates c_x, the number of votes of the candidates b_x, for $x \in [1..m] \setminus \{j_1 \ldots, j_n\}$ jumps to $2m + 11$. As a result of eliminating the n candidates b_x, the number of votes of the candidates c_x, for $x \in \{j_1 \ldots, j_n\}$, jumps to $2m + 5$. Also, since the union of the n sets S_x, $x \in \{j_1 \ldots, j_n\}$, is the universal set, the elimination of b_x in the $6n$ "cover ballots" causes the number of votes of *every* e_i to jump to $2m + 12$.

Next, the n remaining candidates c_x, for $x \in \{j_1 \ldots, j_n\}$, with $2m + 5$ votes are eliminated. This does not change the vote of any other candidate. Lastly, the $m - n$ candidates b_x, for $x \in [1..m] \setminus \{j_1 \ldots, j_n\}$, and v each with $2m + 11$ votes are eliminated. None of the e_i is eliminated at this point because all of them have $2m + 12$ votes. Then, all e_is will be deleted, each with $2m + 12$ votes, and, finally, u wins with $2m + 11 + 2m + 13 = 4m + 24$ votes.

We need to prove the other direction. Namely, if the margin is n then there is an exact cover. Suppose that the outcome of the elections can be changed to be u by at most n ballot changes. Since candidate v has one more vote than each of the $3n$ candidates $e_1, \ldots, e_{3n}$, we need to increase the votes of all the candidates $e_1, \ldots, e_{3n}$ by at least 2 so that none of the e_i is eliminated before v is eliminated. Because if any of e_is is eliminated before v is eliminated, then the second choice of e_i's ballot goes to v and the votes of v increase to $4m + 21$. Then all e_i and u will be eliminated, and v wins the election, and u loses. The only way to ensure that none of e_is is eliminated before v is by eliminating some of the candidates b_j. This can be done by ballot changes that reduce the number of votes of some of the candidates b_j by 1 and increase the number of votes of the respective candidates c_j. This will cause some candidates b_j to be eliminated and thus increase the votes of the resulting elements e_i in the "cover ballots" corresponding to these candidates b_j. Since we can make only n ballot changes and since the cover ballots of any candidate b_j change the votes of only the 3 candidates from $\{e_1, \ldots, e_{3n}\}$ that correspond to the set S_j, the n candidates b_j eliminated first must correspond to an exact set.

THEOREM 0.2. DISTTO *is NP-hard, even when $\ell = 3$.*

PROOF. First, we prove that DISTTO is NP-hard when instead of ballot modifications we consider ballot deletions. The proof is by reduction from the 3-Exact Cover problem (3XC) described earlier. In the 3XC problem we are given a universal set $\{v_1, \ldots, v_{3n}\}$, and $m > n$ subsets $S_1, \ldots, S_m$ of size 3 each. We need to determine whether there are n subsets whose union is the universal set. Given an instance of the 3XC problem, we show how to reduce it to an instance of DISTTO. The instance of DISTToconsists of $3n + 1$ candidates $v_1, v_2, \ldots, v_{3n+1}$, and the elimination order $\pi = v_1, v_2, \ldots, v_{3n+1}$ ($\pi[1] = v_1$ is eliminated first, and $\pi[3n + 1] =$

SIGMOD 2024, Santiago, Chile

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/1122445.1122456

v_{3n+1} is the winner). We show that this elimination order can be achieved with n ballot deletions iff the 3XC instance has a positive answer. The polynomial number of ballots in the instance varies in size from 3 to 1 and is described below.

Ballots of size 3: There are m ballots of size 3, one per every subset S_i, $1 \leq i \leq m$. Consider a subset $S_i = \{v_x, v_y, v_z\}$. From now on, we assume that the subset is "ordered", that is, $1 \leq x < y < z \leq 3n$. For every such subset S_i, the ballot (v_x, v_y, v_z) is added, namely v_x is the top candidate in the ballot, v_y is the second candidate, and v_z is the bottom candidate.

Ballots of size 2: For $1 \leq x < y \leq 3n$, let c_{xy} be the sum of the following 2 numbers: (1) number of ballots of size 3 in which v_x is the top candidate and v_y is the second candidate and (2) the number of ballots of size 3 in which v_x is the second candidate and v_y is the bottom candidate (note that the index of the top candidate in this case is lower than x). Let $maxc_x = \max_{y=x+1}^{3n}\{c_{xy}\}$. For every $x < y \leq 3n$, there are $maxc_x - c_{xy}$ ballots of size 2 consisting of candidate v_x as the top candidate and v_y as the second candidate. There are also $maxc_i$ ballots consisting of candidate v_i as the top candidate and candidate v_{3n+1} as the second candidate.

The total number of size 2 ballots is bounded by $6nm - 2m$ since there are at most $(3n-1) \cdot maxc_x$ size 2 ballots with v_x as the top candidate for $1 \leq i \leq 3n$, and $\sum_{x=1}^{3n} maxc_x \leq 2m$.

Ballots of size 1: For $1 \leq x \leq 3n$, let d_x be the total number of ballots of size 3 and size 2 in which v_x is the top candidate. Let $maxd = \max_{y=1}^{3n}\{d_y\}$. There are $maxd - d_x$ ballots of size 1 consisting only of candidate v_x as the top candidate and the only candidate. There are also $maxd - 1$ ballots consisting of only candidate v_{3n+1} as the top and only candidate. The number of ballots of size 1 is bounded by $18n^2m - 3nm$ since at most $3n$ candidates have single ballots, and for each of these candidates, there are at most $m + 6nm - 2m$ ballots of size 1, since this is an upper bound on the number of ballots of size 2 and 3 per candidate.

We prove that if there is an exact cover, then the margin is n. Suppose that the 3XC instance has an exact cover consisting of n sets. Each such set corresponds to a ballot of size 3. We call these ballots the "cover ballots". For $1 \leq x \leq 3n$, let $b(x)$ be the unique cover ballot containing x. We prove below that after deleting the n cover ballots the IRV process will result in the elimination order $v_1, v_2 ... v_{3n+1}$.

By our construction, before the deletion of the cover ballots, each of the candidates $v_1, \ldots, v_{3n}$ is the top candidate on the $maxd$ ballots and v_{3n+1} is the top candidate on the $maxd - 1$ ballots. Since the candidates on every ballot are ordered, v_1 must be the top candidate in ballot $b(1)$ and thus after the removal of this ballot, v_1 is the top candidate in $maxd - 1$ ballots. Also, since no candidate appears more than once in the cover ballots, after their removal, each of the candidates $v_2, \ldots, v_{3n}$ is the top candidate on either $maxd - 1$ or $maxd$ ballots. Recall that ties are broken arbitrarily, and thus we can eliminate v_1. As a result of the elimination of v_1 the top candidate in all ballots that included v_1 (and are not of size 1) is updated. By our construction, there are exactly $maxc_1$ such ballots for each of the candidates $v_2, \ldots, v_{3n+1}$. After the elimination of v_1, v_2 must be the top candidate in ballot $b(2)$ and therefore after the removal of this ballot v_2 is the top candidate in $maxc_1 + maxd - 1$ ballots. Again, no candidate can be the top candidate in less than $maxc_1 + maxd - 1$ ballots and thus v_2 can be eliminated. Continuing in the same manner, after the elimination of $v_1, \ldots, v_{x-1}$, candidate v_x must be the top candidate in ballot $b(x)$ and thus after the removal of this ballot v_x is the top candidate in $\sum_{y=1}^{x-1} maxc_y + maxd - 1$ ballots and can be eliminated as dictated by the required elimination order.

In the other direction, we prove that if the margin is n then there is an exact cover. To achieve this goal, we show that any set of ballots whose removal results in the elimination order $v_1, v_2 ... v_{3n+1}$ must include the candidates $v_1, v_2 ... v_{3n}$. We prove this by contradiction. Assume that this is not the case and that there exists a set of ballots that do not include a candidate v_x whose removal results in the required elimination order. Let v_x be the candidate with the minimum index that is not included in the deleted ballots. In this case, by our construction, when v_x is about to be eliminated, it is the top candidate of $\sum_{y=1}^{x-1} maxc_y + maxd$ ballots, while v_{3n+1} is the top candidate of $\sum_{y=1}^{x-1} maxc_y + maxd - 1$ ballots. A contradiction. Clearly, the only way to delete n ballots that include all $3n$ candidates $v_1, v_2 ... v_{3n}$ is by choosing ballots of size 3 that correspond to an exact cover.

Next, we extend this proof to the case of ballot modifications. We use the same ballot profile as before with only one difference: candidate v_{3n+1} has $maxd - n - 1$ ballots, that is, $n+1$ fewer ballots than any other candidate (instead of having 1 ballot less than the others). By a similar reduction, it can be shown that in this scenario, the 3XC problem instance has an exact cover iff the optimal solution to the DistToinstance consists of n ballot modifications where the ballots removed in these modifications include candidates $v_1, v_2 ... v_{3n}$ and each of the added n ballots includes candidate v_{3n+1} as the top and only candidate.

□

REFERENCES

[1] John J. Bartholdi and James B. Orlin. 1991. Single transferable vote resists strategic voting. *Social Choice and Welfare* 8, 4 (1991), 341–354. http://www.jstor.org/stable/41105995