Online (Budgeted) Social Choice

Brendan Lucier, Joel Oren

Abstract

We consider classic social choice problems in an online setting. In the problems we consider, a decision-maker must select a subset of candidates in accordance to reported preferences, e.g. to maximize the value of a scoring rule. However, agent preferences cannot be accessed directly; rather, agents arrive one at a time to report their preferences, and each agent cares only about those candidates that have been selected by the time she arrives. On each step, the decision maker must choose whether to irrevocably add candidates to the final selection set given the preferences observed so far, with the goal of maximizing the average score over all agents.

We show that when preferences are arbitrary but agents arrive in uniformly random order, an online selection algorithm can approximate the optimal value of an arbitrary positional scoring function to within a factor of $(1-1/e)-o(1)$ as the number of agents grows large, nearly matching the performance of the best offline polynomial-time algorithm. When agent preferences are drawn from a Mallow's model distribution, a different selection algorithm achieves approximation factor that limits to 1 as the number of agents grows large. Our methods are straightforward to implement, and draw upon connections to online computation and secretary problems.

1 Introduction

Suppose that a manufacturer wishes to focus on a selected set of possible products to offer to incoming consumers. On each day a new client arrives, selecting her favorite product among those being offered. However, the client may also express preferences over *potential* products, including those that are not currently being offered. The manufacturer must then decide whether or not to add new production lines to make available to that consumer (as well as to future consumers). While adding a new product would potentially increase customer welfare, it carries with it some opportunity cost: it would be impractical to offer every possible product, so choices are effectively limited and irrevocable (since new production lines incur substantial overhead). Adding new products may be worthwhile if many future customers would prefer the chosen product as well, though this is not known to the manufacturer in advance. The problem is thus one of online decision-making, where uncertainty of future preferences must be balanced with the necessity of making decisions to realize current gains.

Such a setting gives rise to obvious complications. On one hand, adding an item that is highly ranked by the current user to the list of available items will satisfy the current client. On the other hand, in such settings there is usually an underlying constraint that prohibits the addition of arbitrarily many items. In our study of this problem, we will address settings in which the underlying restriction is a cardinality constraint, which limits the number of chosen alternatives. The main problem that we are facing is therefore an online social choice problem: we are required to choose the most "favourable" set of candidates, while having only a partial view of the objective function. For any given offline social choice problem, such as selecting a candidate to maximize the value of a certain scoring function of the user preferences, one might consider an online variant in which each agent receives value only for those candidates that have been selected at or before the time that he arrives. In this case, the objective function is the average of the agent scores, determined by a prescribed positional scoring function or hidden utility function.

We consider various different models for the manner in which the agents preferences are set. In the distributional model, the player preferences are drawn independently from a distribution over permutations. For example, one might assume that the preferences are sampled from a parametrized Mallows model, which defines a unimodal distribution over permutations. An alternative approach that is common to online algorithm analysis is to suppose that the set of agent preferences is set arbitrarily (i.e. adversarially), but that the order of agent arrival is random[1]. Utilizing previous results in the area of online matching, we show that our methods for this adversarial setting carry over to the case in which the preferences are drawn independently from an *unknown* distribution.

As previously mentioned, our ultimate goal is to maximize the average score of the agents when each agent is matched with his most preferred item available at the time of arrival. Generally speaking, our finding is that if the number of agents is sufficiently large compared to the number of candidates, it is possible to design online algorithms that perform asymptotically as well as the best possible offline algorithms, with high probability. Our approach is reminiscient of those used for well-studied secretary-type problems, in which the candidates arrive online rather than the agents. Our results also suggest a number of potential extensions for future research, which we discuss in our concluding remarks.

Results We first consider adversarial settings, where agent preferences are arbitrary but arrive in uniformly random order. We show that one can approximate the optimal choice of a *single* candidate, with respect to an *arbitrary* positional scoring function, with approximation ratio $(1 - o(1))$ where the asymptotic notation is with respect to the number of agents. In other words, the regret exhibited by the online selection method vanishes as n grows large. If more than one alternative can be chosen, say $k > 1$ in total, we show that for any positional scoring function, combining our sample approach with a standard greedy algorithm for submodular set-function maximization provides a $(1 - \left(\frac{k-1}{k}\right)^k - o(1))$ approximation to the optimal choice. Thus, as n grows large, our online algorithm achieves approximation factor $1 - 1/e$, matching a lower bound for offline algorithms [12].

Moving away from positional scoring functions to arbitrary utility functions, we apply a recent result due to Boutilier et al. [3] who demonstrated that a social choice function can approximate the choice of a candidate to maximize agent utilities to within a factor of $\tilde{O}(\sqrt{m})$ (where m is the number of candidates), even if only preference lists are made available. Using our results for arbitrary positional scoring functions, we obtain similar bounds for the problem of maximizing average agent utility in an online fashion, with vanishing additional errors due to sampling.

Finally, in the distributional setting where preferences are drawn from a parameterized Mallows model, we show that for the selection of $k \geq 1$ alternatives under an arbitrary positional scoring rule, one can obtain an approximation ratio of $(1 - o(1))$, suffering regret that vanishes as n grows. In the particular case of Borda ranking, we show that sampling a logarithmic number of agents is sufficient for approximating the optimal k-set.

2 Related Work

The problem of selecting a single candidate given a sequence of agent preference lists is the traditional social choice problem. The budgeted form of this offline problem was introduced by Chamberlin and Courant [5], and subsequently studied by Boutilier and Lu [12], in which several natural constraints on the allocated set were considered. In particular, it is shown that for the case where producing copies of the alternatives bears no cost, the problem of selecting which candidates to make available is a straightforward case of non-decreasing and

[1] For the problems we consider, as with many others, no algorithm can guarantee reasonable performance if the adversary is also allowed to set the order of the arrival of agents.

submodular set-function maximization, subject to a cardinality constraint, which admits a simple greedy algorithm with approximation ratio $1 - 1/e$. Our work differs in that the agent preferences arrive online, complicating the choice of which alternatives to select, as the complete set of agents preference is not fully known in advance.

In our online setting, we refer to the Mallows model ([14]), a well-studied model for distributions over permutations (e.g. [8, 6]) which has been studied and extended in various ways. In recent work, Braverman and Mossel have shown that the sample complexity required to estimate the maximum-likelihood ordering of a given Mallows model distribution is roughly linear [4]. We make use of some of their results in our analysis.

Adversarial and stochastic analysis in online computation have received considerable attention (e.g. [7]). In our analysis, we make critical use of the assumption that agent arrivals are randomly permuted. This is a common assumption in online algorithms (e.g. [10, 11, 13]). Correspondingly, in our analysis of the adversarial model, we use techniques that resemble methods used in secretary and multi-armed bandits problems (see [2] for a survey), of partially observing some initial data, and bounding the total error.

A recent paper by Boutilier et al. [3] considered the social choice problem from a utilitarian perspective, where agents have underlying utility functions that induce their reported preferences. The authors introduce a measure of *distortion* to compare the performance of their social choice functions to the social welfare of the optimal alternative. We make use of their constructions in our results for the utilitarian model.

The online arrival of preferences has been previously studied by Tennenholtz [16]. This work postulates a set of voting rule axioms that are compatible with online settings.

3 Preliminaries

Given is a ground set of alternatives (candidates) $A = \{a_1, \dots, a_m\}$. An agent $i \in N = \{1, \dots, n\}$, has a preference $\succ_i$ over the alternatives, represented by a permutation π^i. For a permutation π and an alternative $a \in A$, we will let $\pi(a)$ denote the rank of a in π. A *positional scoring function* (PSF) assigns a score v_i to the alternative ranked ith, given a prescribed vector $\mathbf{v} \in \mathbb{R}^m_{\geq 0}$. Given an (implicit) set of agent preferences, we will denote the average score of a *single* element $a \in A$ by $\overline{F}(a) = \frac{1}{n}\sum_{i=1}^{n} F_i(a)$, where $F_i(a) = \mathbf{v}(\pi^i(a))$. Moreover, we will consider the score of a *set* $S \subseteq A$ of candidates w.r.t. to a set of agents as the average positional scores of each of the agents, assuming that each of them selected their highest ranked candidate in the set: $\overline{F}(S) = \frac{1}{n}\sum_{i \in N} \max_{a \in S} F_i(a)$.

The online budgeted social choice problem. We consider the problem of choosing a set of $k \geq 1$ candidates from the set of potential alternatives. An algorithm for this problem starts with an empty "slate" $S_0 = \emptyset$ of alternatives, of prescribed capacity $k \leq m$. In each step $t \in [n]$, an agent arrives and reveals her preference ranking. Given this, the algorithm can either add new candidates $I \subseteq A \setminus S_{t-1}$ to the slate (i.e. set $S_t \leftarrow S_{t-1} \cup I$), if $|S_{t-1}| + |I| \leq k$, or leave it unchanged. Agent i in turn takes a copy of one of the alternatives *currently* on the slate, i.e. S_t. Any addition of alternatives to the slate is *irrevocable*: once an alternative is added, it cannot be removed or replaced by another alternative. The offline version of this problem is called the limited choice model in [12].

Some of our results will make use of algorithms for maximizing non-decreasing submodular set functions subject to a cardinality constraint. A submodular set function $f : 2^U \rightarrow \mathbb{R}_{\geq}$ upholds $f(S \cup \{x\}) - f(S) \geq f(T \cup \{x\}) - f(T)$ for all $S \subseteq T \subseteq U$ and $x \in U \setminus T$.

4 The Adversarial Model

We begin by supposing that the set of agent preference profiles is arbitrary, as might be chosen by an adversary. After the collection of all preference profiles has been fixed, we assume that they are presented to an online algorithm in a uniformly random order. The algorithm can irrevocably choose up to k candidates during any step of this process; each arriving candidate will then receive value corresponding to his most-prefered candidate that has already been chosen. The goal is to maximize the value obtained by the algorithm, with respect to an arbitrary positional scoring function[2].

In general, we cannot hope to achieve an arbitrarily close approximation factor to the optimal (in hindsight) choice of k candidates, as it is **NP**-hard to obtain better than a $(1-\frac{1}{e})$ approximation to this problem even when all profiles are known in advance[3]. Our goal, then, is to provide an algorithm for which the approximation factor approaches $1-\frac{1}{e}$ as n grows, matching the performance of the best-possible algorithm for the offline problem[4].

Let $F(\cdot)$ be an arbitrary PSF; without loss of generality we can scale F so that $F(1)=1$. Note that this implies that $F(a) \in [0,1]$ for each outcome a. If agent i has preference permutation π^i, then we write $F_i(\cdot) = F(\pi^i(\cdot))$ for the scoring function F applied to agent i's permutation of the choices. Also, we will write σ for the permutation of players representing the order in which they are presented to an online algorithm. Thus, for example, $F_{\sigma(1)}(a)$ denotes the value that the first observed player has for object a.

Given a set S of objects and PSF F, we write $F(S) = \max_{a\in S} F(a)$ for the value of the highest-ranked object in S. Given a set T of players, $F_T(S) = \sum_{j\in T} F_j(S)$ is the total score held by the players in T for the objects in S. We also write $\overline{F}_T(S) = \frac{F_T(S)}{|T|}$ for the average score assigned to set S. Let $OPT = \max_{S\subseteq A, |S|\leq k} F_N(S)$ be the optimal outcome value.

Let us first describe a greedy social choice rule for the offline problem that achieves approximation factor $(1-1/e)$, due to [12]. This algorithm proceeds by repeatedly selecting the candidate that maximizes the marginal gain in the objective value, until a total of k candidates have been chosen. As any PSF $F(\cdot)$ can be shown to be a (non-decreasing) submodular set-function over the sets of candidates (see for example, [12]), such an algorithm obtains approximation $1-(\frac{k-1}{k})^k$, which is at most $1-1/e$ for all k. We will write $Greedy(N,k)$ for this algorithm applied to set of players N with cardinality bound k.

We now consider the online algorithm $\mathcal{A}$, listed as Algorithm 1 below.

Algorithm 1: Online Candidate Selection Algorithm

Input: Candidate set A, parameters k and n, online sequence of preference profiles

Let $t \leftarrow n^{2/3}(\log n + k\log m)$;

Observe the first t agents, $T = \{\sigma(1), \ldots, \sigma(t)\}$;

$S \leftarrow Greedy(T,k)$;

Choose all candidates in S and let the process run to completion;

We write $V(\mathcal{A})$ for the value obtained by this algorithm. We claim that the expected value obtained by $\mathcal{A}$ will approximate the optimal offline solution.

[2] A stronger adversary would not only be able to set the preferences of the voters, but also their order, or even set preferences adaptively. However, it is not hard to see that in such cases no non-trivial bounds can be obtained, as the adversary can strategically cause the algorithm to exhaust its budget and then set the preferences to be the worst possible from that point onward.

[3] One can reduce Max-k-Coverage to the budgeted social choice problem for the special case of l-approval: the PSF in which the first l positions receive score 1, and others receive score 0.

[4] For the special case of the Borda scoring rule, it can be shown that the algorithm that simply select a random k-set obtains a $1-O(1/m)$-approximation to the offline problem. Furthermore, this algorithm can be derandomized using the method of conditional expectations. We omit the proof due to space considerations.

Theorem 1. *If $m < n^{1/3-\epsilon}$ for any $\epsilon > 0$, then $E[V(\mathcal{A})] \geq (1 - (\frac{k-1}{k})^k - o(1))OPT$.*

The first step in the proof of Theorem 1 is the following technical lemma, which states that the preferences of the first t players provide a good approximation to the (total) value of every set of candidates, with high probability.

Lemma 2. *$Pr[\exists S, |S| \leq k : |\overline{F}_T(S) - \overline{F}(S)| > n^{-1/3}] < \frac{2}{n}$, where the probability is taken over the order in which the agents arrive.*

Proof. Choose any set S with $|S| \leq k$. For each $j \in [t]$, let X_j be a random variable denoting the value $F_{\sigma(j)}(S)$. Note that $E[X_j] = \overline{F}(S)$ for all j, and that $\overline{F}_T(S) = \frac{1}{t}\sum X_j$. By the Hoeffding inequality (without replacement), for any $\epsilon > 0$, $Pr[|\overline{F}_T(S) - \overline{F}(S)| > \epsilon] < 2e^{-\epsilon^2 t}$. By the union bound over all S with $|S| \leq k$,

$$Pr[\exists S, |S| \leq k : |\overline{F}_T(S) - \overline{F}(S)| > \epsilon] < 2\sum_{\ell=1}^{k} \binom{m}{\ell} e^{-\epsilon^2 t} \leq 2m^k e^{-\epsilon^2 t}.$$

Setting $t = n^{2/3}(\log n + k \log m)$ and $\epsilon = n^{-1/3}$ then yields the desired result. $\square$

With Lemma 2 in hand, we can complete the proof of Theorem 1 as follows. Since $F_T(S)$ approximates $F(S)$ well for every S, our approach will be to sample T, choose the (offline) optimal output set according to the preferences of T, then apply this choice to the remaining bidders. This generates two sources of error: the sampling error bounded in Lemma 2, and the loss due to not serving the agents in T. By setting the value of t judiciously, and noting that OPT cannot be very small (it must be at least $\frac{n}{m}$), one can show that the relative error vanishes as n grows large. The details appear in the full version of the paper.

One special case of note occurs when $k = 1$; that is, there is only a single candidate to be chosen. In this case, the regret experienced by our online algorithm vanishes as n grows.

Corollary 3. *If $k = 1$ and $m < n^{1/3-\epsilon}$ for any $\epsilon > 0$, then $E[V(\mathcal{A})] \geq (1 - o(1))OPT$.*

4.1 A Correspondence with the Unknown Distribution Model

We now note a correspondence between the random order model analyzed above and a model in which rankings are drawn from an underlying distribution over preferences. This observation was first made by Karande et al. ([9]) in the context of online bipartite matching. Suppose there is an underlying distribution $\mathcal{D}$ over the set of rankings over the alternatives A. For each player $i \in N$, suppose the ranking π^i for player i is sampled independently from $\mathcal{D}$.

The following result due to Karande et al. states that our algorithm for the adversarial model with random arrival order applies to this unknown-distribution setting as well.

Claim 4 ([9]). *Let $\mathcal{A}$ be an algorithm for the online social problem under the random order model that obtains a expected competitive ratio of α. Then $\mathcal{A}$ obtains an expected approximation ratio of at least α for the online social choice problem in the unknown distribution model. Furthermore, hardness results in the unknown distribution model hold in the random order model as well.*

This result implies that algorithm $\mathcal{A}$ achieves approximation factor $(1 - (\frac{k-1}{k})^k - o(1))$ to the social choice problem when preferences are drawn from an unknown underlying distribution, and that it is NP-hard to achieve an approximation factor better than $(1 - 1/e)$.

5 A Utilitarian Approach

In the previous section we considered the problem of maximizing the social value of a positional scoring function in an online setting. However, it may be more natural in some circumstances to assume that each agent assigns a non-negative utility to each candidate, even though these utilities are hidden and only the preference lists are revealed to a potential social choice function. In such settings, one would wish to choose candidates that maximize overall social welfare (i.e. sum of utilities), again in an online fashion. However, this goal is hindered by the fact that the utilities themselves are never made available to the algorithm. In this section we adapt a general technique due to Boutillier et al. [3] to show that our result for online PSF maximization extends to approximate online utility maximization.

We assume that each agent $i \in N$ has a latent utility function $u_i \colon A \to \mathbb{R}_{\geq 0}$. A utility function u_i induces a preference profile $\pi(u_i) = \pi^i$ such that $\pi^i(a) > \pi^i(a')$ precisely[5] when $u_i(a) \geq u_i(a')$. We let $\pi(\mathbf{u})$ denote the induced preference profile given a utility profile $\mathbf{u}$.

As in [3], we will assume that utilities can be normalized so that $\sum_{a \in A} u_i(a) = 1$ for each i. This assumption essentially states that each agent has the same total weight assigned to her candidate utilities. Note that without this assumption it would be impossible to approximate the optimal social welfare, since a single agent could have a single utility score that dominates all others, but an algorithm with access only to the preference profiles would have no awareness of this fact.

Intuitively, we would like to choose an alternative $a \in A$ that maximizes the (unknown) social welfare $sw(a, \mathbf{u}) = \sum_{i=1}^{n} u_i(a)$, based solely on the reported vote profile $\overrightarrow{\pi} = \overrightarrow{\pi}(\mathbf{u}) = (\pi^1, \dots, \pi^n)$ induced by the utility profile. Of course, the preference profile $\overrightarrow{\pi}$ does not completely capture all of the information in the utility profile, and hence we should expect some loss.

Our hope will be to find a social choice rule f such that, if it were applied to the preference profile $\overrightarrow{\pi}$, it would return a candidate that approximately maximizes $sw(a, \mathbf{u})$. The *distortion* of f is the worst-case approximation factor incured when f is applied $\overrightarrow{\pi}(\mathbf{u})$. This notion of distortion was first formalized by Procaccia and Rosenschein in [15], and has been used in subsequent studies of the social choice problem with partial (or noisy) information about the underlying utilities (e.g. [3]). The formal definition is as follows.

Definition 5 (distortion)**.** *Let $\overrightarrow{\pi} \in S_m^n$ be a preference profile, and let $f : S_m^n \to A$ be a social choice function. The distortion of f is then given by*

$$dist(\overrightarrow{\pi}, f) = \sup_{\mathbf{u}:\pi(\mathbf{u}) = \overrightarrow{\pi}(\mathbf{u})} \frac{\max_{a \in A} sw(a, \mathbf{u})}{sw(f(\overrightarrow{\pi}), \mathbf{u})} \tag{5.1}$$

In [3], Boutilier et al. proposed a randomized social choice rule f with distortion $O(\sqrt{m \log m})$, and provided a corresponding lower bound of $\Omega(\sqrt{m})$. This rule f makes use of a positional scoring function $H(\cdot)$, that they refer to as the *harmonic scoring function*. In the harmonic scoring function, the score of a candidate ranked in position i is $H_i = 1/i$. Given preference profile $\overrightarrow{\pi}$, rule f either a) with probability 1/2, chooses each candidate a with probability proportional to $H_N(a) = \sum_{i \in N} H(\pi^i(a))$, or b) with the remaining probability 1/2, returns a uniformly random candidate.

We will make use of this social choice rule f to design an online algorithm achieving social welfare within a factor of $O(\sqrt{m \log m})$ of the optimal welfare. As before, we assume an adversarial setting: the collection of agent preferences can be arbitrary, but they are presented to the algorithm in an order determined by a (uniform) random permutation σ. Our algorithm $\mathcal{A}$ is described as Algorithm 2, below.

[5] In keeping with our simplifying assumption that preference profiles do not include indifference, we can assume that ties in utility are broken in some consistent manner.

Algorithm 2: Online Candidate Selection Algorithm for Utility Maximization

Input: Candidate set A, parameter n, sequence of preference profiles arriving online

Let $t \leftarrow n^{2/3} \log n$;
Observe the first t agents, $T = \{\sigma(1), \dots, \sigma(t)\}$;
$a^* \leftarrow f(\pi^{\sigma(1)}, \dots, \pi^{\sigma(t)})$;
Choose candidate a^* and let the process run to completion;

Given a particular utility profile $\mathbf{u}$ we will write $E[sw(\mathcal{A})]$ to denote the expected social welfare of the outcome returned by $\mathcal{A}$, given preference profile $\overrightarrow{\pi}(\mathbf{u})$, over permutations σ and randomness in $\mathcal{A}$. We will also write OPT for the optimal social welfare attainable for $\mathbf{u}$, i.e. $OPT = \max_{a \in A} \sum_i u_i(a)$.

Theorem 6. *Suppose $n > m^3$. Then for all $\mathbf{u}$, $E[sw(\mathcal{A})] \geq \frac{1}{O(\sqrt{m} \log m)} OPT$.*

The idea behind the proof of Theorem 6 is to note that the algorithm for offline utility maximization due to Boutilier et al. [3] works primarily by applying the low-distorition PSF f. However, our Theorem 1 implies that PSF value maximization can be approximated well by an online algorithm. We can therefore approximate the set that maximizes the (offline) value of f in the online setting. As long as the errors due to sampling and omitting the first t agents are not too large, this then implies an approximation to the utility-maximizing candidate set. The details of the proof appear in the full version of the paper.

6 The Distributional Model

We next suppose that agent preferences are distributed according to the well-studied Mallows model, which defines a family of permutation distributions. Roughly speaking, Mallow's model assumes that preferences are aligned according to some base permutation $\hat{\pi}$, but each agent's permutation is (independently) perturbed according to a particular error measure. We begin by giving a formal definition of this distribution.

Let us begin our formal definition by introducing the Kendall-tau distance (which is also known as the Kemeny distance or the bubble-sort distance):

Definition 7 (Kendall-tau distance)**.** *For all $\pi, \pi' \in S_m$, the* Kendall-tau distance *between π and π' is $d_K(\pi, \pi') = \#\{i \neq j : \pi(i) < \pi(j)$ and $\pi'(i) > \pi'(j)\}$.*

Definition 8 (The Mallows model)**.** *Let $\phi \in (0, 1)$ and $\hat{\pi} \in S_m$. The Mallows model distribution $D(\hat{\pi}, \phi)$ is a distribution over permutations of $\{1, \dots, m\}$, such that the probability of a permutation $\pi \in S_m$ is*

$$Pr[\pi] = \phi^{d_K(\pi, \hat{\pi})} / Z \tag{6.1}$$

where Z is a normalization constant: $Z = \sum_{\pi \in S_m} \phi^{d_K(\hat{\pi}, \pi)}$.

Fact 9. *It can be shown that $Z = 1 \cdot (1 + \phi) \cdot \dots \cdot (1 + \dots + \phi^{m-1})$.*

We note that the Mallows model induces a unimodal distribution. Furthermore, the parameter ϕ can be seen as controlling the amplitude of error with respect to permutation $\hat{\pi}$: as ϕ approaches 1 the distribution tends to uniformity, and as ϕ approaches 0 the distribution approaches a point mass at $\hat{\pi}$.

We will assume that the agent preference rankings are drawn independently from a Mallows model distribution $D(\hat{\pi}, \phi)$, where the underlying reference ranking $\hat{\pi}$ is unknown.

We will assume that the *dispersion parameter* ϕ is known in advance. Our optimization task in this model is to select a $S \subseteq A$ of size at most k, in an online fashion, so as to maximize the expected value of S among the remaining agents (with respect to a given positional scoring function).

For simplicity of notation and without loss of generality, from hereon we assume that $\hat{\pi}$ is the identity permutation. That is, $\hat{\pi}(i) = i$. We note that since $D(\hat{\pi}, \phi)$ is a unimodal distribution, Theorem 1 and Claim 4 together imply an immediate corollary for this distributional model.

Theorem 10. *Let $F(\cdot)$ be an arbitrary positional scoring function, and let $\mathcal{A}$ be the online algorithm listed as Algorithm 1. Then if $m < n^{1/3-\epsilon}$ for any $\epsilon > 0$, we have $E[V(\mathcal{A})] \geq (1 - (\frac{k-1}{k})^k - o(1))OPT$.*

Given this result, our motivating question for this section is whether we can obtain improved results by making use of the particular form of the Mallows model.

6.1 An Improved Result for Arbitrary PSFs

Suppose that our goal is to maximize the value of an arbitrary PSF $F(\cdot)$, scaled so that $F(1) = 1$. Write $A_m = \sum_{i=0}^{m-1} \phi^i$. We begin with a lemma about the Mallows model, which shows that in a sampled permutation π, we do not expect any particular candidate to be placed very far from its position in the reference ranking (the proof appears in the appendix of the full version paper):

Lemma 11. *Let $\pi \sim D(\hat{\pi}, \phi)$. Then for any $i \neq j$, $Pr[\pi^{-1}(i) = i] \geq Pr[\pi^{-1}(i) = j] + \frac{1-\phi}{A_m}$.*

Given this lemma, our strategy will be to observe many samples from the distribution, then attempt to guess the identities of the top k elements in the underlying permutation $\hat{\pi}$. Since each candidate is most likely to appear in its position from $\hat{\pi}$, we expect to be able to determine $\hat{\pi}$ after a relatively small number of samples. Our algorithm is provided as Algorithm 3, below.

Algorithm 3: Online Candidate Selection Algorithm for the Mallows Model

---

Input: Candidate set A, Mallows model parameter ϕ, parameter n, sequence of preference profiles arriving online

Let $t \leftarrow 2(\frac{1-\phi}{2A_m})^2 \log m \log n$;

Observe the first t agents, $T = \{\sigma(1), \ldots, \sigma(t)\}$;

For each $i = 1, \ldots, k$, let a_i be the candidate that occurs most often in position i among $\pi^{\sigma(1)}, \ldots, \pi^{\sigma(t)}$.;

Choose candidates $a_1, \ldots, a_k$ and let the process run to completion;

We now show that this algorithm does, indeed, exhibit vanishing regret as n grows large.

Theorem 12. *Suppose that $n > m^{2+\epsilon} \frac{1}{1-\phi}$ for some $\epsilon > 0$. Then algorithm $\mathcal{A}$ satisfies $E[v(\mathcal{A})] \geq (1 - o(1))OPT$.*

The proof of the theorem, which relies on the Hoeffding and the union bound, appears in the full version of the paper.

6.2 The Borda Scoring Rule

We now demonstrate that if our positional scoring function is the canonical Borda scoring function, then we can obtain a good approximation with fewer samples (and hence a weaker restriction on the size of n relative to m). In the Borda positional scoring function, for an agent with preference $\pi \in S_m$, the score is defined as follows: $B_i(a) = m - \pi(a)$; i.e. the scores are evenly spread between 0 and $m-1$.

We begin with a lemma about the Mallows model, which shows that we do not expect the top candidate to be placed very far from its position in the reference ranking:

Claim 13. *Let $\pi \sim D(\hat{\pi}, \phi)$, and let $a = \pi^{-1}(i)$; i.e. the first item in the permutation. Then with high probability $\hat{\pi}(a) = o(m)$.*

Proof. Fix $c \in (0,1)$. Now, consider the probability that any of the elements $\lfloor c \cdot m \rfloor, \ldots, m$ appear in position one in a sampled permutation π:

$$\begin{aligned} Pr[\pi(i) = 1 : i \geq \lfloor c \cdot n \rfloor] &= \sum_{i=\lfloor c\cdot n\rfloor}^{m} \sum_{\pi \in S_m : \pi(i)=1} \frac{\phi^{d_K(\hat{\pi},\pi)}}{Z_m} = \sum_{i=\lfloor c\cdot n\rfloor}^{m} \frac{\phi^{i-1} \cdot Z_{m-1}}{Z_m} \\ &= \sum_{i=\lfloor c\cdot n\rfloor}^{m} \frac{\phi^{i-1}}{1+\phi+\cdots+\phi^{m-1}} \end{aligned} \tag{6.2}$$

The claim follows from the fact that this is essentially a sum of exponentially small terms □

We will complement the above claim by showing that w.h.p. (albeit not necessarily exponentially small), the position of the first element in a sampled permutation in the reference ranking is bounded by $O(\log m)$. We then argue that by sampling more permutations, we can augment our bound. The claims are essentially consequences of the results obtained by Braverman and Mossel. Recall that an equivalent statement of the probability of sampling a permutation is $Pr[\pi] = e^{-\beta i}$, where $\beta = -ln\phi$.

Claim 14 ([4])**.**

$$Pr[\pi^{-1}(1) \geq i] \leq e^{-\beta i}/(1-e^{-\beta}) \tag{6.3}$$

The proof of this claim is similar to the one of Claim 13.

Corollary 15.

$$Pr[\pi^{-1}(1) \geq \ln m] \leq m^{-\beta}/(1-e^{-\beta}) \tag{6.4}$$

The following claim argues that the error in our estimate for the first element in $\hat{\pi}$ goes linearly small with the number of sampled permutations $\sigma^1, \ldots, \sigma^r \sim D(\hat{\pi}, \phi)$.

Claim 16 ([4])**.** *Suppose that the permutations $\pi^1, \ldots, \pi^r$ are drawn from $D(\hat{\pi}, \phi)$, and let $\overline{\pi}(a) = \frac{1}{r}\sum_{i=1}^{r} \pi^i(a)$.*

$$Pr[|\overline{\pi}(\ell) - \ell| \geq i] \leq 2 \cdot \left(\tfrac{(5i+1)\cdot e^{-\beta i}}{1-e^{-\beta}}\right)^r \quad , \quad \textit{for all } i \in [m]. \tag{6.5}$$

Setting $i = \ln n$, we obtain the following corollary:

Corollary 17. *Let $\alpha > 0$. Then for sufficiently large n,*

$$Pr[|\overline{\pi}(a_\ell) - \ell| \geq \frac{\alpha+2}{\beta \cdot r} \ln n] < n^{-\alpha} \tag{6.6}$$

Despite the above results that imply that using the top-ranked element in even a single sample should get us close to the top-ranked element in the reference ranking, we still have to argue that w.h.p., this estimate also approximates the expected top-ranked element, induced by the distribution. The following result provides an affirmative answer to this question.

Theorem 18 ([4])**.** *Let $L = \max\left(6 \cdot \frac{\alpha+2}{\beta \cdot r} \log m, 6 \cdot \frac{\alpha+2+1/\beta}{\beta}\right)$. Then except with probability $< 2 \cdot m^{-\alpha}$, for any maximum-likelihood π^m and for all ℓ, we have*

$$|\pi^m(a_\ell) - \hat{\pi}(a_\ell)| \leq 32L \tag{6.7}$$

where $\hat{\pi}$ is the reference ranking.

So in total, with probability $n^{-\alpha}$, $|\overline{\pi}(a_\ell) - \pi^m(a_\ell)| \leq O(1)$. Thus, we get a natural algorithm for maximizing the average Borda score for all but the first $\log n$ agents:

Theorem 19. *The algorithm that samples the first $\log n$ permutations and puts on the slate the element from A with the highest average score obtains a $1 - O(1/n)$-approximation of the optimal average Borda score.*

The theorem follows from the previous conclusion and by recalling that the maximum value any element can receive is $m - 1$.

6.3 The case of $k \geq 1$

Here, we show that by allowing the selection of k elements from A, the probability of maximizing the expected Borda rank, increases exponentially.

Theorem 20. *Let $\pi^1, \ldots, \pi^{\log n}$ be a a set of $\log n$ sample permutations, randomly drawn from distribution $D(\hat{\pi}, \phi)$. And let $\overline{\pi}$ be their average ranking. Then*

$$Pr[\overline{\pi}(a_i) > \log n + i : \forall i \in [k]] < n^{-O(k)} \tag{6.8}$$

Proof. Let π be a permutation over A such that for all $i \in [k]$, $\pi(a_i) \geq \log n + i$. Then consider the i'th element a in π. The number of pairwise inversions that exist in π w.r.t it are at least $\log n$, by our assumption that $\hat{\pi}(a) > \log n + i$. Then by definition of the distribution, the probability of sampling such a permutation π is at most $\frac{Z_{m-k} \prod_{i=1}^{k} \phi^{\log n}}{Z_m} \leq \phi^{k \cdot \log n} = n^{-O(k)}$ □

Note that the above theorem needs to be complemented with an upper bound on the gap between the reference ranking position of and the maximum-likelihood of each candidate. However, we can easily get this by sampling $r = \log n$ permutations and applying Theorem 18, which gives a maximal $O(1)$ gap between the maximum-likelihood position and the reference rank, for any element in A, with polynomially (in n) small probability. I do believe however, that the polynomially small probability of an error could be shown to be in fact exponentially small in k (i.e. $n^{-O(k)}$).

7 Conclusions and Future Directions

We have given two methods for choosing the (approximately) best candidate in two natural and standard settings for the online choice problem at hand. As we have demonstrated, even with just a budget of 1, one could obtain very good results in the Mallows model, with a relatively small sample set. In the adversarial setting, we have shown that with a relatively small sample set (albeit not logarithmically small) one can approximate the optimal choice of a candidate with up to an $o(1)$ multiplicative error, with high probability. More importantly, we have shown that by taking a sampling approach we can approximate the social optimum, whenever the voting rule is a positional scoring function. As a result, this gives a useful tool when moving to a utilitarian setting.

One direction for future investigation would be to improve the rate at which the regret vanishes as n grows, both in the distributional setting as well as in the adversarial setting. Another direction that our study raises is the study of more involved constraints. In particular, we believe that if the alternatives have associated costs, then one could extend our work to cases in which there is a knapsack constraint. In terms of the our original example, we could imagine that there are costs attributed to the construction of the manufacturing lines. Furthermore, we can imagine that there are unit costs for producing copies of the alternatives in their production lines. More precisely, the decision maker will pay an initial price t_a for setting up the production line for alternative a, as well as an additional price of l_a for manufacturing a copy of alternative a for each agent who selects it.

The majority of our work in this paper deals with voting rules that are based on positional scoring functions, and we have shown how to extend our approaches to settings in which there are underlying utilities that induce the agent preferences. However, it will be interesting to consider settings where the voting rule is based on a non-positional scoring function.

Also, one could lift the constraint that requires the decision to be irrevocable, i.e. once an item is added, it cannot be replaced by another item. In this case, one could observe that such a model resembles the online learning setting (e.g. [17]). Alternatively, as previously studied in [1] for related settings, we can consider a setting in which the irrevocability of the decisions is relaxed. Specifically, we would like to consider the case that the decision maker is allowed to remove alternatives from the slate at a cost.

We could also extend our work by considering cases in which the agents can strategically delay their arrival, so as to increase their payoffs due to having a larger set of selected alternatives. Clearly, the pure sampling approach we have taken in this paper would be problematic, as none of the agents would like to take part in the initial sampling of preferences, and would thus delay their arrival in order to avoid it. Also, this scenario may tie-in with the previous extension, so that the agents, who can delay their arrivals, will be somewhat discouraged to do so due to a more powerful algorithm.

Acknowledgements

We would like to thank Ariel Procaccia for his many helpful comments and for pointing out the connection between this work and the results on low-distortion choice rules in [3]. We would also like to thank Craig Boutilier for many fruitful discussions.

References

[1] Moshe Babaioff, Jason D. Hartline, and Robert D. Kleinberg. Selling ad campaigns: online algorithms with cancellations. In *ACM Conference on Electronic Commerce*, pages 61–70, 2009.

[2] Moshe Babaioff, Nicole Immorlica, David Kempe, and Robert Kleinberg. Online auctions and generalized secretary problems. *SIGecom Exchanges*, 7(2), 2008.

[3] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *To appear on Proc. 13th ACM Conference on Electronic Commerce, Jun 2012*, 2012.

[4] Mark Braverman and Elchanan Mossel. Noisy sorting without resampling. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '08, pages 268–276, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[5] John R. Chamberlin and Paul N. Courant. Representative deliberations and representative decisions: Proportional representation and the borda rule. *The American Political Science Review*, 77(3):pp. 718–733, 1983.

[6] Jean-Paul Doignon, Aleksandar Pekeč, and Michel Regenwetter. The repeated insertion model for rankings: Missing link between two subset choice models. *Psychometrika*, 69:33–54, 2004. 10.1007/BF02295838.

[7] Eyal Even-Dar, Robert Kleinberg, Shie Mannor, and Yishay Mansour. Online learning for global cost functions. In *COLT*, 2009.

[8] M. A. Fligner and J. S. Verducci. Distance based ranking models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):pp. 359–369, 1986.

[9] Chinmay Karande, Aranyak Mehta, and Pushkar Tripathi. Online bipartite matching with unknown distributions. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 587–596, New York, NY, USA, 2011. ACM.

[10] Richard M. Karp, Umesh V. Vazirani, and Vijay V. Vazirani. An optimal algorithm for on-line bipartite matching. In *STOC*, pages 352–358, 1990.

[11] Robert D. Kleinberg. A multiple-choice secretary algorithm with applications to online auctions. In *SODA*, pages 630–631, 2005.

[12] Tyler Lu and Craig Boutilier. Budgeted social choice: From consensus to personalized decision making. In *IJCAI*, pages 280–286, 2011.

[13] Mohammad Mahdian and Qiqi Yan. Online bipartite matching with random arrivals: an approach based on strongly factor-revealing lps. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 597–606, New York, NY, USA, 2011. ACM.

[14] C. L. Mallows. Non-null ranking models. *Biometrika*, 44(1/2):pp. 114–130, 1957.

[15] Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *The Tenth International Workshop on Cooperative Information Agents (CIA)*, 2006.

[16] Moshe Tennenholtz. Transitive voting. In *Proceedings of the 5th ACM conference on Electronic commerce*, EC '04, pages 230–231, New York, NY, USA, 2004. ACM.

[17] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

Brendan Lucier
Microsoft Research, New England
Email: `brlucier@microsoft.edu`

Joel Oren
Department of Computer Science,
University of Toronto
Email: `oren@cs.toronto.edu`