Uniformly Bounded Regret in Dynamic Fair Allocation

Santiago R. Balseiro, Shangzhou Xia

Graduate School of Business, Columbia University, New York, NY 10027, srb2155@columbia.edu, sx2182@columbia.edu

Extended Abstract[1]

Fair resource allocation has been widely studied in economics and operations research, with applications in school seat allocation, work visa lotteries, affordable housing allocation, machine load balancing, etc. Involving multiple agents stipulates fairness as a crucial concern to the central planner. To this end, allocations are often evaluated on a balance between efficiency and fairness across the agents. Common fairness criteria studied in the literature of fair allocation include Pareto efficiency, envy-freeness along with its variants, and proportional fairness.

Cardinal welfare metrics incorporate efficiency and fairness concerns of allocations by mapping agents' individual utilities to a real-valued numeric representation of collective welfare. These metrics are wieldy in operations research and computer science because they allow to cast allocation problems as optimization problems, and different allocations may be compared against one another, yielding a quantifiable welfare difference. One particular cardinal welfare metric is the egalitarian welfare, based on the theory of distributive justice by Rawls (2001), whose difference principle maximizes the welfare of the worst-off group of agents. The egalitarian welfare objective defines a max-min welfare optimization problem, whose indivisible variant is dubbed the Santa Claus problem. Another well-studied cardinal example is the Nash social welfare, mathematically the geometric mean of the agents' utilities. Optimal allocations under the Nash social welfare are invariant to scaling of any single agent, for it is homogeneous on each agent's utility. In this work, we consider a broad parameterized collection of welfare metrics, which subsumes the two examples above, called the *Hölder means* or q-th power means.

Allocation problems in real life, however, often concern dynamic and irrevocable decisions where either agents or resources arrive in a sequential manner. Most prior works on dynamic resource allocation aim to maximize the utilitarian welfare, i.e., the efficiency of the allocation, which may result in unfair concentration of resources on certain high-utility agents while leaving others' demands under-fulfilled.

This work focuses on the specific variant of allocating resources arriving sequentially to heterogeneous agents, which is particularly relevant when resources are perishable or urgently needed and require immediate allocation decisions. Applications include allocating donated organs to hospitals, donated food to local charities, and online advertisement slots to advertisers, etc. Our objective is to design computationally efficient dynamic policies that maximize the overall welfare.

Contributions In this work, we consider the dynamic fair allocation problem, where T resources arriving sequentially are to be immediately and irrevocably allocated to n agents with linear utilities. The joint marginal utilities of each resource to the agents, drawn independently from a finite joint distribution known to the central planner, which are revealed before allocation decisions are made. An optimal online policy could in principle be computed using dynamic programming but, because of the so-called *curse of dimensionality*, solving the problem to optimality is impractical when the time horizon is large. We therefore seek to design policies that are computationally efficient and have provable performance guarantees in terms of the welfare loss against that of the hindsight optimal solution OPT, i.e., the optimal allocation if all marginal utilities were known in advance of deciding on allocations.

We first consider a static policy that is obtained by solving a fluid relaxation in which all stochastic quantities are replaced by their mean values. Simple as it is, we show under all smooth

[1] The full version of the paper can be found here.

Hölder-mean welfare metrics that are smooth, it attains a resounding $O(1)$ regret over any T against the hindsight optimum. The sole non-smooth exception is the egalitarian welfare objective, under which fluid policies attain a regret on the order of $\Theta(\sqrt{T})$. We denote by FLU the optimum of the fluid relaxation problem and by F the corresponding fluid policy.

Our next main contribution is proposing the Backward Infrequent Re-solving with Thresholding (BIRT) policy for dynamic fair allocation under the egalitarian welfare. Our policy is inspired by Bumpensanti and Wang (2020), who designed similar policies for utilitarian objectives based on fluid policies, and consists of updating and re-solving the fluid problem at most $O(\log\log T)$ times and thresholding the resulting solution before deciding subsequent allocations. Utilitarian objectives studied by prior works on dynamic resource allocation are simpler to handle because they are separable over time, i.e., the total utility is simply the sum of the utilities of each time period. By contrast, the egalitarian objective discussed in our paper is non-linear and not time separable, and so requires a different analysis. We provide a novel analysis to show that the BIRT policy achieves $O(1)$ regret against the hindsight optimum, namely uniformly bounded over the time horizon length T. The uniformly bounded regret guarantee implies that the BIRT policy performs almost comparably to a clairvoyant who foresees all arrivals in advance of acting. A salient feature of the BIRT policy is that its $O(1)$ regret guarantee is insensitive to whether or not the original fluid problem is degenerate (more technically, that any fluid policy is nondegenerate) or nearly degenerate. This is in drastic contrast to previous works on online stochastic optimization whose similar $O(1)$-loss results rely heavily on non-degeneracy conditions on the underlying fluid problem.

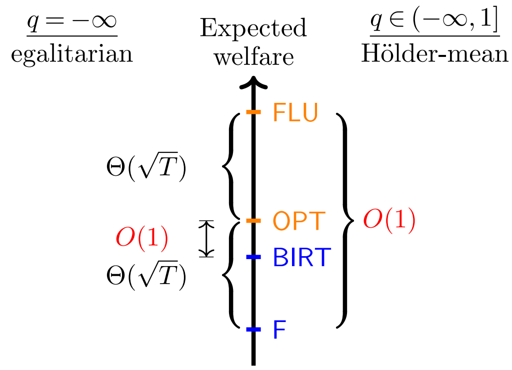

Figure 1 Summary of benchmark values and policy performances. Results relevant to the egalitarian welfare are shown on the left, and that relevant to all other Hölder-mean welfare metrics is shown on the right.

We summarize findings on the benchmarks (FLU and OPT) and policy performances (BIRT and F) on the vertical axis in Figure 1 with respective suprema of the differences (i.e., worst-case arrival distribution) between them. All in all, we provide computationally efficient online policies that attain uniformly bounded regret in for all Hölder-mean welfare metrics. We corroborate our theoretical results with simulation experiments and, in particular, illustrate the sufficiency of infrequent re-solving and the necessity of thresholding in guaranteeing an $O(1)$ regret under the egalitarian welfare.

References

Bumpensanti P, Wang H (2020) A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* 66(7):2993–3009.

Rawls J (2001) *Justice as fairness: A restatement* (Harvard University Press).