Enabling Long-term Fairness in Dynamic Resource Allocation
Tareq Si Salem, Georgios Iosifidis, Giovanni Neglia

To cite this version:

HAL Id: hal-04386354
https://inria.hal.science/hal-04386354
Submitted on 10 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Enabling Long-term Fairness in Dynamic Resource Allocation

Tareq Si Salem
tareq.si-salem@inria.fr
Inria
Sophia Antipolis, France

George Iosifidis
G.Iosifidis@tudelft.nl
TU Delft
Delft, The Netherlands

Giovanni Neglia
giovanni.neglia@inria.fr
Inria
Sophia Antipolis, France

ABSTRACT

We study the fairness of dynamic resource allocation problem under the $\alpha$-fairness criterion. We recognize two different fairness objectives that naturally arise in this problem: the well-understood slot-fairness objective that aims to ensure fairness at every timeslot, and the less explored horizon-fairness objective that aims to ensure fairness across utilities accumulated over a time horizon. We argue that horizon-fairness comes at a lower price in terms of social welfare. We study horizon-fairness with the regret as a performance metric and show that vanishing regret cannot be achieved in presence of an unrestricted adversary. We propose restrictions on the adversary’s capabilities corresponding to realistic scenarios and an online policy that indeed guarantees vanishing regret under these restrictions.

CCS CONCEPTS

• Applied computing → Multi-criterion optimization and decision-making; • Theory of computation → Online learning algorithms; Algorithmic game theory.

KEYWORDS

Online Learning, Multi-timescale Fairness, Axiomatic Bargaining, Dynamic Resource Allocation

ACM Reference Format:

1 INTRODUCTION

Achieving fairness when allocating resources in communication and computing systems has been a subject of extensive research, and has been successfully applied in numerous practical problems. Fairness is leveraged to perform congestion control in the Internet [6, 13], to select transmission power in multi-user wireless networks [17, 22], and to allocate multidimensional resources in cloud computing platforms [3, 20, 21]. Depending on the problem at hand, the criterion of fairness can be expressed in terms of how the service performance is distributed across the end-users, or in terms of how the costs are balanced across the servicing nodes. A prevalent fairness metric is $\alpha$-fairness, which encompasses the utilitarian principle (Bentham-Edgeworth solution [5]), proportional fairness (Nash bargaining solution [15]), max-min fairness (Kalai–Smorodinsky bargaining solution [10]). All these fairness metrics have been used in different cases for the design of resource management mechanisms [14, 16].

A common limitation of the above works is that they consider static environments. That is, the resources to be allocated and, importantly, the users’ utility functions, are fixed and known to the decision maker. This assumption is very often unrealistic for today’s communication and computing systems.

2 MAIN RESULTS

This paper makes the next step towards enabling long-term fairness in dynamic systems. We consider a system that serves a set of agents $I$, where a controller selects at each timeslot $t \in \mathbb{N}$ a resource allocation profile $x_t$ from a set of eligible allocations $X$ based on past agents’ utility functions $u_t^i : X \to \mathbb{R}$ for $t < t$ and of $\alpha$-fairness function $F_\alpha : \mathbb{R}_{\geq 0}^I \to \mathbb{R}$. The utilities might change due to unknown, unpredictable, and (possibly) non-stationary perturbations that are revealed to the controller only after it decides $x_t$.

We employ the terms horizon-fairness (HF) and slot-fairness (SF) to distinguish the different ways fairness can be enforced in a such time-slotted dynamic system (see the illustration in Fig. 1). Under horizon-fairness, the controller enforces fairness on the aggregate utilities for a given time horizon $T$, whereas under slot-fairness, it enforces fairness on the utilities at each timeslot separately. Both metrics have been studied in previous work, e.g., see [7, 9, 11, 19]. Our focus is on horizon-fairness, which raises novel technical challenges and subsumes slot-fairness as a special case.

We design the online horizon-fair (O HF) policy by leveraging online convex optimization (OCO) [8], to handle this reduced information setting under a powerful adversarial perturbation model. Adversarial analysis is a modeling technique to characterize a system’s

Figure 1: The different ways fairness can be enforced in a time-slotted dynamic system. The decision maker can either consider the $\alpha$-fairness objective $F_\alpha$ at every timeslot (slot-fairness), or at the end of the time horizon $T$ (horizon-fairness).
performance under unknown and hard-to-characterized exogenous parameters. In our context, the performance of a resource allocation policy \( \mathcal{A} \) is evaluated by the fairness regret, which is defined as the difference between the \( \alpha \)-fairness, over the time-averaged utilities, achieved by a static optimum-in-hindsight (benchmark) and the one achieved by the policy:

\[
\mathcal{R}_T \left( F_\alpha, \mathcal{A} \right) \triangleq \sup_{\{u_t\}_{t \in [T]} \in \mathcal{U}} \left( \max \limits_{x \in \mathcal{X}} \frac{1}{T} \sum_{t \in [T]} u_t(x) - F_\alpha \left( \frac{1}{T} \sum_{t \in [T]} u_t(x_t) \right) \right). \tag{1}
\]

If the fairness regret vanishes over time (i.e., \( \lim_{T \to \infty} \mathcal{R}_T \left( F_\alpha, \mathcal{A} \right) = 0 \)), policy \( \mathcal{A} \) will attain the same fairness value as the static benchmark under any possible sequence of utility functions. A policy that achieves sublinear regret under these adversarial conditions, can also succeed in more benign conditions where the perturbations are not adversarial, or the utility functions are revealed at the beginning of each slot.

The fairness regret metric (1) departs from the template of OCO. In particular, the scalarization of the vector-valued utilities, through the \( \alpha \)-fairness function, is not applied at every timeslot to allow for the controller to easily adapt its allocations, instead is only applied at the end of the time horizon \( T \). Our first result characterizes the challenges in tackling this learning problem. Namely, we prove that when utility perturbations are only subject to four mild technical conditions, such as in standard OCO, it is impossible to achieve vanishing fairness-regret. Similar negative results were obtained under different setups of primal-dual learning and online saddle point learning [1, 12, 18], but they have been devised for specific problem structures (e.g., online matrix games) and thus do not apply to our setting.

In light of this negative result, we introduce additional necessary conditions on the adversary to obtain a vanishing regret guarantee. Namely, the adversary can only induce perturbations to the time-averaged utilities we call budgeted-severity or partitioned-severity constrained. These conditions capture several practical utility patterns, such as non-stationary corruptions, ergodic and periodic inputs [2, 4, 11, 23]. We proceed to propose the OHF policy which adapts dynamically the allocation decisions and provably achieves \( \mathcal{R}_T \left( F_\alpha, \mathcal{A} \right) = O(1) \).

The OHF policy employs a novel learning approach that operates concurrently, and in a synchronized fashion, in a primal and a dual (conjugate) space. Intuitively, OHF learns the weighted time-varying utilities in a primal space, and learns the weights accounting for the global fairness metric in some dual space. To achieve this, we develop novel techniques through a convex conjugate approach.

Finally, we apply our fairness framework to a representative resource management problem in virtualized caching systems where different caches cooperate by serving jointly the received content requests. We evaluate the performance of OHF with its slot-fairness counterpart policy through numerical examples. We evaluate the price of fairness of OHF, which quantifies the efficiency loss due to fairness, across different network topologies and participating agents. Lastly, we apply OHF to a Nash bargaining scenario, a concept that has been widely used in resource allocation to distribute to a set of agents the utility of their cooperation.

### 3 CONCLUSION AND FUTURE WORK

In this work, we proposed a novel OHF policy that achieves horizon-fairness in dynamic resource allocation problems. Our work paves the road for several interesting next steps. A future research direction is to consider decentralized versions of the policy under which each agent selects an allocation with limited information exchange across agents. Another important future research direction is to bridge the horizon-fairness and slot-fairness criteria to target applications where the agents are interested in ensuring fairness within a target time window. A final interesting research direction is to consider a limited feedback scenario where only part of the utility is revealed to the agents (e.g., bandit feedback). Our policy could be extended to this setting through gradient estimation techniques [8].

### REFERENCES