IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 31, NO. 5, OCTOBER 2023

Optimal Routing to Parallel Servers With Unknown Utilities—Multi-Armed Bandit With Queues

Xinzhe Fu and Eytan Modiano, *Fellow, IEEE*

***Abstract*—We consider the optimal routing problem in a discrete-time system with a job dispatcher connected to M parallel servers. At every time slot, the job dispatcher sends the incoming jobs to a server for execution, with each server having a queue that stores the jobs. The arrival process of incoming jobs, and the service processes of the servers are stochastic with unknown and possibly heterogeneous rates. Each server s_m is associated with an underlying utility v_m that is initially unknown. Whenever server s_m completes a job, a utility of v_m is obtained and a noisy observation of v_m is received. The goal is to design a policy that makes routing decisions to maximize the total utility obtained by the end of a finite time horizon T. The performance of policies is measured in terms of regret, which is the additive difference between the expected total utility obtained by the policy and the supremum of the expected total utility over all the policies. The optimal routing problem can be interpreted as a problem of multi-armed bandit with queues where each server is viewed as an arm and the**

I. INTRODUCTION

CONSIDER a system consisting of a job dispatcher and parallel servers. Incoming jobs arrive at the job dispatcher and get immediately routed to a server where they get queued up for execution. Such system model captures a wide range of applications in communication networks [1], [2], production lines [3], and web server farms [4], and there has been extensive research on the routing problem under this model. Previous works have proposed and analyzed routing policies that aim at optimizing delays [5], minimizing holding costs [6], or achieving desirable load balancing properties [7], [8].

In this paper, we consider the routing problem for parallel servers from the perspective of optimizing system utility. We study the setting where a certain utility is obtained when