Accepted Manuscript

Max-weight scheduling across multiple timescales

Mark S. Squillante, Peter M. van de Ven

PII: S0167-6377(17)30003-2
DOI: https://doi.org/10.1016/j.orl.2018.01.012
Reference: OPERES 6335

To appear in: *Operations Research Letters*

Received date : 2 January 2017
Revised date : 1 September 2017
Accepted date : 30 January 2018

Please cite this article as: M.S. Squillante, P.M. van de Ven, Max-weight scheduling across multiple timescales, *Operations Research Letters* (2018), https://doi.org/10.1016/j.orl.2018.01.012

Max-Weight Scheduling across Multiple Timescales

Mark S. Squillante[a], Peter M. van de Ven[b,*]

[a]Mathematical Sciences Department, IBM Research
[b]Centrum Wiskunde & Informatica (CWI)

Abstract

Many systems consist of a mixture of various resource types that together support better performance relative to those with a single resource type. One important characteristic of these systems is the fact that the various comprising resource types can operate on different timescales, implying that the corresponding control decisions are not made simultaneously. To address the resulting scheduling problem, we present and analyze two variants of max-weight scheduling that are designed to deal with the different timescales of such systems.

Keywords: Multiple timescale systems, Hybrid networks, Markov processes, Scheduling, Stability

1. Introduction

Present-day systems increasingly include adaptive infrastructures that can involve multiple types of resources. Examples include dual-sourcing inventory systems [1] (e.g., one supplier faster than another supplier), communication networks with noisy channels [2] (e.g., time variations in sources and channels), workforce management [3] (e.g., more experienced employees faster than less experienced employees), smart power grids [4] (e.g., time variations in load and generation with renewable energy), wireless communication with energy harvesting devices [5] (e.g., time variations in wireless transmission and energy harvesting), and hybrid communication networks [6]. In the latter case of hybrid communication networks, which served as the original motivation for the present study, we observe that present-day communication networks increasingly consist of an adaptive infrastructure that may encompass multiple instances of mobile networks (e.g., wireless and cellular networks, public Internet, and various private intranets) together with a wide range of communication technologies (e.g., 3G and 4G cellular, wired and satellite networks) [7]. Many of these systems exploit and interoperate with a diversity of resource types in order to increase both capacity and robustness.

One overlooked property in the literature relevant to systems involving multiple types of resources concerns the operation of the comprising resource types at different timescales. To address this and related issues, there is an important need for mathematical frameworks that support the design and adaptive control of these systems. Such a framework includes fundamental performance limits and how these limits impact the design of adaptive control algorithms, convergence of these algorithms in a multiple timescale environment, and assurance of good performance under these algorithms. Our present study seeks to establish fundamental properties of adaptive control policies for the optimal scheduling across various resource types operating at different timescales that ensures low delays and high throughput.

We consider a general system that includes a mixture of various types of resources that operate at different timescales. To elucidate the exposition of our analysis, we focus on a system environment comprising two resource types with one working at a faster timescale than the other. For example, in the context of hybrid networks, the types of resources could be based on satellite and wireless networks where the former operates at a slower timescale than the latter [8], due to its larger propagation delay and round trip time. We note that our analysis extends in a straightforward manner to support an arbitrary number of resource types working at different timescales.

Given the popularity of the max-weight scheduling algorithm [9] in operations research and related fields, we restrict our attention to this class of adaptive control policies. Our goal is to tailor this scheduling algorithm to system environments with timescale differences and to derive its stability, throughput and delay properties. While these fundamental performance properties of the max-weight scheduling algorithm in single-timescale systems have received a great deal of attention in the research literature (see, e.g., [10, 11]), no previous work to our knowledge has studied the performance properties of max-weight scheduling in a system environment that comprises multiple types of resources operating at different timescales.

We speak of a fast resource and a slow resource, where the allocation and feasible service rate over the slow resource remains fixed for a long period of time relative to that of the fast resource. This suggests an interesting new class of scheduling problems in which the longer-term commitment of the slow resource requires more careful scheduling of both resources than would otherwise be the case, possibly taking into account future service on the slow resource when deciding on the usage of the fast resource. Our study, for the first time, explores and models this scheduling problem involving different timescales, and provides initial results on how to handle such environments. We

*Corresponding author (ven@cwi.nl)