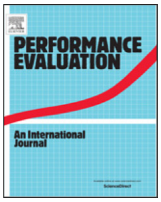
Performance Evaluation 121–122 (2018) 18–37

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva



Reduced-complexity delay-efficient throughput-optimal distributed scheduling with heterogeneously delayed network-state information

Srinath Narasimha [*], Joy Kuri, Albert Sunny

Department of Electronic Systems Engineering, Indian Institute of Science, Bengaluru 560012, India

ARTICLE INFO

Article history:
Received 25 March 2017
Received in revised form 16 December 2017
Accepted 17 December 2017
Available online 23 March 2018

Keywords:
Distributed scheduling
Heterogeneous delays
Network state information
Throughput optimality

ABSTRACT

We consider the problem of distributed scheduling in wireless communication networks where heterogeneously delayed queue lengths and channel states of all links are available at all the transmitters. In an earlier work (by Reddy et al. in *Queueing Systems*, 2012), a throughput-optimal scheduling policy (which we refer to henceforth as the R policy) for this setting was proposed. We study the R policy, and examine its two drawbacks — (i) its huge computational complexity, and (ii) its non-optimal average per-packet queueing delay. We show that the R policy unnecessarily constrains itself to work with information that is more delayed than that afforded by the system. We propose a new distributed scheduling policy that fully exploits the common state information available to all transmitters, thereby greatly improving upon the computational complexity and the delay performance relative to those of the R policy. We also establish the throughput optimality of our policy analytically. We evaluate the performance of the proposed policy and validate our analytical results through extensive numerical simulation. Thus, our work enlarges the ambit of networks for which throughput-optimal scheduling is practicable.

1. Introduction

Scheduling is a central problem in multiple-access wireless communication networks, where the objective is to schedule link transmissions so as to optimize some desired metric (e.g., maximize the system throughput) in the presence of challenges that are unique to the wireless medium — namely, channel fading and interference due to transmissions from other users in the network. This problem has been studied extensively in the literature. A highly influential and often cited work in this area is the work by Tassiulas and Ephremides [1], who proposed the *Back-Pressure* scheduling algorithm (a version of the *Max-Weight* algorithm [2,3]), which is a centralized algorithm that schedules the links in the network based on global knowledge of the instantaneous queue lengths at all the links. Even though this algorithm is provably throughput-optimal (see Section 2.5), it is a centralized algorithm that requires solving a global optimization problem in each time slot, and it also requires knowledge of instantaneous queue lengths at all links in the network to determine the schedule [1,4].

The Max-Weight algorithm, being a centralized policy, involves the computationally prohibitive task of finding a maximal independent (i.e., non-interfering) set of links that can be activated simultaneously and whose summation of link-weights is maximum [5]. To circumvent this limitation, two broad approaches have been considered in the literature [5] — namely, design of random-access algorithms in which access probabilities are dependent on queue sizes [5,6] or on arrival rates [7–12], and design of distributed implementations of the Max-Weight algorithm [13,14]. Some of these approaches require

---

[*] Corresponding author.
E-mail addresses: srinath@iisc.ac.in (S. Narasimha), kuri@iisc.ac.in (J. Kuri), albert.sunny@inria.fr (A. Sunny).

https://doi.org/10.1016/j.peva.2017.12.006
0166-5316/© 2018 Elsevier B.V. All rights reserved.

knowledge of instantaneous queue lengths or instantaneous channel states (or both) to attain their objective. Scheduling in multi-hop wireless networks that drop the packets that are not delivered to their intended destinations within their deadlines is considered in [15,16], and policies which not only maximize the network's timely-throughput (see [15]) but also ensure that the end-to-end network delays are bounded, are proposed. These works are extended in [17] to multi-hop networks with wireless interference constraints where a decentralized policy is proposed and is shown to outperform the timely-throughput of a version of the Q-CSMA policy [6] adapted to the problem of maximizing the network's timely-throughput.

Even though it may be reasonable to assume that each node has knowledge of instantaneous queue lengths and channel states (at all times) for its own links (i.e., links emanating from itself), it is less pragmatic to assume that any node possesses instantaneous information about any link in the network other than its own links (at any time instant). This could be because, for example, these quantities vary quickly with time (e.g., fast fading), or because the propagation delay of the feedback channel is large [18]. In [19], the authors consider networks where each node possesses knowledge of instantaneous queue lengths and channel states for its own links, but only has these information from other links in the network with some globally fixed delay (commonly referred to as *homogeneous delay*). The assumption of homogeneous delays, however, is not satisfied in many networks where there is often a mismatch in the delays with which each node can acquire information about queue lengths and channel states of other links in the network [4]. Non-homogeneous delays are commonly referred to as *heterogeneous delays*.

One serious issue in networks with heterogeneous delays is that the nodes could potentially have different views of the state of the network [4]. In [4], the authors consider distributed scheduling in a wireless network where information about the queue lengths and channel states of links in the network, available at the nodes, are heterogeneously delayed. They characterize the system throughput region for this setting, propose a scheduling policy, and show that their policy is throughput optimal. We start by studying the limitations of the policy proposed in [4].

1.1. Our contributions

As in [4], we consider the problem of distributed scheduling in wireless networks where delayed queue length and delayed channel state information of each wireless link in the network, are available at all the transmitters in the network, but with possibly different delays. We refer to these as heterogeneously delayed queue state information (QSI) and heterogeneously delayed channel state information (CSI) respectively, and collectively refer to them as heterogeneously delayed network state information (NSI). We refer the reader to Section 2.4 for detailed information about the structure of heterogeneously delayed NSI that we have considered in this work. Our contributions in this work are three-fold and are summarized below:

1. First, we study the limitations of the R policy and obtain analytical expressions for the general-case and worst-case computational complexities of the R policy, and show that the R policy is computationally very costly.
2. We obtain an analytical expression for the average per-packet queueing delay of the R policy and show that it is non-optimal, and that it can be improved upon significantly.
3. Next, we bring a new insight into this problem by showing that the structure of heterogeneously delayed NSI as defined in the system model in [4] affords each node access to NSI that is *less delayed* than that exploited there, and yet commonly available at all nodes in the network. We propose a new scheduling policy (henceforth referred to as the H policy) that incorporates this insight and show that the H policy has hugely improved computational complexity and substantially improved delay performance relative to those of the R policy. We establish that the H policy is throughput optimal.

2. System model and performance metrics

We use a slotted-time system and restrict our focus to single-hop transmissions. The notations we define in this section are similar to those in [4] and are summarized in Table 4.

2.1. Network model

Our model of the wireless network has L transmitter–receiver pairs (or links); the set of links is denoted $\mathcal{L}$. We abstract the channel condition on each of the wireless links by the link's capacity. We model the time-varying capacity of each link l as a separate discrete-time Markov chain (DTMC) denoted $\{C_l[t]\}$ on the same state space $\mathcal{C} = \{c_1, c_2, \ldots, c_M\}$, where $c_1 < c_2 < \cdots < c_M$ are non-negative integers.[1,2] The channel conditions on each wireless link are assumed to be independent of the conditions on the other links, but identically distributed, with transition probabilities $p_{ij} := \Pr\big[C_l[t+1] = c_j \mid C_l[t] = c_i\big]$. We assume that these DTMCs are irreducible and aperiodic, and therefore have a stationary distribution, with the stationary probability of being in state $c_j, j \in \{1, 2, \ldots, M\}$, denoted $\pi(c_j)$.

---

[1] We use $\mathcal{C}$ to denote both the set and its cardinality. Thus, as set, $\mathcal{C} = \{c_1, c_2, \ldots, c_M\}$, and as cardinality, $\mathcal{C} = M$.

[2] We remark that the above channel model is assumed for making notations simpler and to enhance clarity. Our results hold even for the case of networks where each link is modeled as a separate DTMC (with different state spaces and/or different transition probabilities).