

Decomposition of Large-Scale MDPs for Wireless Scheduling with Load- and Channel-Awareness

Huasen Wu^{*†}, Xiaojun Lin[‡], Xin Liu^{†§}, Kun Tan[†], and Yongguang Zhang[†]

^{*}School of Electronic and Information Engineering, Beihang University, Beijing 100191, China

[†]Microsoft Research Asia, Beijing 100080, China

[‡]School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA

[§]Department of Computer Science, University of California, Davis, CA 95616, USA

Abstract—Scheduling delay-tolerant tasks based on both load- and channel-awareness can significantly reduce the peak demand in cellular networks. However, solving the optimal scheduling problem leads to a large-scale Markov Decision Process (MDP) with extremely high complexity. In this work, we propose a scalable and distributed approach to this problem, called Coordinated Scheduling (CoSchd). CoSchd decomposes the large-scale MDP problem into many individual MDP problems, each of which can be solved independently by each user under a limited amount of coordination signal from the BS. We show that CoSchd is close to optimal when the number of users becomes large. Further, we propose an online version of CoSchd that iteratively updates the scheduling policy based on online measurements. Simulation results demonstrate that exploiting load- and channel-awareness with CoSchd can effectively alleviate cellular network congestion.

I. INTRODUCTION

A grand challenge facing today’s mobile service providers is to meet the exponentially increasing demand for mobile broadband services. This problem is particularly severe at the so-called “peak”, where the network is heavily loaded at specific times and locations. Currently, wireless providers invest heavily in new spectrum and infrastructure to accommodate the *peak* demand, but such efforts are costly and inefficient: since the network traffic at non-peak times is orders-of-magnitude lower than peaks, provisioning network capacity for peak demand will lead to poor utilization of network resources.

An alternative approach is to exploit the delay tolerance of mobile applications to improve the network utilization. Prior work has identified a class of applications that can tolerate some delay, ranging from a few minutes to hours [1–4]. For example, the analysis in [3] shows that more than 55% of multimedia contents in cellular networks are uploaded more than one day after their creation time. More recently, the survey conducted in the TUBE project indicates that users are willing to delay their data transmissions if appropriate incentives are provided, *i.e.* a discounted price [1]. Motivated by these findings, in this paper we study the scheduling of delay-tolerant traffic to minimize network congestion and improve resource utilization in wireless networks.

There are two directions where delay-tolerance can potentially be exploited to alleviate network congestion: *load-awareness* and *channel-awareness*. On one hand, approaches

such as TUBE [1,5] move delay-tolerant traffic to the time and location where the network is less loaded, *i.e.*, *being load-aware*, and thus alleviate network congestion. However, TUBE does not consider users’ time-varying wireless channels – hence we classify it as a “*load-only*” approach. On the other hand, noting the temporal variation of channel conditions in wireless networks, a number of channel-aware scheduling schemes have been proposed at the mobile device to improve spectrum efficiency [2,4,6]. While this line of work takes advantage of the opportunistic nature of wireless networks, it has been limited to optimizing on a single mobile device. As a result, these schemes are oblivious to traffic-load levels and thus we refer to them as “*channel-only*” approaches. The recent work in [7] proposes mobile-side mechanisms to estimate and react to both channel condition and network load. However, it mainly focuses on reducing the energy consumption of the mobile. To the best of our knowledge, the above two directions have not been investigated jointly for the purpose of reducing network congestion.

In this paper, we study joint load- and channel-aware scheduling policies for delay-tolerant traffic to reduce network congestion. We consider the scenario of a cellular network serving a sequence of data transfer requests. Each data transfer request has a pre-specified deadline, which is directly tied to the users’ overall experience. The network’s objective is to schedule such data transfers intelligently to minimize the network congestion cost, subject to the deadline constraints of the data transfers. We define the network congestion cost as the sum of (strictly-)convex functions of the load at each BS/WiFi-hotspot and at each time. With the strict-convexity, the cost function naturally penalizes high peak demand and thus a cost-minimizing solution will tend to smooth out the traffic load across time and location.

The above scheduling problem is a sequential decision problem and can theoretically be cast as a Markov Decision Process (MDP). However, solving such an MDP problem faces challenges of both computational complexity and information collection. First, as the system size increases, the complexity of the MDP problem increases exponentially due to the curse of dimensionality. Compared to the channel-only approach that only considers one mobile device [2,4], here the size of the problem is very large, as a typical network may have hundreds of thousands of requests and a large number of BSs and Wi-

Fi hotspots. Compared to the load-only approach [1], here the channel uncertainty leads to significant difficulty in determining the amount of load that can be moved under a given policy. Second, if we were to solve the MDP in a centralized manner, the scheduler needs to know all requests and channel evolution statistics for each individual user. Collecting this information may require the BS to track the behaviors of all users, which raises concerns on both signaling overhead and privacy. Thus, decomposition technique and distributed scheduling policies are highly desirable to effectively solve such a large-scale MDP.

In this paper, we propose distributed schemes for solving this type of large-scale MDP problems. We refer to our distributed solution as Coordinated Scheduling (CoSchd). Under CoSchd, the network does not need to know the statistics of all requests before hand, but instead updates a set of congestion signals based on the aggregated network load. At the same time, each user executes an individual decision policy based on the congestion signal and its own channel statistics. The key to this decomposition is to approximate the original problem by exchanging the order of the expectation and the cost function. Specifically, we replace the minimization of the expectation of the cost function with a minimization of the cost of the expected load (See Section III for details). This approximation allows us to apply duality to decompose the network control, which addresses both the complexity issue and the signaling/privacy issues discussed earlier. Under certain conditions, we show that as the number of users in the system tends to infinity, this decomposition is tight in the sense that it approaches to close to the optimal solution of the original problem. We further propose an online version of CoSchd that iteratively updates the scheduling policy based on online measurements. Finally, we have performed trace-driven simulations to evaluate the performance gains of exploiting load- and/or channel-awareness. Our simulation results demonstrate the asymptotic optimality of the proposed CoSchd and the benefits of scheduling with load- and channel-awareness.

To the best of our knowledge, this is the first unified framework that exploits both the load- and channel-fluctuations to alleviate network congestion and improve resource efficiency. Note that load-only and channel-only policies can also be derived from our framework. For example, we can obtain the load-only policy by using identical congestion signals at all time and locations. Moreover, the decomposition approach proposed in this paper could potentially be used to solve other types of large-scale MDP problems, where multiple agents are weakly coupled by sharing common resources.

II. SYSTEM MODEL

We start by considering one BS, where the proposed approach can also be generalized to include multiple BSs and WiFi-hotspots, as discussed in [8]. The problem stated here applies to both the uplink and downlink in cellular networks.

Assume that time is slotted and indexed by $t \in \{0, 1, \dots\}$. Let N be the number of time-slots in each day. A typical

time-slot length ranges from tens of seconds to a few minutes. Because of the large time scale, we assume that a data transfer request will be completed in one time-slot when the request is accepted, as in [1].

Data Traffic. In every day, a sequence of data transfer requests enter the network with user-specified deadlines. We use the words “user” and “request” interchangeably. The requests depart upon completion or deadline expiration.

Consider the scheduling problem in one day, where $t \in \{0, 1, \dots, N - 1\}$. Let $\mathcal{I} = \{1, 2, \dots, m\}$ be the index set of all users that may request transfers from the BS. For each user $i \in \mathcal{I}$, denote the arrival time and the file size of its request by A_i and B_i , respectively. Assume that A_i 's and B_i 's are i.i.d. across users. A_i follows a distribution that reflects the typical traffic pattern of the day [1, 9]. Note that a user may not request transfers every day and we let $A_i = N$ when user i does not request any transfers. We assume that the file size B_i is bounded [10] and is given as soon as the request arrives.

Each request i is associated with a user- or application-specific deadline D_i , i.e., the maximum delay that a user can tolerate. The deadline ranges from minutes to hours for delay-tolerant traffic [1, 3]. Such a deadline requirement depends on specific applications and can be set in various ways. For example, it could be a default setting in an application, e.g., syncing emails every half an hour; or, it can be learned from user preference. We assume that all transmission tasks should be completed at the end of the day, i.e., $A_i + D_i \leq N - 1$, for simplicity. To guarantee the quality of user experience, we need to constrain the deadline violation probability when scheduling delay-tolerant traffic, as will be discussed later. Note that in this model we also allow real-time traffic that needs to be transmitted immediately, in which case the deadline is set to be zero.

Channel Dynamics. Each user experiences time-varying network availability and channel conditions. This is captured by a stochastic process $R_i(t)$ ($t \in \{0, 1, \dots\}$), where $R_i(t) \geq 0$ denotes the instantaneous rate per unit spectrum resource (e.g., a time-frequency block in LTE) at which the BS can communicate with user i in time-slot t . We assume that $R_i(t)$ is independent across users and model $R_i(t)$ as a homogeneous Markov chain over a finite set of the possible transmission rates, i.e., $R_i(t) \in \{r_1, r_2, \dots, r_J\}$, where J is the number of possible rates, and $0 = r_1 < r_2 < \dots < r_J$. For user i , the transition probability matrix is given by

$$\mathbf{P}_i = [p_{j_1 j_2}^{(i)}]_{J \times J}, i \in \mathcal{I}, \quad (1)$$

where $p_{j_1 j_2}^{(i)} \in [0, 1]$, $1 \leq j_1, j_2 \leq J$, is the transition probability from state j_1 to state j_2 for user i . We assume that all channel processes have reached the steady state, i.e., with the stationary distribution $\boldsymbol{\pi}^{(i)}$, where $\boldsymbol{\pi}^{(i)} = [\pi_1^{(i)}, \pi_2^{(i)}, \dots, \pi_J^{(i)}]$ is the stationary distribution for the Markov chain of user i .

When user i in channel condition $R_i(t)$ ($R_i(t) > 0$) is scheduled to transmit a file of size B_i , it consumes $B_i/R_i(t)$ units of spectrum resource. We assume that each user can estimate its current channel condition via measurements of received signal strength and interference levels. Further, the

user can learn the transition probability of its channel dynamics based on historical measurements, as in [6, 11, 12].

Scheduling Policy and Base-Station Load. Let Γ denote a general scheduling policy that decides which users to transmit at a given time-slot. We consider the set of all causal policies. Corresponding to each Γ , we let $L_t(\Gamma)$ be the aggregate amount of spectrum resource consumed by the users transmitting in time-slot t under policy Γ . We express $L_t(\Gamma)$ as

$$L_t(\Gamma) = \sum_{i \in \mathcal{I}} Y_{i,t}(\Gamma), \quad t = 0, 1, \dots, N-1, \quad (2)$$

where $Y_{i,t}(\Gamma)$ is the amount of resource consumed by user i in time-slot t . More precisely,

$$Y_{i,t}(\Gamma) = \begin{cases} B_i/R_i(t), & \text{if user } i \text{ transmits in slot } t, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Objective. From the network's point of view, the objective is to minimize the total congestion cost in the horizon of N time-slots subject to the deadline violation constraints. Let $f(\cdot)$ be a strictly-convex congestion-cost function and $v_i(\Gamma)$ be the deadline violation probability of user i . The scheduling problem is then

$$(\mathcal{P}_0) \quad \begin{aligned} & \underset{\Gamma}{\text{minimize}} && F = \sum_{t=0}^{N-1} \mathbb{E}[f(L_t(\Gamma))], \\ & \text{subject to} && v_i(\Gamma) \leq \eta_i, \quad \forall i \in \mathcal{I}, \end{aligned} \quad (4)$$

where η_i is the maximum deadline violation probability tolerated by user i .

In problem \mathcal{P}_0 , the convexity of $f(\cdot)$ penalizes peaks and thus favors load that is smoothed over time, which is desirable for network operators. In our numerical results, we use the following function $f(l) = (l/\bar{C})^\nu$, where \bar{C} is a positive constant and $\nu > 1$ is a factor for controlling the penalty.

Note that in principle, \mathcal{P}_0 can be viewed as an MDP by taking the waiting time and channel condition of all users as system state. However, solving such an MDP problem in a centralized manner is forbiddingly complex. First, the size of the problem is very large, as a typical network may have hundreds of thousands of users, over a time horizon of a day. In addition, deadline constraint is notoriously difficult to solve in general because of the resource coupling across time and among users. Second, the problem formulation assumes knowledge of all jobs and their detailed channel information. In practice, it is not feasible to gather such detailed information in a central entity because of both signaling overhead and privacy concerns. Next, we will focus on the regime where the number of users is large, and develop a distributed approach for (approximately) solving problem \mathcal{P}_0 . Our main intuition is the following. In our system, each user can be seen as interacting with the set of all other users. When the number of users is large, the impact of any given user's decision on the overall system should be minimal. Thus, it would be as if each user is interacting with a common entity that includes all users in the system. If we can summarize the effect of all users by some kind of "congestion signal," we may then

be able to approximate the original system by one where each user independently reacts to such a common congestion signal. The challenges are how to design such a common congestion signal and how to establish the (asymptotic) optimality of the decomposition, which will be the focus of the following section.

III. ASYMPTOTICALLY OPTIMAL DECOMPOSITION

This section studies asymptotically optimal policies for solving the large scale MDP \mathcal{P}_0 . Note that the objective in (4) is to minimize the *expectation of total cost*. We first propose a lower bound of \mathcal{P}_0 by introducing a new problem \mathcal{P}_1 that minimizes the *total cost of expectation*. We show the asymptotic optimality of the decomposition approach in the many-source regime and a distributed implementation of the approach, referred to as Coordinated Scheduling (CoSchd).

A. Lower Bound

In the original MDP problem \mathcal{P}_0 , the cost is a function of the instantaneous load level $L_t(\Gamma)$ and the objective is to minimize the expected total cost. Because the cost function $f(\cdot)$ is convex, the optimal value of \mathcal{P}_0 can be lower bounded by exchanging the order of the expectation and the cost function. Specifically, consider the following problem that minimizes the total cost of the expected load level:

$$(\mathcal{P}_1) \quad \begin{aligned} & \underset{\Gamma}{\text{minimize}} && \tilde{F} = \sum_{t=0}^{N-1} f(l_t(\Gamma)), \\ & \text{subject to} && v_i(\Gamma) \leq \eta_i, \forall i \in \mathcal{I}, \end{aligned}$$

where $l_t(\Gamma) = \mathbb{E}[L_t(\Gamma)]$ is the expected value of load level. Let F^* be the optimal value of the original problem \mathcal{P}_0 and let F_{lb} be the optimal value of \mathcal{P}_1 . Because the constraints of \mathcal{P}_0 and \mathcal{P}_1 are identical and the only difference lies in the objective function, we can easily show the following proposition by the convexity of $f(\cdot)$ and Jensen's inequality [13].

Proposition 1 *The optimal value of problem \mathcal{P}_1 provides a lower bound on the value of the original problem \mathcal{P}_0 , i.e., $F_{\text{lb}} \leq F^*$.*

As we will see later, thanks to the linearity of expectation, the cost of expected load is much easier to deal with than the expectation of cost. Hence, problem \mathcal{P}_1 and its lower-bound property are critical in the design and analysis of asymptotically optimal policies. Next, we will study the optimal solution for \mathcal{P}_1 , and show its asymptotic optimality for the original problem \mathcal{P}_0 in the many-source regime.

B. Dual Decomposition

This subsection proposes a decomposition approach for solving problem \mathcal{P}_1 based on dual decomposition. To use dual decomposition, we first introduce auxiliary variables $h_t \geq 0$ ($t = 0, 1, \dots, N-1$). Let $\mathbf{h} = \{h_t, t = 0, 1, \dots, N-1\}$. We

can rewrite problem \mathcal{P}_1 as

$$\begin{aligned} & \underset{\Gamma, \mathbf{h}}{\text{minimize}} && \tilde{F} = \sum_{t=0}^{N-1} f(h_t) \\ & \text{subject to} && l_t(\Gamma) \leq h_t, \quad \text{for all } t, \\ & && v_i(\Gamma) \leq \eta_i, \quad \forall i \in \mathcal{I}. \end{aligned} \quad (5)$$

Let $\beta = [\beta_0, \beta_1, \dots, \beta_{N-1}]$ be the Lagrange multiplier vector corresponding to the constraints in Eq. (5). It will be clear that β serves as the congestion signal provided by the BS over time (in a day). Given β , we formulate and decompose the Lagrangian as follows:

$$\begin{aligned} \mathcal{L}(\Gamma, \mathbf{h}, \beta) &= \sum_{t=0}^{N-1} f(h_t) - \sum_{t=0}^{N-1} \beta_t [h_t - l_t(\Gamma)] \\ &= \sum_{t=0}^{N-1} [f(h_t) - \beta_t h_t] + \sum_{i \in \mathcal{I}} \sum_{t=0}^{N-1} \beta_t y_{i,t}(\Gamma), \end{aligned} \quad (7)$$

where $y_{i,t}(\Gamma) = \mathbb{E}[Y_{i,t}(\Gamma)]$ is the expected amount of resource consumed by user i in slot t . Let the objective function of the dual problem be $g(\beta)$, i.e.,

$$g(\beta) = \inf_{\Gamma, \mathbf{h}} \mathcal{L}(\Gamma, \mathbf{h}, \beta). \quad (8)$$

Since the Lagrangian has been decomposed, we can use a distributed policy Γ_i to minimize the expected consumed resource of user i such that the latter term in (7) is minimized. Therefore, for given β , the dual objective function can be obtained by solving the following subproblems:

$$\begin{aligned} (\mathcal{SP}_0) \quad & \underset{\mathbf{h} \geq \mathbf{0}}{\text{minimize}} && \sum_{t=0}^{N-1} [f(h_t) - \beta_t h_t], \\ (\mathcal{SP}_i) \quad & \underset{\Gamma_i}{\text{minimize}} && \sum_{t=0}^{N-1} \beta_t y_{i,t}(\Gamma_i) \\ & \text{subject to} && v_i(\Gamma_i) \leq \eta_i, \quad i \in \mathcal{I} \end{aligned}$$

The master dual-problem is

$$(\mathcal{D}_1) \quad \underset{\beta}{\text{maximize}} \quad g(\beta) \\ \text{subject to} \quad \beta \geq \mathbf{0}.$$

Since $f(\cdot)$ is convex, subproblem \mathcal{SP}_0 can be easily solved by convex optimization algorithms [13]. For subproblem \mathcal{SP}_i , we can view it as a constrained sequential decision problem and obtain the optimal policy Γ_i by backward induction [14]. Therefore, the dual problem can be solved efficiently by using (sub-)gradient approach, as will be discussed later.

For a general optimization problem, dual decomposition only guarantees weak duality, i.e., the dual solution only provides a lower bound to the original problem. However, we show below that the duality gap between \mathcal{P}_1 and \mathcal{D}_1 is zero, and hence there exists an optimal β such that the algorithms \mathcal{SP}_0 and \mathcal{SP}_i combined provides an optimal solution to \mathcal{P}_1 .

Proposition 2 *Given that the cost function $f(\cdot)$ is convex, the dual problem \mathcal{D}_1 have zero duality gap, and thus the dual decomposition approach provides an optimal value to \mathcal{P}_1 .*

Proof: Strong duality holds for convex optimization problem. To prove the proposition, we convert \mathcal{P}_1 to a convex problem with exponentially large number of control variables, and thus show the dual problem \mathcal{D}_1 have zero duality gap. See Appendix I for details. ■

The proof in Appendix I uses a transformation of policy representations, which will also be useful for design scheduling policies. Hence, we briefly introduce the transformation here. Let Ω_i be the set of possible realizations of channel process $R_i(t)$ for user i . For each realization denoted by $\mathbf{r} = [r(0), r(1), \dots, r(N-1)] \in \Omega_i$, let $\mathbf{r}(0:t) = [r(0), r(1), \dots, r(t)]$ be the first $t+1$ elements of \mathbf{r} . We only focus on causal policies, and thus the decision is made based on the history of the channel conditions. Let $x_{a_i, w, \mathbf{r}(0:a_i+w)} \in [0, 1]$ denote the transmission probability of user i when its arrival time is a_i , waiting time is w slots and the channel condition history is $\mathbf{r}(0:a_i+w)$. Then, a policy Γ_i for solving subproblem \mathcal{SP}_i can be represented by a decision matrix $\mathbf{x}_i = \{\mathbf{x}_{a_i}(\mathbf{r}) : 0 \leq a_i \leq N-1, \mathbf{r} \in \Omega_i\}$, where each submatrix $\mathbf{x}_{a_i}(\mathbf{r}) = [x_{a_i, w, \mathbf{r}(0:a_i+w)}]$ represents the policy for each pair of arrival time a_i and channel realization \mathbf{r} . For each decision matrix \mathbf{x}_i , we define the following transformation, denoted by $\varphi_i = \mathcal{T}(\mathbf{x}_i)$, as follows: for each realization $\mathbf{r} \in \Omega_i$,

$$\varphi_{a_i, w, \mathbf{r}(0:a_i+w)} = \begin{cases} x_{a_i, w, \mathbf{r}(0:a_i+w)}, & \text{if } w = 0, \\ x_{a_i, w, \mathbf{r}(0:a_i+w)} \prod_{w'=0}^{w-1} [1 - x_{a_i, w', \mathbf{r}(0:a_i+w')}], & \text{if } w = 1, 2, \dots, D_i - 1. \end{cases} \quad (9)$$

Note that $\varphi_{a_i, w, \mathbf{r}(0:a_i+w)}$ can be interpreted as the probability that user i under a particular channel realization \mathbf{r} transmits at time $a_i + w$. The transformation \mathcal{T} is invertible, where the inverse transformation $\varphi_i = \mathcal{T}^{-1}(\mathbf{x}_i)$ can be defined as follows: for a realization $\mathbf{r} \in \Omega_i$,

$$\begin{aligned} x_{a_i, w, \mathbf{r}(0:a_i+w)} &= \\ & \begin{cases} \varphi_{a_i, w, \mathbf{r}(0:a_i+w)}, & \text{if } w = 0, \\ \frac{\varphi_{a_i, w, \mathbf{r}(0:a_i+w)}}{\prod_{w'=0}^{w-1} (1 - x_{a_i, w', \mathbf{r}(0:a_i+w')})}, & \text{if } 0 < w \leq D_i - 1 \text{ and } \prod_{w'=0}^{w-1} (1 - x_{a_i, w', \mathbf{r}(0:a_i+w')}) > 0, \\ 0, & \text{if } 0 < w \leq D_i - 1 \text{ and } \prod_{w'=0}^{w-1} (1 - x_{a_i, w', \mathbf{r}(0:a_i+w')}) = 0. \end{cases} \end{aligned} \quad (10)$$

C. CoSchd: Coordinated Scheduling

Based on the dual decomposition discussed in the previous subsection, we propose the following distributed algorithm, referred to as Coordinated Scheduling (CoSchd), to solve the approximate problem \mathcal{P}_1 . In this distributed algorithm, the BS decides its congestion signal vector in an iterative fashion and each user individually decides its transmission schedule based on the congestion signal from the BS and its own channel characteristics.

In Algorithm 1, we follow a (sub-)gradient method to solve the dual problem \mathcal{D}_1 :

$$\beta_t^{(d+1)} = \left[\beta_t^{(d)} + \alpha^{(d)} (l_t^{(d)} - h_t^{(d)}) \right]^+, \quad \forall t \quad (13)$$

Algorithm 1 Coordinated Scheduling (CoSchd).

Input: Distributions of A_i, B_i, D_i ;

Transition probability matrix \mathbf{P}_i .

Output: Transmission probability $\bar{\mathbf{x}}_i$.

Init: Set $d = 0$ and $\beta_t^{(0)} = 1$ for all $t = 0, 1, \dots, N - 1$.

for $d \leftarrow 1, 2, \dots, d_{\max}$

1) **Mobile-side:** each user $i \in \mathcal{I}$ solves \mathcal{SP}_i and obtains

$$\mathbf{x}_i^{(d)} \leftarrow \arg \min \sum_{t=0}^{N-1} \beta_t y_{i,t}(\mathbf{x}_i),$$

$$\boldsymbol{\varphi}_i^{(d)} \leftarrow \mathcal{T}(\mathbf{x}_i^{(d)});$$

Each user estimates its expected load in each slot according to Eq. (16) and reports to the BS;

2) **Network-side:** the BS collects the load of each user and calculates the aggregated load $l_t^{(d)}$:

The BS solves \mathcal{SP}_0 and updates $\beta_t^{(d)}$ using Eq. (13) below;

endfor

Averaging: Calculate the average transition probability

$$\bar{\boldsymbol{\varphi}}_i \leftarrow \frac{1}{d_{\max}} \sum_{d=1}^{d_{\max}} \boldsymbol{\varphi}_i^{(d)}; \quad (11)$$

$$\bar{\mathbf{x}}_i \leftarrow \mathcal{T}^{-1}(\bar{\boldsymbol{\varphi}}_i). \quad (12)$$

where d is the iteration index, $\alpha^{(d)}$ is the step-size, and $[\cdot]^+$ denotes the projection to non-negative numbers.

Based on the framework of CoSchd, the network-side operation is simple: first, the BS solves subproblem \mathcal{SP}_0 and obtains the optimal value of $\mathbf{h}^{(d)}$; second, the BS updates congestion signals according to Eq. (13) based on load level $l^{(d)}$ and $\mathbf{h}^{(d)}$. Next, we focus on the operation on the mobile side.

1) Mobile-side Operation

On the mobile-side, each user operates independently as follows: it generates policies based on its channel characteristics and the congestion signals, and then executes the policy based on the instantaneous channel condition.

For a given $\boldsymbol{\beta}$ (i.e., the congestion signal vector), the subproblem \mathcal{SP}_i turns out to be a constrained sequential decision problem [14]. In the proof of Proposition 2, we consider general causal policies and assume that each user makes decisions based on its channel condition history. In fact, for Markov channel processes, each user only needs to make decision based on the waiting time and the current channel state. In particular, one can introduce a cost for deadline violation. The mobile minimizes \mathcal{SP}_i plus the deadline violation cost by backward induction. We now discuss the specifics of the deterministic deadline-constraint case as follows and refer the readers to [14] for the probabilistic deadline constraint case.

For user i arriving at a_i , let $x_{a_i,w,j} \in [0, 1]$ ($w = 0, 1, \dots, D_i - 1; j = 1, 2, \dots, J$) be the probability that user i requests transmission when its waiting time is w and

channel state is j . (Thus, the probability $x_{a_i,w,r(0:a_i+w)} = x_{a_i,w,j}$ if $r(a_i + w) = r_j$.) In the deterministic deadline-constraint case, i.e., $\eta_i = 0$, all data must be transmitted before expiration. Therefore, for user i arriving at a_i , it requires that $x_{a_i,a_i+D_i-1,j} = 1$. To guarantee a finite transmission cost, we assume that for each user,

$$\mathbb{E}\{B_i/R_i(a_i + D_i - 1)|\mathcal{E}_{i,D_i-1}\} < +\infty, \quad i \in \mathcal{I}, \quad (14)$$

where \mathcal{E}_{i,D_i-1} represents the event that user i does not transmit before $a_i + D_i - 1$. In the case with temporally-Markovian channels, using the principle of optimality and taking the multipliers $\boldsymbol{\beta}$ into account, we can obtain the optimal decision

$$x_{a_i,w,j} = \begin{cases} 1, & \text{if } \frac{\beta_{a_i+w}}{R_i(a_i+w)} \leq \mathbb{E}[V_{a_i,w+1}^*|r_j] \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where $\mathbb{E}[V_{a_i,w+1}^*|r_j]$ is the expected future cost conditioned on $R_i(a_i + w) = r_j$, which can be calculated by backward induction:

$$\begin{aligned} & \mathbb{E}[V_{a_i,w+1}^*|r_j] \\ &= \begin{cases} \mathbb{E}\left[\frac{\beta_{a_i+D_i-1}}{R_i(a_i+D_i-1)}|r_j\right] & \text{for } w = D_i - 2; \\ \mathbb{E}\left[\min\left(\frac{\beta_{a_i+w+1}}{R_i(a_i+w+1)}, V_{a_i,w+2}^*\right)|r_j\right], & \text{for } w = D_i - 3, D_i - 4, \dots, 0. \end{cases} \end{aligned}$$

As a special case of Markovian channels, when the channel process is independent across time-slots, it is easy to verify that the policy becomes a threshold policy, i.e., there exists a threshold T_w for each w , the transfer occurs if $R_i(a_i + w) \geq T_w$.

After obtaining \mathbf{x}_i , each user can estimate the amount of required resource as follows:

$$\begin{aligned} y_{i,t} &= \sum_{a=0}^{N-1} \mathbb{P}(A_i = a) y_{i,a,t} \\ &= \sum_{a=0}^{N-1} \mathbb{P}(A_i = a) \mathbb{E}[B_i] \sum_{j=2}^J \pi'_{i,a,t,j} / r_j, \quad (16) \end{aligned}$$

where $\pi'_{i,a,t,j}$ is the probability that the user i with arrival time a transmits at slot t under channel condition r_j , i.e.,

$$\pi'_{i,a,t,j} = \begin{cases} \pi_j^{(i)}, & \text{if } t = a, \\ \sum_{j'=1}^J (1 - x_{a,t-a,j'}) \pi'_{i,a,t-1,j'} p_{j'j} & \text{otherwise.} \end{cases}$$

Finally, the averaging operation given by (11) and (12) is taken to deal with the possible oscillation issues of the subgradient method, as in [15]. Note that the constraints of \mathcal{P}_1 are linear in $\boldsymbol{\varphi}_i$ and the subgradient is bounded. Using the results in [15], we can show that CoSchd provides a approximate solution of problem \mathcal{P}_1 . Specifically, we use a constant step-size in (13) and let $\alpha^{(d)} = \alpha$. Let $\bar{F}_{\text{CoSchd}(\alpha)}$ be the cost value of \mathcal{P}_1 under CoSchd with $\alpha^{(d)} = \alpha$. Then, the following lemma states that CoSchd(α) provides a near

optimal solution of \mathcal{P}_1 . The proof is similar to Proposition 2 in [15] and is omitted here.

Lemma 1 For CoSchd with constant step-size $\alpha^{(d)} = \alpha$, the cost of problem \mathcal{P}_1 is bounded as follows:

$$\tilde{F}_{\text{CoSchd}(\alpha)} \leq F_{\text{lb}} + \mu_0 \alpha, \quad (17)$$

where F_{lb} is the optimal value of \mathcal{P}_1 , and $\mu_0 > 0$ is a constant depending on the maximum value of $\|l_t - h(t)\|$.

2) Asymptotic Optimality in the Many-Source Regime

According to Proposition 2, \mathcal{P}_1 can be solved by dual decomposition approach. However, \mathcal{P}_1 is not equivalent to the original problem \mathcal{P}_0 . Fortunately, because all users independently solve individual MDPs under CoSchd, as the number of users increases, the instantaneous load level is close to its expectation. Using this property, we can show that the proposed approach is asymptotically optimal for \mathcal{P}_0 in the many-source regime.

Consider the many-source regime. To study the asymptotic properties of the proposed approach, we consider the following m -scaled system.

Assumption 1 All users in \mathcal{I} can be divided into K classes. For each class k ,

- the number of users m_k ($k = 1, 2, \dots, K$) is proportional to the total number of users m , i.e., $m_k = m\lambda_k$, where $0 < \lambda_k < 1$ is the ratio of class- k users and $\sum_{k=1}^K \lambda_k = 1$;
- users in class- k have the same deadline requirements, and the same statistics of arrival time and channel dynamics that do not change with m .

Further, we make the following assumption on the cost function:

Assumption 2 The cost function $f(\cdot)$ in the m -scaled system is continuous, and is a function of the normalized load, i.e., $f(l) = \tilde{f}(\tilde{l})$, where $\tilde{l} = l/m$.

For the above m -scaled system, we let $F_{\text{CoSchd}(\alpha)}^{(m)}$ be the cost value of the original problem \mathcal{P}_0 under CoSchd with $\alpha^{(d)} = \alpha$, and let $F_{\text{lb}}^{(m)}$ be the optimal value of problem \mathcal{P}_1 . Note that by optimizing on the normalized load-level \tilde{l}_t , the constant μ_0 in Lemma 1 for the m -scaled system is independent of m . The following proposition then shows the performance of CoSchd in the many-source regime.

Proposition 3 Under Assumptions 1 and 2, $F_{\text{CoSchd}(\alpha)}^{(m)}$ converges to a value near $F_{\text{lb}}^{(m)}$ as m increases, i.e.,

$$\lim_{m \rightarrow \infty} F_{\text{CoSchd}(\alpha)}^{(m)} \leq \lim_{m \rightarrow \infty} F_{\text{lb}}^{(m)} + \mu_0 \alpha, \quad (18)$$

where μ_0 is the constant introduced in Lemma 1.

Proof: See Appendix II. ■

According to Proposition 1, $F_{\text{lb}}^{(m)}$ provides a lower bound on the optimal value of \mathcal{P}_0 . The above proposition states that $F_{\text{CoSchd}(\alpha)}^{(m)}$ will be in a neighborhood of $F_{\text{lb}}^{(m)}$ as m increases, and thus the decomposition approach is asymptotically optimal for \mathcal{P}_0 in the many-source regime.

D. Online Implementation of CoSchd

The CoSchd approach proposed in the previous section can only be implemented before really running the network. Next, we propose an online version of CoSchd (online-CoSchd) in Algorithm 2, where in every day, the scheduling decision is iteratively computed based on the actual traffic-load measurements.

Algorithm 2 Online-CoSchd.

Init:

set $d = 0$ and $\beta_t^{(0)} = 1$ for all $t = 0, 1, \dots, N - 1$.

Iteration: (day d)

1) At time $t = 0, \beta_t^{(d)}$ ($t = 0, 1, \dots, N - 1$) is announced to all users;

Each user $i \in \mathcal{I}$ solves \mathcal{SP}_i as the request arrives and calculates the average decision matrix according to (19);

2) For $t = 0 \rightarrow N - 1$,

Each user makes decision based on its channel states;

The BS serves requested users and observes the load level $L_t^{(d)}$;

The BS solves \mathcal{SP}_0 and updates $\beta_t^{(d)}$ using Eq. (13) with $l_t^{(d)} = L_t^{(d)}$;

3) Set $d \leftarrow d + 1$ and go to step 1).

Recall that Algorithm 1 needs to consider all possible realizations of channel process for each user when calculating the average value of the decision matrices, and it becomes very complex as the deadline increases (because $|\Omega_i| = J^{D_i}$). In contrast, in Algorithm 2, here each user directly applies a Exponential-Moving-Averaging policy to deal with the oscillation issues, i.e.,

$$\bar{\mathbf{x}}_i^{(d)} = \vartheta \bar{\mathbf{x}}_i^{(d-1)} + (1 - \vartheta) \mathbf{x}_i^{(d)}, \quad (19)$$

where $\vartheta \in [0, 1]$ is a memory factor for trading-off the convergence speed and fluctuation, as will be discussed in Section IV. Without considering all possible realizations, the number of control variables for each user is reduced to $D_i J$, which is much smaller than J_i^D for large D_i .

With the online policy, each user also estimates its instantaneous channel condition and then makes decision using the decision table, as shown in Eq. (15). We will evaluate the performance of online-CoSchd through simulation in Section IV. In the current stage, we have not obtained theoretic properties for the online version of CoSchd and leave it as our future work.

IV. EVALUATION

Based on the framework proposed in our paper, we can also derive load-only and channel-only policies as special cases (with constant congestion signal and data rate, respectively). In this section, We evaluate the performance of load-only, channel-only, and CoSchd approaches via trace-driven simulations. Since the offline-CoSchd and online-CoSchd have similar performance and online-CoSchd is more scalable as discussed in Section III, we only consider online-CoSchd here

and refer it to as CoSchd for simplicity. As a baseline, we also consider *ImTrans*, where all users immediately transfer the data when the requests arrive.

A. Simulation Setup

We use a slot length of 10 minutes and each day is divided into 144 time-slots. For the cost function, we set ν to be $\nu = 8$ which is large enough for smoothing out the load level according to our experiments. We consider both a single-cell scenario and a two-cell scenario, except that the load-only policy will only be evaluated in the single-cell scenario similar to [1].

1) Traffic Arrival Pattern

We assume a time-dependent Poisson arrival process, i.e., the total number of requests arriving in time-slot t is a Poisson random variable with mean value λ_t ($t = 0, 1, \dots, N-1$). For the single-cell network, the mean arrival rates are set based on the weekday traffic profile of the center BS from Fig. 1 in [16]. For the multi-cell network, we use the weekday traffic profile of the center BS and the neighbor BS 1 (again from Fig. 1 in [16]). To capture the delay-tolerance of traffic, we apply the waiting function proposed in [1], and use the patience indices for the different traffic classes estimated from the U.S. survey in [1]. Specifically, for the delay-tolerant traffic (“Time-Dependent Pricing” traffic in [1]), the probability that user i wants to wait D_i slots is proportional to $\frac{1}{(D_i+1)^\rho}$, where the patience index ρ is 2.0 for video traffic and 0.6 for others. In addition, the usage distribution of the different traffic classes is taken from recent estimates [17], where the proportion of video traffic is about 65%.

2) Channel Profile

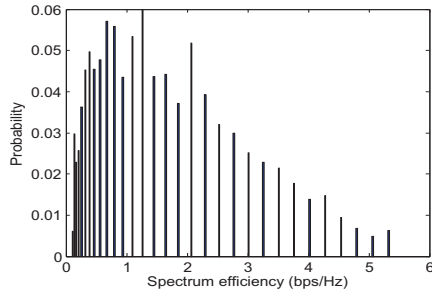


Fig. 1. Distribution of spectrum efficiency.

We collected a set of Received Signal Strength Indication (RSSI) values from a group of anonymous mobile users to best emulate the spectrum efficiency in cellular networks. We assume that the interference strength is a constant and thus the RSSI value represents the SINR, which determines the spectrum efficiency. We follow the LTE-Advanced standards [18], and map the measured RSSI to the proper modes of Modulation and Coding Scheme (MCS). We use the 5-bit CQI and the distribution of the corresponding spectrum efficiency is shown in Fig. 1. To capture time-varying and location-dependent channel conditions, we use a Markov model where Markovian transitions between adjacent channel states (RSSI

values) are assumed in each time-slot [19]. We assume that all users use the same channel model. One limitation of the model is that the parameters (e.g., transition probabilities) do not change over time while real human users may have more time-dependent behavior (e.g., 2am at home vs. 2pm at work). We hope to further collect real-life channel profile traces for a more realistic evaluation of real networks in the future.

B. Convergence of CoSchd

We first demonstrate the asymptotic behavior of the system and the convergence of CoSchd.

Fig. 2 shows the difference between the values of the original problem \mathcal{P}_0 and its approximate version \mathcal{P}_1 . Note that the cost of expected load $f[\mathbb{E}(L_t)]$ is close to the lower bound (Lemma 1) and the expected cost under CoSchd $\mathbb{E}[f(L_t)]$ provides an upper bound on the original problem \mathcal{P}_0 . As we can see from the figure, the gap between the upper- and lower-bounds becomes smaller as the network scale increases. The two values are close to each other in medium-sized systems. For example, when the average number of users in each slot is 400, the gap is about 15% of the value for the original problem. Hence, minimizing the cost of expected load approximately solves the original problem \mathcal{P}_0 .

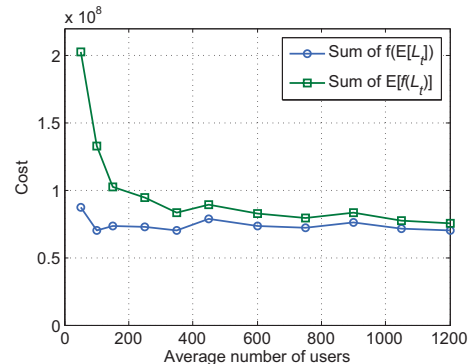


Fig. 2. The difference between the cost of expected load and the expectation of cost.

Fig. 3 shows the evolution of the duality gap between problem \mathcal{P}_1 and its duality. The duality gap decreases as the number of iterations increases, and the duality gap is small after several iterations. Comparing the evolutions with different memory factor ϑ , we can see that with smaller ϑ , the duality gap decreases faster, but the fluctuation is larger. We set $\vartheta = 0.9$ for the rest of simulations which seems to strike a balance between the convergence speed and fluctuation.

C. Network Load

Fig. 4 shows the network load in one day obtained by different approaches. The three subfigures represent different settings. Fig. 4(a) and Fig. 4(b) are for single-cell systems with, respectively, 50% and 75% of load being delay-tolerant. In contrast, Fig. 4(c) is for a multi-cell system with 50% of load being delay-tolerant. We can make a number of interesting observations from Figs. 4(a) and 4(b). Specifically,

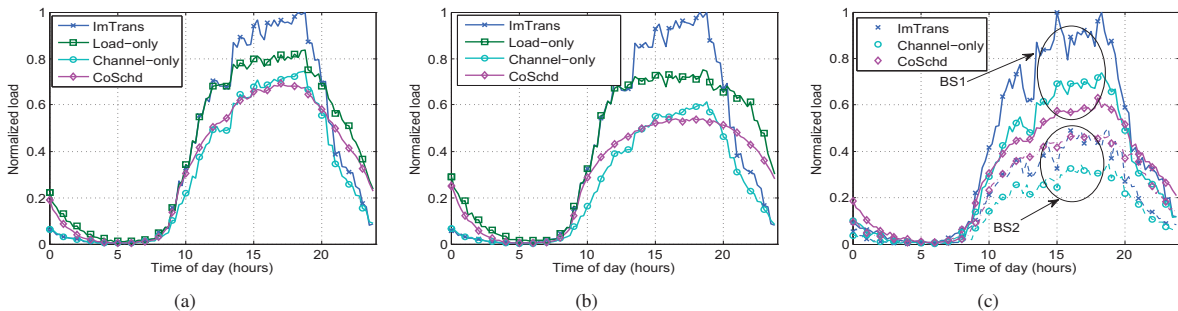


Fig. 4. Load level under different approaches and settings (a) single-cell with 50% delay-tolerant traffic, (b) single-cell with 75% delay-tolerant traffic, and (c) multi-cell with 50% delay-tolerant traffic (solid-line for BS1 and dash-line for BS2).

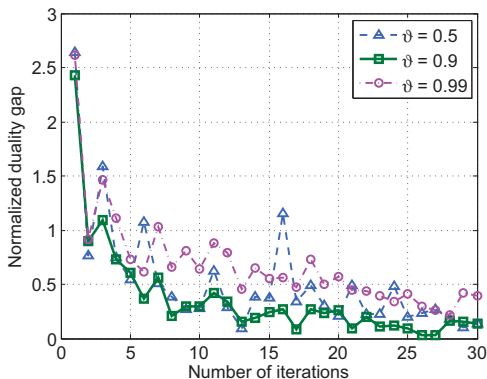


Fig. 3. The convergence of CoSchd.

from Fig. 4(a), we can see that by moving the delay-tolerant traffic into “valleys”, the peak load obtained by the load-only policy is about 80% of that under ImTrans. On the other hand, using the channel-only policy, the peak is reduced to about 75% of ImTrans. A similar observation can be made from Fig. 4(b), while the peak load reduction is more significant since there is more delay-tolerant traffic. This finding suggests that channel-awareness can be more effective than load-awareness in wireless systems.

Further, although CoSchd leads to even lower peak consumption by considering both load-awareness and channel-awareness, the additional gain compared to the channel-only policy is relatively marginal in the single-cell setting in Fig. 4(a) and Fig. 4(b) (about 8% reduction in both figures). We note that, under the channel-only policy, users defer their transmissions when waiting for good channels. Therefore, a “peak-shedding” effect also occurs under the channel-only approach. Since the traffic fluctuation is not large, the room for CoSchd to further move traffic is relatively small. However, the multi-cell simulation in Fig. 4(c) illustrates different behaviors. By moving the delay-tolerant traffic to the neighbor BS (i.e., BS 2), the peak of network load (corresponding to the load in BS 1 at about 18:00) is reduced by about 20% by CoSchd compared to the channel-only policy.

V. CONCLUSIONS

In this paper, we present a decomposition technique for solving a large-scale MDP that models the problem of scheduling delay-tolerant traffic with both load- and channel-awareness. Despite the high complexity of the large-scale MDP, we develop a distributed framework, called CoSchd, and shows its asymptotic-optimality in the many-source regime. An online version of CoSchd is also proposed to reduce the complexity.

The results in this paper are of both practical and theoretical values. Practically, our proposed policy can be implemented in a distributed manner in real systems. Further, our comparative evaluations provide cellular operators with operation guidelines to decide the most appropriate approaches. Specifically, our numerical results suggest that channel-awareness is rather important in wireless networks. For single-cell systems, channel-only may be preferred due to its simplicity and relatively good performance. For multi-cell systems with load variations, CoSchd can attain significant additional gains. Theoretically, the joint approach provides an optimal benchmark for comparing with other solutions. Moreover, the decomposition technique and the proposed CoSchd algorithm can potentially be applied to other large-scale MDP, where multiple agents are weakly coupled through sharing common resources.

APPENDIX I PROOF OF PROPOSITION 2

To prove the optimality of the proposed dual decomposition approach, we show that problem \mathcal{P}_1 can be reformulated to a convex optimization problem \mathcal{P}_2 , albeit with an exponentially large number of decision variables. Thus, strong duality holds between \mathcal{P}_2 and its dual, named \mathcal{D}_2 .

First, we show that any policy Γ can be represented by a stochastic policy Ψ as follows¹. Note that each causal policy Γ makes decision based on the history of the arrival sequence and channel processes. To represent the history, for each user $i \in \mathcal{I}$, we introduce $\tilde{A}_i(t)$ to represent its present status in time-slot t . Namely, if the arrival time A_i of user i is equal

¹The representation and transformation are slightly different from that in Section III, because here scheduling policies are made based on the state of the entire network before decomposition.

to a_i (we let $a_i = N$ represent the event that user i does not appear), then $\tilde{A}_i(t) = -1$ if $a_i > t$, and $\tilde{A}_i(t) = a_i$ if $a_i \leq t$. Recall that $R_i(t)$ ($i = 0, 1, \dots, N-1$) is the channel process of user i . Hence, the history of the system up to time-slot t is given by

$$\mathbf{S}_t = [\tilde{\mathbf{A}}_t \tilde{\mathbf{R}}_t],$$

where $\tilde{\mathbf{A}}_t = [\tilde{A}_1(t), \tilde{A}_2(t), \dots, \tilde{A}_{|\mathcal{I}|}(t)]^T$ and

$$\tilde{\mathbf{R}}_t = \begin{bmatrix} R_1(0) & R_1(1) & \dots & R_1(t) \\ R_2(0) & R_2(1) & \dots & R_2(t) \\ \vdots & \vdots & \ddots & \vdots \\ R_{|\mathcal{I}|}(0) & R_{|\mathcal{I}|}(1) & \dots & R_{|\mathcal{I}|}(t) \end{bmatrix}.$$

Let Ω be the set of possible realizations of arrival sequence and channel processes, i.e., the possible realization of \mathbf{S}_{N-1} . Then, each policy Γ can be represented by a stochastic policy Ψ , which is a $\Omega \mapsto [0, 1]^{|\mathcal{I}| \times N}$ mapping: for each $\mathbf{s} \in \Omega$,

$$\Psi(\mathbf{s}) = \begin{bmatrix} \psi_1(\mathbf{s}_0) & \psi_1(\mathbf{s}_1) & \dots & \psi_1(\mathbf{s}_{N-1}) \\ \psi_2(\mathbf{s}_0) & \psi_2(\mathbf{s}_1) & \dots & \psi_2(\mathbf{s}_{N-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{|\mathcal{I}|}(\mathbf{s}_0) & \psi_{|\mathcal{I}|}(\mathbf{s}_1) & \dots & \psi_{|\mathcal{I}|}(\mathbf{s}_{N-1}) \end{bmatrix},$$

where \mathbf{s}_t is the history of arrival sequence and channel processes up to time-slot t for the realization \mathbf{s} , and $\psi_i(\mathbf{s}_t) \in [0, 1]$ is the transmission probability of user i in time-slot t .

Second, we study the expected resource consumed by user i under $\Psi(\mathbf{s})$. For each $\mathbf{s} \in \Omega$ where user i arrives in time-slot a_i , we can calculate the probability that user i transmits in slot $a_i + w$ as follows

$$\varphi_i(\mathbf{s}_t) = \begin{cases} \psi_i(\mathbf{s}_{a_i}), & t = a_i \\ \psi_i(\mathbf{s}_{a_i+w}) \prod_{w'=0}^{w-1} [1 - \psi_i(\mathbf{s}_{a_i+w'})], & t = a_i + w, 0 < w \leq D_i - 1 \\ 0, & \text{otherwise.} \end{cases}$$

For given \mathbf{s} , the expected consumed resource of user i in time-slot t is

$$c'_{i,t}(\mathbf{s}, \Psi) = \frac{b_i \varphi_i(\mathbf{s}_t)}{R_i(t)}.$$

In addition, note that all users should transmit before expiration. Hence,

$$\sum_{w=0}^{D_i-1} \varphi_i(\mathbf{s}_{a_i+w}) = 1, \quad \mathbf{s} \in \Omega, i \in \mathcal{I}. \quad (20)$$

Moreover, using the relationship between $\varphi_i(\cdot)$ and $\psi_i(\cdot)$, a $\varphi_i(\cdot)$ satisfying (20) can be mapped to a policy Ψ ².

²If $\sum_{w'=0}^w \varphi_i(\mathbf{s}_{0,a_i+w'}) = 1$ for some $w < D_i - 1$, then for $w' > w$, $\psi_i(\mathbf{s}_{i,w'})$ can be artificially set to be 0, which will not affect the behavior of Ψ .

Consequently, problem \mathcal{P}_1 is equivalent to

$$\begin{aligned} (\mathcal{P}_2) \quad & \underset{\Psi, h'_t}{\text{minimize}} && F = \sum_{t=0}^{N-1} f(h'_t), \\ & \text{subject to} && \sum_{w=0}^{D_i-1} \varphi_i(\mathbf{s}_{a_i+w}) = 1, \quad \mathbf{s} \in \Omega, i \in \mathcal{I}, \\ & && l'_t(\Psi) \leq h'_t, \quad t = 0, 1, \dots, N-1, \end{aligned}$$

where

$$l'_t(\Psi) = \sum_{\mathbf{s} \in \Omega} \sum_{i \in \mathcal{I}} \pi(\mathbf{s}) c'_{i,t}(\mathbf{s}, \Psi). \quad (21)$$

We can verify that \mathcal{P}_2 is a convex optimization problem because $f(\cdot)$ is a convex function and all the constraints are linear. However, we do note that it is impractical to solve \mathcal{P}_2 directly because of its large number of variables. Recall that there are $|\mathcal{I}| \times N$ decision variable for each possible state. Assume the channel state of each user can be quantized to J values, then there are $J^{|\mathcal{I}| \times N}$ possible states, and thus $|\mathcal{I}| \times N \times J^{|\mathcal{I}| \times N}$ decision variables, which is clearly intractable. We note that the formulation can be considered as a linear representation of a centralized Markov Decision Policy, which clearly suffers the curse of dimensionality.

Again, we resort to the dual decomposition approach to study \mathcal{P}_2 . Similar to the approach in Section III, we can introduce a dual variable for each time slot, and then rearrange the variables that belong to each user. Then, we have a similar format as in \mathcal{SP}_0 and \mathcal{SP}_i . The dual decomposition approach can also be applied to solve problem \mathcal{P}_2 and the strong duality holds.

APPENDIX II PROOF OF PROPOSITION 3

Let $\tilde{F}_{\text{CoSched}(\alpha)}^{(m)}$ be the cost value of \mathcal{P}_1 under $\text{CoSched}(\alpha)$. Because $\tilde{F}_{\text{CoSched}(\alpha)}^{(m)} \leq F_{\text{lb}}^{(m)} + \mu_0 \alpha$ according to Lemma 1, we only need to show that $\lim_{m \rightarrow \infty} F_{\text{CoSched}(\alpha)}^{(m)} = \lim_{m \rightarrow \infty} \tilde{F}_{\text{CoSched}(\alpha)}^{(m)}$. To achieve this, we first consider the single-class system, i.e., $K = 1$. Since $F_{\text{CoSched}(\alpha)}^{(m)}$ is the sum of the expected costs in each slot, i.e., $F_{\text{CoSched}(\alpha)}^{(m)} = \sum_{t=0}^{N-1} \mathbb{E}[f(L_t)]$, we can prove Proposition 3 if we can show that under $\text{CoSched}(\alpha)$,

$$\lim_{m \rightarrow \infty} \mathbb{E}[f(L_t)] = f(\mathbb{E}[L_t]), \quad (22)$$

which implies that the ‘‘expectation of the cost’’ approaches the ‘‘cost of the expectation’’ as m increases. As will be seen shortly, this can be verified by using the fact that, under $\text{CoSched}(\alpha)$, each user operates independently when the congestion signal β is fixed.

Specifically, fix a time-slot t . Let Y_i ($i = 1, 2, \dots, m$) be the amount of resource required by the i -th user in slot t . Since all users in the same class have identical traffic and channel statistics, Y_i 's ($i = 1, 2, \dots, m$) are i.i.d. random variables. Let $\mathbb{E}[Y_i] = \mu_Y$. The load level is $L_t = \sum_{i=1}^m Y_i$ and the normalized load level is $\tilde{L}_t = \frac{1}{m} \sum_{i=1}^m Y_i$ with $\mathbb{E}[\tilde{L}_t] = \mu_Y$.

Since the file size B_i is bounded, the amount of resource Y_i is bounded and we let $y_{\max} = \max Y_i = \frac{\max B_i}{r_2}$.

Using the Chernoff bound, we have that for a given $\delta > 0$,

$$\mathbb{P}\{\tilde{L}_t \leq \mu_Y - \delta\} \leq e^{-mI_Y(\delta)} \quad (23)$$

$$\mathbb{P}\{\tilde{L}_t \geq \mu_Y + \delta\} \leq e^{-mI_Y(\delta)}, \quad (24)$$

where

$$I_Y(\delta) = \min\{D(\mu_Y + \delta || \mu_Y), D(\mu_Y - \delta || \mu_Y)\}, \quad (25)$$

and $D(x||y) = x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}$ is the Kullback-Leibler divergence.

Next, we bound $\mathbb{E}[f(L_t)]$ using the above results and the properties of the cost function. $\mathbb{E}[f(L_t)]$ can be calculated as follows

$$\begin{aligned} \mathbb{E}[f(L_t)] &= \mathbb{E}[\tilde{f}(\tilde{L}_t)] \\ &= \int_0^\infty \tilde{f}(l) \phi_{\tilde{L}_t}(l) dl \\ &= \left[\int_0^{\mu_Y - \delta} + \int_{\mu_Y - \delta}^{\mu_Y + \delta} + \int_{\mu_Y + \delta}^\infty \right] \tilde{f}(l) \phi_{\tilde{L}_t}(l) dl, \end{aligned}$$

where $\phi_{\tilde{L}_t}(\cdot)$ is the probability density function of \tilde{L}_t . Note that $\tilde{f}(l)$ is increasing in l . Thus, from (23) and (24), we have

$$\mathbb{E}[f(L_t)] \geq (1 - 2e^{-mI_Y(\delta)}) \tilde{f}(\mu_Y - \delta), \quad (26)$$

and

$$\mathbb{E}[f(L_t)] \leq \tilde{f}(\mu_Y + \delta) + 2e^{-mI_Y(\delta)} \tilde{f}(y_{\max}). \quad (27)$$

For any $\epsilon > 0$, by the continuity of $\tilde{f}(l)$, we can choose a $\delta > 0$ such that $\tilde{f}(\mu_Y - \delta) \geq \tilde{f}(\mu_Y) - \epsilon/2$ and $\tilde{f}(\mu_Y + \delta) \leq \tilde{f}(\mu_Y) + \epsilon/2$. Combining with (26) and (27), we know that there exists an m_1 such that for all $m \geq m_1$, we have

$$|\mathbb{E}[f(L_t)] - f(\mathbb{E}[L_t])| = |\mathbb{E}[f(L_t)] - \tilde{f}(\mu_Y)| \leq \epsilon,$$

and thus (22) holds by taking $\epsilon \rightarrow 0$.

For multi-class systems, similar properties can be obtained. Let \mathcal{I}_k ($k = 1, 2, \dots, K$) be the index set of class- k users. For all $i \in \mathcal{I}_k$, Y_i are i.i.d. random variables because within one class, all users have identical traffic and channel characteristics. Let $\mu_Y^{(k)} = \mathbb{E}[Y_i]$ for $i \in \mathcal{I}_k$. Then, the load level is $L_t = \sum_{k=1}^K \sum_{i \in \mathcal{I}_k} Y_i$ and the normalized load level is given by

$$\tilde{L}_t = \frac{L_t}{m} = \sum_{k=1}^K \lambda_k \tilde{L}_t^{(k)},$$

where $\tilde{L}_t^{(k)} = \frac{1}{m_k} \sum_{i \in \mathcal{I}_k} Y_i$, and $\mathbb{E}[\tilde{L}_t] = \sum_{k=1}^K \lambda_k \mu_Y^{(k)}$. Because the event $\tilde{L}_t \leq \mathbb{E}[\tilde{L}_t] - \delta$ implies that there exists at least one k satisfying that $\tilde{L}_t^{(k)} \leq \mu_Y^{(k)} - \delta$, we have

$$\begin{aligned} &\mathbb{P}\{\tilde{L}_t \leq \mathbb{E}[\tilde{L}_t] - \delta\} \\ &\leq \sum_{k=1}^K \mathbb{P}\{\tilde{L}_t^{(k)} \leq \mu_Y^{(k)} - \delta\} \leq K e^{-mI_Y^*(\delta)} \quad (28) \end{aligned}$$

where

$$I_Y^*(\delta) = \min_{k \in \{1, 2, \dots, K\}} \lambda_k \min\{D(\mu_Y^{(k)} + \delta || \mu_Y^{(k)}), D(\mu_Y^{(k)} - \delta || \mu_Y^{(k)})\}.$$

Similarly,

$$\mathbb{P}\{\tilde{L}_t \geq \mathbb{E}[\tilde{L}_t] + \delta\} \leq K e^{-mI_Y^*(\delta)}. \quad (29)$$

Therefore, using the same approach in the single-class case, $\mathbb{E}[f(L_t)]$ can be made as close to $\tilde{f}(\mathbb{E}[\tilde{L}_t])$ as desired.

The conclusion then follows by adding the expected costs from time-slot 0 to $N - 1$.

REFERENCES

- [1] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, "TUBE: Time-dependent pricing for mobile data," in *Proc. ACM SIGCOMM'12*, Helsinki, Finland, Aug. 2012.
- [2] M. Ra, J. Paek, A. Sharma, R. Govindan, M. Krieger, and M. Neely., "Energy-delay tradeoffs in smartphone applications," in *Proc. ACM MobiSys'10*, San Francisco, CA, June 2010, pp. 255 - 270.
- [3] I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci, "Taming user-generated content in mobile networks via drop zones," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 2840 - 2848.
- [4] A. Schulman, V. Navda, R. Ramjee, N. Spring, P. Deshpande, C. Grunewald, K. Jain, and V. N. Padmanabhan, "Bartendr: A practical approach to energy-aware cellular data scheduling," in *Proc. ACM MobiCom'10*, Chicago, Sept. 2010, pp. 85 - 96.
- [5] TUBE, "TUBE: Time-Dependent Pricing." [Online]. Available: <http://scenic.princeton.edu/tube/tdp.html>
- [6] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting mobile 3G using WiFi," in *Proc. of ACM MobiSys'10*, San Francisco, CA, June 2010, pp. 209-222.
- [7] A. Chakraborty, V. Navda, V. N. Padmanabhan, and R. Ramjee, "Coordinating cellular background transfers using LoadSense," in *Proc. ACM MobiCom'13*, Miami, FL, USA, 2013.
- [8] H. Wu, X. Lin, X. Liu, K. Tan, and Y. Zhang, "CoSchd: the virtue of channel and load awareness in alleviating cellular congestion," *Technical Report*, 2014. [Online]. Available: <https://engineering.purdue.edu/%7elinx/papers.html>
- [9] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Towards dynamic energy-efficient operation of cellular network infrastructure," *IEEE Communications Magazine*, vol. 49, no. 6, 2011.
- [10] R. Irmer (Editor in charge), "NGMN radio access performance evaluation methodology," NGMN White Paper, Jan. 2008.
- [11] A. J. Nicholson and B. D. Noble, "Breadcrumbs: forecasting mobile connectivity," in *Proc of ACM MobiCom'08*, New York, NY, USA, 2008, pp. 46-57.
- [12] A. Rahmati and L. Zhong, "Context-for-wireless: context-sensitive energy-efficient wireless data transfer," in *Proceedings of ACM MobiSys'07*, New York, NY, USA, 2007, pp. 165-178.
- [13] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [14] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed. Wiley, 2005.
- [15] A. Nedic and A. Ozdaglar, "Subgradient methods in network resource allocation: Rate analysis," in *Proc. the 42nd Annual Conference on Information Sciences and Systems*, Mar 2008, pp. 1189 - 1194.
- [16] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Towards dynamic energy-efficient operation of cellular network infrastructure," *IEEE Communications Magazine*, vol. 49, no. 6, 2011.
- [17] "Cisco visual networking index: Global traffic forecast update, 2012 - 2017," Feb. 2013. [Online]. Available: <http://www.cisco.com>
- [18] 3GPP TS 36.213 V11.0.0, "Physical layer procedures," Sept. 2012.
- [19] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE Trans. Vehicular Technology*, vol. 44, no. 1, 1995.