Online Scheduling for Energy Efficiency in Real-Time Wireless Networks

Shuai Zuo
CESG and Department of ECE
Texas A&M University
College Station, TX 77843, USA
zuosh029@neo.tamu.edu

I-Hong Hou
CESG and Department of ECE
Texas A&M University
College Station, TX 77843, USA
ihou@tamu.edu

***Abstract*—This paper studies the problem of using minimum power to provide satisfactory performance for real-time applications over unreliable and fading wireless channels. We demonstrate that this problem can be formulated as a linear programming problem. However, this formulation involves exponentially many constraints, and many parameters are either unavailable or difficult to compute, which makes it infeasible to employ standard techniques to solve the linear programming problem. Instead, we propose a simple online scheduling algorithm for this problem. This algorithm has very low complexity and makes scheduling decisions solely based on system history and current channel conditions. It is also compatible with any power control algorithms. We prove that our algorithm provides satisfactory performance to all real-time applications, and the total power consumption can be made arbitrarily close to the theoretical lower bound. Simulation results show that our scheduling algorithm indeed achieves small power consumption with fast convergence.**

I. Introduction

Wireless networks have been widely applied to serve real-time applications with strict per-packet delay bounds like VoIP, interactive multimedia, video streaming and real-time surveillance. With the recent emphasis on green communications, reducing energy consumption while providing satisfactory performance to all real-time applications has become an important concern for future wireless networks.

One major challenge for energy minimization in wireless networks is that wireless transmissions are unreliable, which incurs significant energy consumption for failed transmissions. The channel conditions can change over time, and it may be infeasible to obtain the distributions of channel conditions. Wireless transmitters may employ power control algorithms to alleviate the impact of fading wireless channels.

In this paper, we study the problem of designing an optimal scheduling policy that minimizes power consumption for serving real-time applications over fading wireless channels. We consider an analytical model that incorporates the hard per-packet delay bounds and throughput requirements of real-time clients, the fading and unreliable wireless channels, and the usage of power control algorithms. We demonstrate that minimizing power consumption can be formulated as a linear programming problem. However, such a problem involves exponentially many constraints, and many parameters are either not available or difficult to compute. Therefore, standard techniques for solving linear programming problems cannot be applied.

We then propose a simple online scheduling algorithm that jointly controls various variables involved in serving real-time applications. The algorithm has very low complexity, and it only requires the knowledge of a small number of parameters that are directly available without any computations. We prove that this simple online algorithm satisfies the requirements of all real-time clients. We also prove the difference between the total power consumption of this algorithm and that of an optimum offline algorithm is upper-bounded by a constant, and the constant can be made arbitrarily small. In other words, our algorithm solves the linear programming problem without knowing its parameters.

The performance of the online algorithm is further evaluated simulations. We have implemented the algorithms in ns-2. Simulation results suggest important insights in choosing various parameters of the online algorithm. They also demonstrate that the proposed algorithm indeed achieves low energy consumption and provides satisfactory services to all real-time clients.

The rest of the paper is organized as follows. Section II summarizes existing related work. Section III introduces our system model and the linear programming formulation. Section IV establishes technical backgrounds for developing the online algorithm. Section V introduces our online algorithm, and Section VI studies the performance of the algorithm. Section VII demonstrates our simulation results. Finally, Section VIII concludes the paper.

II. Related Work

There have been several studies about real-time wireless networks. Hou, Borkar, and Kumar [4] investigated the problem of scheduling multiple unicast flows with hard deadlines, and some extensions of this work are included in [5] and [6]. Jaramillo and Srikant [7] studied the real time scheduling problem for wireless networks with both real-time and non-real-time clients. Li and Eryilmaz [8] and Mao and Shroff [11] investigated the real time scheduling problem in multihop wireless networks.

Energy optimal control in wireless has been wildly studied. Neely, Modiano, and Rohrs [13] studied the problem for time varying wireless networks and introducing Lyapunov optimization theory. Lin, Lin, and Shroff [9] developed energy efficient algorithm for multihop wireless networks. Yao and Giannakis [19] proposed a low complexity energy efficient scheduling for wireless sensor network and they investigated the tradeoff between energy consumption and delay. Hossain, Koufos, and Jantti [3] studied the tradeoff between energy and average packet delay. Nevertheless, none of these works take into consideration of networks with hard deadlines.

Berry and Gallager [1] explored the problem of energy-efficient scheduling with delay constraint, nevertheless, their work only considered the wireless channel with a single user. Chen, Mitra, and Neely [2] proposed energy efficient scheduling with individual delay constraints. Their work does not consider the unreliable nature of wireless transmissions. Wang and Li [18] studied the optimal offline policy for power allocation to meet the deadlines of packets.

Stine and Veciana [16] proposed algorithms to improve energy efficiency in centralized wireless networks. Qin et al [14] studied the energy minimizing problem for periodic traffic. Hassan and Assaad [17] provided a solution within a lower bound and a upper bound for the optimal energy consumption in wireless networks with hard deadlines for each packet. Salodkar et al [15] proposed an online energy efficient algorithm under the framework of constrained Markov decision processes. Nevertheless, none of them provided the computational efficiency in their works.

III. System Model and Problem Formulation

Consider a wireless system with N clients with real-time traffic, numbered as $\{1, 2, ..., N\}$, and one access point (AP). Time is slotted with slot $t \in \{1, 2, ...\}$ such that each slot is the duration of one packet transmission. Time slots are further grouped into intervals where each interval consists of T consecutive time slots in $(kT, (k+1)T]$. We assume that each client generate one packet at the beginning of each interval, and each packet has a hard delay bound of T time slots, i.e., packets generated at the beginning of an interval need to be delivered before the end of the interval. Packets that are not delivered on time are dropped.

We consider fading wireless channels and assume that the AP employ some power control algorithm to determine the transmission power for each client. In each interval, the channel is in state $c \in C$ with probability α_c, and changes its state independently at the beginning of the next interval. The AP's power control algorithm determines a transmission power based on the channel state, which in turn determines the link reliability for each client. Under channel state c, each transmission for n consumes an amount $e_{n,c}$ of energy, and the transmission is successful with probability $p_{n,c}$. In the event that a transmission fails, the AP may spend another $e_{n,c}$ energy to retransmit the packet if there are time slots left in the interval.

The performance of a client is measured by its *timely-throughput*, defined as the long-term average number of packets delivered for the client per interval. Each client n requires a hard timely-throughput bound of at least q_n packets per interval.

In this paper, we aim to design scheduling algorithm that minimizes the total power consumption while satisfying the timely-throughput requirements of all clients. We can solve this problem by decomposing different channel states and assigning $q_{n,c}$ as the timely-throughput of client n under a particular channel state c. Since channel state c occurs with probability α_c, the overall timely-throughput of n is $\sum_c \alpha_c q_{n,c}$, and hence we require $\sum_c \alpha_c q_{n,c} \geq q_n$. We also need to evaluate whether it is indeed feasible to provide $q_{n,c}$ timely-throughput for each n, c. Hou et al. [4] have established the following theorems for feasibility:

Theorem 1: Let $w_{n,c}$ be the long-term average number of time slots that the AP transmits for client n under channel state c per interval. The timely-throughput of client n under channel state c is at least $q_{n,c}$ if and only if $w_{n,c} \geq \frac{q_{n,c}}{p_{n,c}}$.

Theorem 2: It is feasible to provide a timely-throughput of $q_{n,c}$ to each client n under channel state c if and only if $\sum_{n \in S} \frac{q_{n,c}}{p_{n,c}} \leq T - I_{S,c}$, for all subsets of clients S. In the formula, $I_{S,c} = \mathbb{E}[(T - \sum_{n\in S} \gamma_n)^+]$, where γ_n is a geometric random variable with mean $1/p_n$, and $x^+ \triangleq \max\{0, x\}$.

By Theorem 1, the total power consumption needed to provide a timely-throughput of $q_{n,c}$ for each n and c is $\sum_{n,c} \frac{q_{n,c}}{p_{n,c}} e_{n,c}\alpha_c$. Therefore, the problem of minimizing power consumption while satisfying timely-throughput requirements can be written as the following optimization problem:

$$\text{Minimize} \quad \sum_{n,c} \frac{q_{n,c}}{p_{n,c}} e_{n,c}\alpha_c \tag{1}$$

$$\text{subject to} \quad \sum_{n\in S} \frac{q_{n,c}}{p_{n,c}} \leq T - I_{S,c}, \quad \forall S,\ \forall c, \tag{2}$$

$$\sum_{c} q_{n,c}\alpha_c \geq q_n, \quad \forall n. \tag{3}$$

While this problem is a linear optimization problem, it can be intractable to solve due to the following reasons: First, the channel state distribution α_c may not be available to the clients and AP. Second, there are exponentially many feasibility constraints in (2). Third, it may be difficult to compute each $I_{S,c}$ directly. However, in the following, we demonstrate that there exists a simple online algorithm that not only satisfies the timely-throughput requirements of all clients, but also achieve a power consumption that is within a constant gap from the optimum.

IV. Preliminaries

Finding an online algorithm for the problem (1) - (3) involves two parts: finding suitable $q_{n,c}$ for all n, c, and finding a scheduling policy that actually provides a timely-throughput of $q_{n,c}$ for each n, c. We call the former problem the *virtual flow control problem* and the latter the *scheduling problem*.

We introduces two variables $q_{n,c}(k)$ and $w_{n,c}(k)$ to capture the AP's solutions to the virtual flow control problem and the scheduling problem, defined as follows. At the beginning of each interval k, the AP observes the current channel state $c(k)$, and it sets the value $q_{n,c(k)}(k) \in [0,1]$. The value of $q_{n,c(k)}(k)$ can be interpreted as the adjustment of $q_{n,c(k)}$. In particular, the AP effectively chooses $q_{n,c} = \liminf_{K\to\infty} \frac{\sum_{k=1}^{K} q_{n,c}(k)}{\sum_{k=1}^{K} \mathbb{I}_{c(k)=c}}$ as its solutions to the problem (1) - (3), where $q_{n,c}(k) = 0$, if $c \neq c(k)$, and $\mathbb{I}$ is the indicator function.

The AP also determines a scheduling policy of transmitting packets in interval k. We let $w_{n,c(k)}(k)$ be the number of time slots that the AP transmits for client n in interval k. Recall that each transmission for n is successful with probability $p_{n,c}$, and there is only one packet to be delivered for n in each interval. Hence, the value of $w_{n,c(k)}(k)$ is determined by not only the scheduling policy but also random events in interval k. We set $w_{n,c}(k) = 0$ if $c \neq c(k)$. By the definition of $I_{S,c}$, we have that $\mathbb{E}[\sum_{n\in S} w_{n,c(k)}(k)] \leq T - I_{S,c(k)}$, for all $S \subseteq \{1, 2, \ldots, N\}$, under any scheduling policy.

To capture whether the solutions to the virtual flow control problem and the scheduling problem satisfy the timely-throughput requirements of all clients, we define two *virtual queues* as follows:

$$H_{S,c}(k+1) = H_{S,c}(k) + \sum_{n\in S} \frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k), \quad (4)$$

and

$$Q_n(k+1) = [Q_n(k) + q_n - q_{n,c(k)}(k)]^+, \quad (5)$$

with $H_{S,c}(0) = 0$, and $Q_n(0) = 0$, for all S, n, and c.

We note the definitions of virtual queues are very similar to those of Lagrange multipliers corresponding to (2) and (3), as used in many online algorithms based on dual decomposition [10] or Lyapunov optimization [12]. However, using the standard definition of Lagrange multipliers, we should have defined $H_{S,c}(k+1) = [H_{S,c}(k) + \sum_{n\in S} \frac{q_{n,c}(k)}{p_{n,c}} - (T - I_{S,c(k)})]^+$, which involves the computation of $I_{S,c}$. On the other hand, our definition of $H_{S,c}$ in (4) is very easy to calculate, as it is solely based on the control decisions of the AP and random events in the interval. Further, we have

Lemma 1: If $\limsup_{K\to\infty} \mathbb{E}[|H_{S,c}(K)|]/K = 0$, and $\limsup_{K\to\infty} \mathbb{E}[Q_n(K)]/K = 0$, for all S, c and n, then the timely-throughput of client n is at least q_n. Further, the vector $[q_{n,c}] \triangleq [\liminf_{K\to\infty} \frac{\sum_{k=1}^{K} q_{n,c}(k)}{\sum_{k=1}^{K} \mathbb{I}_{c(k)=c}}]$ satisfies constraints in (2) and (3), and the total power consumption is at most $\sum_c \alpha_c e_{n,c} \frac{\bar{q}_{n,c}}{p_{n,c}}$, where $[\bar{q}_{n,c}] \triangleq [\limsup_{K\to\infty} \frac{\sum_{k=1}^{K} q_{n,c}(k)}{\sum_{k=1}^{K} \mathbb{I}_{c(k)=c}}]$.

Proof: Fix $S = \{n\}$, then we have

$$\begin{aligned}
0 &= \limsup_{K\to\infty} \frac{\mathbb{E}[|H_{\{n\},c}(K)|]}{K} \\
&\geq \limsup_{K\to\infty} \frac{1}{K}\left(\sum_{k=0}^{K-1} \frac{\mathbb{E}[q_{n,c}(k)]}{p_{n,c}} - \sum_{k=0}^{K-1} \mathbb{E}[w_{n,c}(k)]\right), \\
&\Rightarrow \liminf_{K\to\infty} \frac{\sum_{k=0}^{K-1} \mathbb{E}[w_{n,c}(k)]}{K} \geq \liminf_{K\to\infty} \sum_{k=0}^{K-1} \frac{\mathbb{E}[q_{n,c}(k)]}{p_{n,c}}/K \\
&\Rightarrow w_{n,c} \geq \frac{q_{n,c}}{p_{n,c}}.
\end{aligned}$$

Therefore, the timely-throughput of n under c is at least $q_{n,c}$, by Theorem 1. We also have

$$\begin{aligned}
&0 = \limsup_{K\to\infty} \frac{\mathbb{E}[Q_n(K)]}{K} \geq q_n - \frac{1}{K}\mathbb{E}[\sum_{k=0}^{K-1} q_{n,c(k)}(k)] \\
&\Rightarrow q_n \leq \sum_c \alpha_c q_{n,c},
\end{aligned}$$

and the overall timely-throughput of n is at least q_n.

On the other hand,

$$\begin{aligned}
0 &= \limsup_{K\to\infty} \frac{\mathbb{E}[|H_{\{n\},c}(K)|]}{K} \\
&\geq \limsup_{K\to\infty} \frac{1}{K}\left(\sum_{k=0}^{K-1} \mathbb{E}[w_{n,c}(k)] - \sum_{k=0}^{K-1} \frac{\mathbb{E}[q_{n,c}(k)]}{p_{n,c}}\right), \\
&\Rightarrow \limsup_{K\to\infty} \frac{\sum_{k=0}^{K-1} \mathbb{E}[w_{n,c}(k)]}{K} \leq \limsup_{K\to\infty} \sum_{k=0}^{K-1} \frac{\mathbb{E}[q_{n,c}(k)]}{p_{n,c}}/K.
\end{aligned}$$

The total power consumption is then at most $\sum_{n,c} e_{n,c} \limsup_{K\to\infty} \frac{\sum_{k=0}^{K-1} \mathbb{E}[w_{n,c}(k)]}{K} \leq \sum_c \alpha_c e_{n,c} \frac{\bar{q}_{n,c}}{p_{n,c}}$.

Finally, we show that $\sum_{n\in S} \frac{q_{n,c}}{p_{n,c}} \leq T - I_{S,c}$. For any S, c, we have

$$\begin{aligned}
0 &= \limsup_{K\to\infty} \frac{\mathbb{E}[|H_{S,c}(K)|]}{K} \\
&\geq \limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}[\sum_{n\in S} \frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k)] \\
&\geq \limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}[(\sum_{n\in S} \frac{q_{n,c}(k)}{p_{n,c}} - (T - I_{S,c}))\mathbb{I}_{c(k)=c}] \\
&\Rightarrow \sum_{n\in S} \frac{q_{n,c}}{p_{n,c}} \leq T - I_{S,c}.
\end{aligned}$$

■

We say that $H_{S,c}(k)$ and $Q_n(k)$ are *mean rate stable* if $\limsup_{K\to\infty}\mathbb{E}[|H_{S,c}(K)|]/K = 0$, and $\limsup_{K\to\infty}\mathbb{E}[Q_n(K)]/K = 0$.

V. Online Algorithms

In this section, we develop an online algorithm that jointly solves the virtual flow control problem and the scheduling problem based on Lyapunov theory. Let $\Theta(t) \triangleq [H(k),\ Q(k)]$ be a collective vector of the virtual queues $H(k) = [H_{S,c}(k)]$, and $Q(k) = [Q_n(k)]$. Define a Lyapunov function as:

$$L(\Theta(k)) \triangleq \frac{1}{2}\left(\delta\sum_{S,c} H^2_{S,c}(k) + \sum_n Q^2_n(k)\right), \qquad (6)$$

where $\delta \geq 0$ is a predefined parameter. We have the following lemma:

Lemma 2: At each time interval k, for any control decisions applied in the network and constant V, we have the following inequality:

$$\begin{aligned}
&\mathbb{E}[L(\Theta(k+1)) - L(\Theta(k))|\Theta(k)] \\
&+ V\mathbb{E}[\sum_{n,c}\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}|\Theta(k),\ c(k)] \\
\leq\ &\mathbb{E}[\delta\sum_{S,c} H_{S,c(k)}(k)(\sum_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k))|\Theta(k),\ c(k)] \\
&+ \mathbb{E}[\sum_n Q_n(k)(q_n - q_{n,c(k)}(k))|\Theta(k),\ c(k)] \\
&+ V\mathbb{E}[\sum_{n,c}\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}|\Theta(k),\ c(k)] + B,
\end{aligned} \qquad (7)$$

where B is a finite constant.

Proof: First, we have:

$$\begin{aligned}
&H^2_{S,c}(k+1) - H^2_{S,c}(k) \\
=\ &(H_{S,c}(k) + (\sum_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c(k)}(k)))^2 - H^2_{S,c}(k) \\
=\ &(\sum_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k))^2 \\
&+ 2H_{S,c}(k)(\sum_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k)) \\
\leq\ &B_1 + 2H_{S,c}(k)(\sum_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}} - \sum_{n\in S} w_{n,c}(k)),
\end{aligned} \qquad (8)$$

where the last inequality holds because both $\sum\limits_{n\in S}\frac{q_{n,c}(k)}{p_{n,c}}$ and $\sum\limits_{n\in S} w_{n,c}(k)$ are bounded. Similarly, we also have:

$$\begin{aligned}
&Q^2_n(k+1) - Q^2_n(k) \\
=\ &([Q_n(k) + q_n - q_{n,c(t)}(k)]^+)^2 - Q^2_n(k) \\
\leq\ &(Q_n(k) + q_n - q_{n,c(k)}(k))^2 - Q^2_n(k) \\
=\ &(q_n - q_{n,c(k)}(k))^2 + 2Q_n(k)(q_n - q_{n,c(k)}(k)) \\
\leq\ &B_2 + 2Q_n(k)(q_n - q_{n,c(k)}(k)),
\end{aligned} \qquad (9)$$

where the last inequality holds because $q_{n,c(k)}(k)$ is bounded. Summing (8) and (9) proves the lemma. ■

The first term in the left hand side (LHS) of (7) is known as the Lyapunov drift. The second term can be explained as follows: if $H_{S,c}(k)$ and $Q_n(k)$ are mean rate stable for all S, n and c, then, as shown in Lemma 1, the long-term average of $\sum_{n,c}\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}$ is the average power consumption. We therefore refer the LHS of (7) as "drift plus weighted power consumption".

We aim to reduce the LHS of (7) by minimizing the right hand side (RHS) of (7). We will prove in Section VI that minimizing the RHS of (7) for all k ensures that $H_{S,c}(k)$ and $Q_n(k)$ are mean rate stable, and the average power consumption is within a constant difference from the optimum.

Given $\Theta(k)$ and $c(k)$, the RHS of (7) can be rewritten as

$$\begin{aligned}
&\text{RHS of (7)} \\
=&\sum_n \mathbb{E}[\frac{q_{n,c(k)}}{p_{n,c(k)}}(\sum_{S:n\in S}\delta H_{S,c(k)}(k) + Ve_{n,c(k)} - Q_n(k)p_{n,c(k)})] \\
&- \delta\sum_n \mathbb{E}[w_{n,c(k)}\sum_{S:n\in S} H_{S,c(k)}(k)] + \mathbb{E}[\sum_n Q_n(k)q_n] + B.
\end{aligned}$$

With this representation of the RHS of (7), it is clear that the following policies minimize the RHS of (7):

Virtual Flow Control: At each time interval k, the AP sets $q_{n,c(k)}(k)$ as:

$$q_{n,c(k)}(k) = \mathbb{I}_{\{\sum_{S:n\in S}\delta H_{S,c(k)}(k) + Ve_{n,c(k)} \leq Q_n(k)p_{n,c(k)}\}} \qquad (10)$$

Scheduling: At each time interval k, the AP employs a scheduling policy that maximizes

$$\mathbb{E}[\sum_n\sum_{S:n\in S} H_{S,c(k)}(k)w_{n,c(k)}(k)]. \qquad (11)$$

To employ such policies, it seems that the AP needs to track the exponentially many variables $[H_{S,k}]$. Instead, we let $D_{n,c}(k) = \sum\limits_{S:n\in S} H_{S,c}(k)$, and we have the following evolution of $D_{n,c}(k)$:

$$\begin{aligned}
D_{n,c}(k+1) =\ &D_{n,c}(k) + \sum_{S:n\in S}(\sum_{m\in S}\frac{q_{m,c}(k)}{p_{m,c}} - \sum_{m\in S} w_{m,c}(k)) \\
=\ &D_{n,c}(k) + 2^{N-1}(\frac{q_{n,c}(k)}{p_{n,c}} - w_{n,c}(k)) \\
&+ 2^{N-2}(\sum_{m\neq n}\frac{q_{m,c}(k)}{p_{m,c}} - \sum_{m\neq n} w_{m,c}(k)).
\end{aligned} \qquad (12)$$

The last equality holds as there are 2^{N-1} subsets that contain n, and 2^{N-2} subsets that contain both n and m, for some $m \neq n$. We can then rewrite (10) and (11) as $q_{n,c(k)}(k) = \mathbb{I}_{\{\delta D_{n,c(k)}(k) + Ve_{n,c(k)} \leq Q_n(k)p_{n,c(k)}\}}$ and $\mathbb{E}[\sum_n D_{n,c(k)}(k)w_{n,c(k)}(k)]$, respectively.

We propose employing the *largest debt first policy*, which was proposed in [4] in a different context, to maximize $\mathbb{E}[\sum_n D_{n,c(k)}(k)w_{n,c(k)}(k)]$. In the largest debt first policy, the AP sorts all clients in descending order of $D_{n,c(k)}$ and schedules transmissions according to this ordering. In other words, a client n is only scheduled by the AP if the packets for all clients m with $D_{m,c(k)} > D_{n,c(k)}$ have been delivered. Clients with negative $D_{n,c(k)}$ are not scheduled by the AP.

Theorem 3: The largest debt first policy maximizes $\mathbb{E}[\sum_n D_{n,c}(k)w_{n,c}(k)|\Theta(k),\ c(k)]$ in each interval k.

Proof: Without loss of generality, assume $D_{1,c}(k) \geq D_{2,c}(k) \geq \cdots \geq D_{M,c}(k) \geq 0 \geq \ldots$. Let γ_n be the random variable denoting the number of transmissions the AP has to make for client n before a successful transmission. Recall that $w_{n,c(k)}(k)$ is the number of transmissions that the AP actually makes for client n during the interval. Assuming that $[\gamma_n]$ is known, maximizing $\mathbb{E}[\sum_n D_{n,c}(k)w_{n,c}(k)]$ reduces to solving the following linear programming problem:

$$\begin{aligned} \text{Max } & \sum_n D_{n,c(k)}(k)w_{n,c(k)}(k) \\ \text{s.t. } & 0 \leq w_{n,c(k)}(k) \leq \gamma_n, \forall n \\ & \sum_n w_{n,c(k)}(k) \leq T. \end{aligned}$$

One obvious solution is to allocate the first γ_1 time slots to client 1, the next γ_2 time slots to client 2, etc., until all the T time slots are allocated or all packets for clients 1 through M are delivered. This solution is consistent with the largest debt first policy. ∎

Our joint algorithm for virtual flow control and scheduling is summarized in Alg. 1. We note that our joint algorithm has the following properties: First, its complexity is $O(N \log N + T)$ per interval. Second, it does not need to know the distributions of channel states, α_c. Third, it does not need to compute the values of $I_{S,c}$.

Algorithm 1 Virtual flow control and scheduling

$Q_n \leftarrow 0, \forall n$
$D_{n,c} \leftarrow 0, \forall n, c$
for each interval **do**
  $c \leftarrow$ current channel
  $q_n \leftarrow \mathbb{I}_{\{\delta D_{n,c} + V e_{n,c} \leq Q_n p_{n,c}\}}, \forall n$
  $w_n \leftarrow 0, \forall n$
  Sort all clients such that $D_{1,c} \geq D_{2,c} \geq \ldots$
  $m \leftarrow 0$
  for each time slot in the interval **do**
    if $D_{m,c} > 0$ **then**
      Transmit for client m
      $w_m \leftarrow w_m + 1$
      if the transmission is successful **then**
        $m \leftarrow m + 1$
      end if
    end if
  end for
  $Tot \leftarrow 2^{N-2}(\sum_n \frac{q_n}{p_{n,c}} - \sum_n w_n)$
  $D_{n,c} \leftarrow D_{n,c} + 2^{N-2}(\frac{q_n}{p_{n,c}} - w_n) + Tot, \forall n$
end for

VI. Performance Analysis

In this section, we study the performance of Alg. 1.

Theorem 4: For a controlled parameter $V > 0$, the values of $q_{n,c}(k)$ and $w_{n,c}(k)$ under Alg. 1 satisfy:

$$\sum_n \frac{\bar{q}_{n,c}}{p_{n,c}} e_{n,c}\alpha_c \leq \sum_{n,c} \frac{q^*_{n,c}}{p_{n,c}} e_{n,c}\alpha_c + \frac{B}{V}, \tag{13}$$

where $\bar{q}_{n,c} = \limsup_{K\to\infty} \frac{\sum_{k=0}^{K-1} \mathbb{E}[q_{n,c}(k)]}{\sum_{k=0}^{K-1} \mathbb{I}_{\{c(k)=c\}}}$ and $q^*_{n,c}$ is the optimal solution for the problem (1)-(3).

Proof: Since Alg. 1 minimizes the RHS of (7), taking the expected value under Alg. 1 on the LHS of (7) over all $\Theta(k), c(k)$ yields:

$$\begin{aligned} & \mathbb{E}[L(\Theta(k+1))] - \mathbb{E}[L(\Theta(k))] + V\mathbb{E}[\sum_{n,c} \frac{q_{n,c}(k)}{p_{n,c}} e_{n,c}] \\ \leq\ & B + \mathbb{E}[\sum_{n,c} D_{n,c}(k)(\frac{q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}}{p_{n,c}} - w_{n,c}(k))] \\ & + \mathbb{E}[\sum_n Q_n(k)(q_n - \sum_c q^*_{n,c}\mathbb{I}_{\{c(k)=c\}})] \\ & + V\mathbb{E}[\sum_{n,c} \frac{q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}}{p_{n,c}} e_{n,c}] \end{aligned} \tag{14}$$

Since $q^*_{n,c}$ is the optimal solution for the problem (1)-(3), we have $\mathbb{E}[q_n - \sum_c q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}] \leq 0$, for all n. We also have $\mathbb{E}[\sum_{n,c} \frac{q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}}{p_{n,c}} e_{n,c}] = \sum_{n,c} \frac{q^*_{n,c}}{p_{n,c}} e_{n,c}\alpha_c$, as channel state c occurs with probability α_c. We can then further simplify the above inequality as

$$\begin{aligned} & \mathbb{E}[L(\Theta(k+1))] - \mathbb{E}[L(\Theta(k))] + V\mathbb{E}[\sum_{n,c} \frac{q_{n,c}(k)}{p_{n,c}} e_{n,c}] \\ \leq\ & B + \mathbb{E}[\sum_{n,c} D_{n,c}(k)(\frac{q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}}{p_{n,c}} - w_{n,c}(k))] \\ & + V\sum_{n,c} \frac{q^*_{n,c}}{p_{n,c}} e_{n,c}\alpha_c \\ =\ & B + \mathbb{E}[\sum_{n,c} [D_{n,c}(k)]^+(\frac{q^*_{n,c}\mathbb{I}_{\{c(k)=c\}}}{p_{n,c}} - w_{n,c}(k))] \\ & + V\sum_{n,c} \frac{q^*_{n,c}}{p_{n,c}} e_{n,c}\alpha_c, \end{aligned} \tag{15}$$

where the last equality holds because $w_{n,c}(k) = 0$ if $D_{n,c}(k) < 0$ by Alg. 1.

Since $[q^*_{n,c}]$ satisfies the constraints (2) and (3), there exists a randomized stationary policy that makes scheduling decisions solely based on the channel state and provides a timely-throughput of $q^*_{n,c}$ for each n and c. Let $w^*_{n,c(k)}(k)$ be the number of time slots that the AP transmits for n in interval k, and $w^*_{n,c}(k) = 0$ if $c \neq c(k)$. We then have $\mathbb{E}[\hat{w}_{n,c}(k)] \geq \frac{q^*_{n,c}}{p_{n,c}}\alpha_c$, and:

$$\begin{aligned}
&\mathbb{E}[L(\Theta(k+1))] - \mathbb{E}[L(\Theta(k))] + V\mathbb{E}[\sum_{n,c}\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}] \\
\leq\ & B + \sum_{n,c}\mathbb{E}[[D_{n,c}(k)]^+](\frac{q^*_{n,c}}{p_{n,c}}\alpha_c - \mathbb{E}[\hat{w}_{n,c}(k)]) \\
&+ V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c \\
\leq\ & B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c. \qquad (16)
\end{aligned}$$

Summing the above inequality over $k \in \{0, 1, ..., K-1\}$ and dividing by K yields:

$$\begin{aligned}
&\frac{\mathbb{E}[L(\Theta(K))] - \mathbb{E}[L(\Theta(0))]}{K} + \frac{V}{K}\sum_{k=0}^{K-1}\sum_{n,c}\mathbb{E}[\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}] \\
\leq\ & B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c. \qquad (17)
\end{aligned}$$

Since $L(\Theta(0)) = 0$ and $L(\Theta(k)) \geq 0$, we further have

$$\begin{aligned}
&\limsup_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\sum_{n,c}\mathbb{E}[\frac{q_{n,c}(k)}{p_{n,c}}e_{n,c}] \\
=\ & \sum_{n,c}\limsup_{K\to\infty}\frac{\sum_{k=0}^{K-1}\mathbb{E}[q_{n,c}(k)]e_{n,c}}{\sum_{k=0}^{K-1}\mathbb{I}_{\{c(k)=c\}}p_{n,c}}\frac{\sum_{k=0}^{K-1}\mathbb{I}_{\{c(k)=c\}}}{K} \\
=\ & \sum_{n,c}\frac{\bar{q}_{n,c}}{p_{n,c}}e_{n,c}\alpha_c \\
\leq\ & \sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c + \frac{B}{V}. \qquad (18)
\end{aligned}$$

■

Theorem 5: The virtual queues $[H_{S,c}(k)]$ and $[Q_n(k)]$ under Alg. 1 are mean rate stable.

Proof: Since $q_{n,c}(k) \geq 0$ for all n, c, and k, from (16), we get:

$$\begin{aligned}
&\frac{1}{2}\left(\sum_{S,c}\mathbb{E}[H^2_{S,c}(k)] + \sum_n\mathbb{E}[Q^2_n(k)]\right) = \mathbb{E}[L(\Theta(k))] \\
\leq& \mathbb{E}[L(\Theta(0))] + (B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c)k, \qquad (19)
\end{aligned}$$

and hence

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}[H^2_{S,c}(k)] \leq \mathbb{E}[L(\Theta(0))] + (B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c)k \\
\Rightarrow& \mathbb{E}[|H_{S,c}(k)|] \\
&\leq \sqrt{2\mathbb{E}[L(\Theta(0))] + 2(B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c)k}.
\end{aligned}$$

Dividing by k and let $k \to \infty$ yields:

$$\begin{aligned}
&\lim_{k\to\infty}\frac{\mathbb{E}[|H_{S,c}(k)|]}{k} \\
\leq& \lim_{k\to\infty}\frac{\sqrt{2\mathbb{E}[L(\Theta(0))] + 2(B + V\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c)k}}{k} = 0.
\end{aligned} \qquad (20)$$

The proof with $Q_n(k)$ is similar, thus, all virtual queues $[H_{S,c}(k)]$ and $[Q_n(k)]$ are mean rate stable. ■

Finally, since $Q_n(k)$ and $H_{S,c}(k)$ are mean rate stable under Alg. 1, we can establish the following theorem, which is a direct result of Lemma 1 and Theorem 4.

Theorem 6: Alg. 1 provides a timely-throughput of q_n to each client n, and its total power consumption is at most $\sum_{n,c}\frac{q^*_{n,c}}{p_{n,c}}e_{n,c}\alpha_c + \frac{B}{V}$.

VII. Simulation Results

We have implemented Alg. 1 in ns-2. In this section, we demonstrate our simulation results. We consider a WiFi system where an AP serves 10 real-time clients. We set the length of an interval to be 20ms, which consists of 32 time slots when the AP transmits at 11 Mb/s. Client n requires a timely-throughput of $(70+n)\%$.

We assume that there are 10 channel states, and the AP may, or may not, employ power control algorithms. When power control algorithms are not used, the AP transmits all packets with unit energy, $e_{n,c} = 1$. The channel reliability of client n under channel state c is $(30+10\times((c+n) \mod 7))\%$. On the other hand, when the AP employs power control algorithms, it chooses proper transmission powers such that the channel reliabilities of all clients are $p_{c,n} = 80\%$. The transmission power used for client n under channel state c is $((c+n) \mod 4+1)$.

We measure the performance of various policies with two metrics: average total power consumption and average total deficiency. The deficiency of a client n is defined as $(q_n - \text{actual timely-throughput of } n)^+$. Total deficiency is then defined as the sum of deficiencies of all clients. All simulation results are the average of five runs.

A. The Choice of δ

We first study the choice of the control parameter δ. We chose various values of δ and study the total deficiency. As we are not concerned about power consumption in this simulation, we set $V = 0$, and consider the scenario where the AP uses a fixed transmission

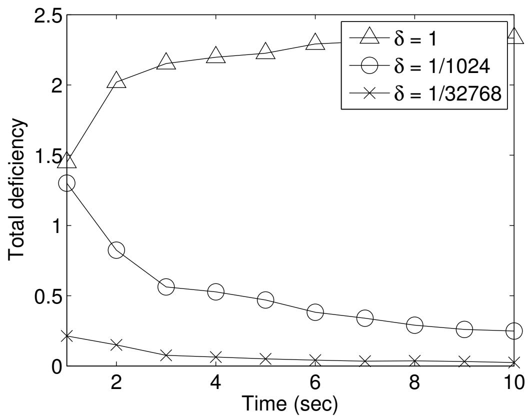


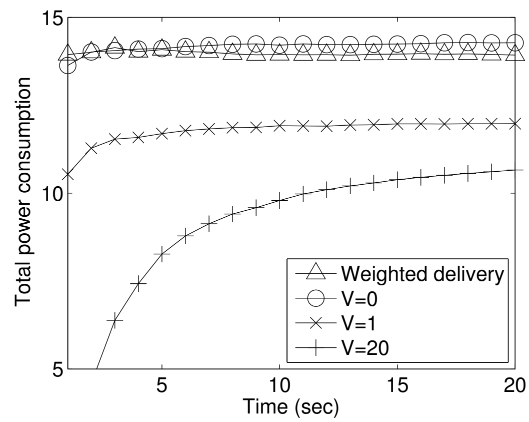
Fig. 1: Total deficiency under various δ.

power. As shown in Fig. 1, setting $\delta = 1$ results in poor performance. The total deficiency remains above 2. In fact, it takes more than 300 seconds for the total deficiency to be below 2. The reason that setting $\delta = 1$ results in such poor performance is because $D_{n,c}(k)$ and $Q_n(k)$ change at different rates. By definition, $|D_{n,c}(k+1) - D_{n,c}(k)|$ can be on the order of $2^N T$. On the other hand, $|Q_n(k+1) - Q_n(k)|$ is bounded by 1. With this observation, we envision that setting $\delta = \frac{1}{2^N T}$ and making $\delta D_{n,c}(k)$ and $Q_n(k)$ change at similar rates may lead to better performance. As shown in Fig. 1, choosing $\delta = \frac{1}{2^N T} = 1/32768$ indeed achieves the best performance, and the total deficiency converges to 0 very quickly. Therefore, we set $\delta = 1/32768$ throughout the rest of the section.

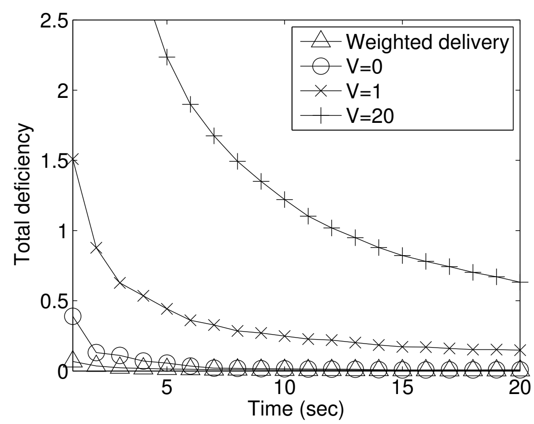
B. Performance Comparison

We compare the performance of Alg. 1 for various values of V. In addition, we also compare the performance of the weighted delivery policy, which was proposed in [6] and proved to satisfy the requirements of all clients under fading wireless channels as long as they are feasible.

Simulation results when the AP uses a fixed transmission power, and when the AP employs power control algorithms, are presented in Fig. 2 and Fig. 3, respectively. Not surprisingly, the weighted delivery policy and our policy with $V = 0$ have the worst total power consumption, as these two policies do not consider power consumption and only focus on satisfying clients' requirements. Setting $V = 1$ can significantly reduce power consumption, and setting $V = 20$ can further reduce it. On the other hand, when V is large, the total deficiency converges to zero very slowly. Hence, the choice of V is indeed a tradeoff between power consumption and short-term system performance. To achieve both low power consumption and fast convergence in total deficiency, one may choose V to be a variable that slowly increases with a rate of $o(k)$.



(a) Total power consumption



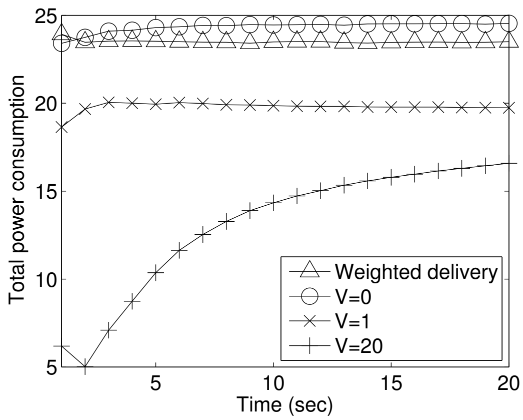
(b) Total deficiency

Fig. 2: Simulation results when the AP uses a fixed transmission power.

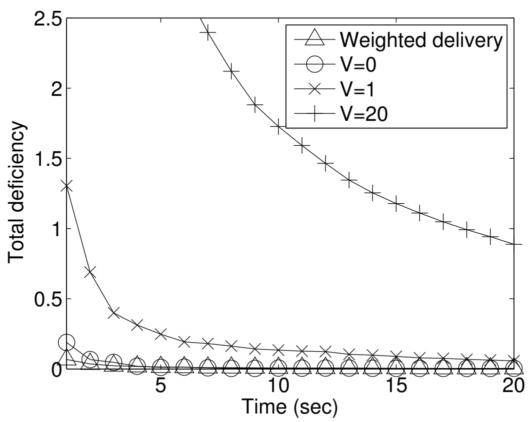
VIII. Conclusion

We have studied the problem of providing timely-throughput guarantees to real-time clients with minimum power consumption. While this problem can be formulated as a linear programming problem, such a formulation involves exponentially many constraints, the knowledge of distributions of channel states, and the computation of some complicated parameters. To concur these challenges, we have proposed a simple algorithm that jointly solves the virtual flow control problem and the scheduling problem. We have proved that this algorithm satisfies the timely-throughput requirement of each client, and its power consumption can be made arbitrarily close to optimum.

We have implemented and evaluated the performance of our algorithm in ns-2. Simulation results provide insights in choosing appropriate parameters for our



(a) Total power consumption



(b) Total deficiency

Fig. 3: Simulation results when the AP employs power control algorithms.

algorithm to converge quickly. Simulation results also show that our algorithm achieve a substantial reduction on total power consumption compared to a policy that only aims to satisfy the timely-throughput requirements of all real-time clients.

References

[1] Berry, R. A., and Gallager, R. G. Communication over fading channels with delay constraints. *IEEE Trans. on Information Theory 48*, 5 (May 2002), 1135–1149.

[2] Chen, W., Mitra, U., and Neely, M. Energy-efficient scheduling with individual delay constraints over a fading channel. In *Prof. of WiOpt* (Apr 2007), pp. 1 –10.

[3] Hossain, M. A., Koufos, K., and Jantti, R. Minimum-energy power and rate control for fair scheduling in the cellular downlink under flow level delay constraint. *Wireless Communications, IEEE Transactions on 12*, 7 (2013), 3253–3263.

[4] Hou, I.-H., Borkar, V., and Kumar, P. A theory of QoS in wireless. In *Proc. of IEEE INFOCOM* (2009), pp. 486–494.

[5] Hou, I.-H., and Kumar, P. Admission control and scheduling for qos guarantees for variable-bit-rate applications on wireless channels. In *Proc. of MOBIHOC* (2009).

[6] Hou, I.-H., and Kumar, P. Scheduling heterogeneous real-time traffic over fading wireless channels. In *Proc. of IEEE INFOCOM* (2010).

[7] Jaramillo, J., and Srikant, R. Optimal scheduling for fair resource allocation in ad hoc networks with elastic and inelastic traffic. In *Proc. of IEEE INFOCOM* (Mar. 2010), pp. 1 –9.

[8] Li, R., and Eryilmaz, A. Scheduling for end-to-end deadline-constrained traffic with reliability requirements in multi-hop networks. In *Proc. of IEEE INFOCOM* (Apr. 2011), pp. 3065 –3073.

[9] Lin, L., Lin, X., and Shroff, N. B. Low-complexity and distributed energy minimization in multihop wireless networks. *IEEE/ACM Trans. on Networking 18*, 2 (Apr. 2010), 501 – 514.

[10] Lin, X., Shroff, N., and Srikant, R. A tutorial on cross-layer optimization in wireless networks. *Selected Areas in Communications, IEEE Journal on 24*, 8 (2006), 1452–1463.

[11] Mao, Z., Koksal, C. E., and Shroff, N. B. Online packet scheduling with hard deadlines in multihop communication networks. In *Proc. of IEEE INFOCOM* (2013).

[12] Neely, M. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan and Claypool, 2010.

[13] Neely, M. J., Modiano, E., and Rohrs, C. E. Dynamic power allocation and routing for time varying wireless networks. *IEEE Journal on Selected Areas in Communications 23*, 1 (Jan 2005), 89 –103.

[14] Qin, X., Alghamdiand, M., Nijim, M., Zong, Z., and Bellam, K. Scheduling of periodic packets in energy-aware wireless networks. In *Proc. of IPCCC* (Apr 2007), pp. 210 –217.

[15] Salodkar, N., Karandikar, A., and Borkar, V. A stable online algorithm for energy-efficient multiuser scheduling. *IEEE Trans. on Mobile Computing 9*, 10 (Oct. 2010), 1391 – 1406.

[16] Stine, J. A., and Veciana, G. D. Improving energy efficiency of centrally controlled wireless data networks. *IEEE Trans. on Information Theory 8*, 6 (Nov 2002), 681–700.

[17] Ul Hassan, N.and Assaad, M. Energy efficient causal packet scheduling in wireless fading channels with hard delay constraints. In *Prof. of WCNC* (Apr 2010), pp. 1 –5.

[18] Wang, X., and Li, Z. Energy-efficient transmissions of bursty data packets with strict deadlines over time-varying wireless channels. *Wireless Communications, IEEE Transactions on 12*, 5 (2013), 2533–2543.

[19] Yao, Y., and Giannakis, G. B. Energy-efficient scheduling for wireless sensor networks. *IEEE Trans. on Communications 53*, 8 (2005).