Optimal Universal Schedules for Discrete Broadcast

Michael Langberg, Alexander Sprintson, *Member, IEEE*, and Jehoshua Bruck, *Fellow, IEEE*

***Abstract*—We study broadcast systems that distribute a series of data updates to a large number of passive clients. The updates are sent over a broadcast channel in the form of discrete packets. We assume that clients periodically access the channel to obtain the most recent update. Such scenarios arise in many practical applications, such as distribution of traffic information and market updates to mobile wireless devices.**

Our goal is to design broadcast schedules that minimize the *waiting time*, i.e., the amount of time the client needs to wait in order to obtain the most recent update. We assume that each client has a different access pattern depending on the channel conditions, computing power, and storage capabilities. We introduce and analyze optimal *universal schedules* that guarantee low waiting time for any client, regardless of its behavior.

***Index Terms*—Broadcast systems, data updates, universal schedules, waiting time.**

I. Introduction

High availability and low cost of wireless mobile devices have sparked a renewed interest in data broadcast systems [8], [20], [25]. Such systems facilitate ubiquitous information access by enabling a large number of users to receive dynamic information in a scalable and efficient way, minimizing power consumption and keeping the clients' locations secret.

In this correspondence, we focus on systems that enable each client to access a series of data updates. Such systems can benefit a variety of wireless applications that need, for example, a constant access to the stock market information or the current traffic conditions. In particular, we focus on systems that contain a large number of heterogeneous clients with different access patterns.

Fig. 1 depicts an example of a data broadcast system we consider. The system includes four major components: a dynamic database, a server (scheduler), a dedicated broadcast channel, and a large number of wireless clients. The server periodically accesses the database, retrieves the most recent data, encapsulates it into packets, and sends the packets over the broadcast channel. We assume that each client continuously performs the following operations.

(i) Tune in to the channel to obtain the most recent update.
(ii) Process the obtained update and go to step (i).

These operation allow the client to maintain up-to-date local database to meet the requirements of the application.

The processing step may include several operations such as the decryption of the received packet, updating the internal database, and

Manuscript received August 4, 2006; revised January 17, 2008. Published August 27, 2008 (projected). The work of J. Bruck was supported in part by the Caltech Lee Center for Advanced Networking and by the National Science Foundation under Grants ANI-0322475 and CCF-0346991. Part of this work was performed while M. Langberg and A. Sprintson were Postdoctoral Scholars at the California Institute of Technology, Pasadena.

M. Langberg is with the Computer Science Division, The Open University of Israel, Raanana 43107, Israel (e-mail: mikel@openu.ac.il).

A. Sprintson is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: spalex@tamu.edu).

J. Bruck is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: bruck@caltech.edu).

Communicated by E. Modiano, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2008.928296

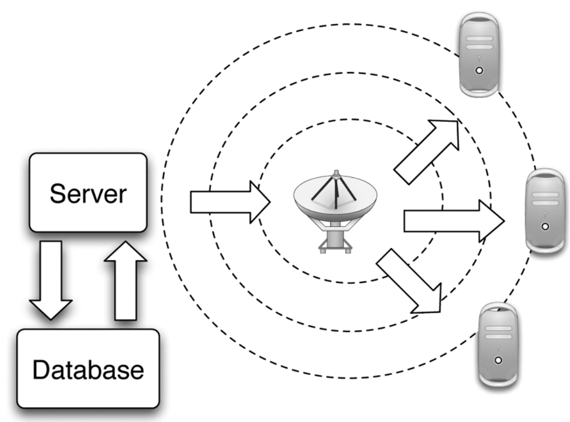

Fig. 1. A broadcast data distribution system.

waiting for the user input. In addition, if the channel conditions are poor, the client needs to wait until the conditions improve before it will be able to receive the next packet. We refer to the time interval required for a client to get ready to receive the next update as the *idle time*. Since the duration of this interval is affected by many factors, it can vary over time. Moreover, the idle time can vary for different clients.

Once a client is ready to receive a new update, it tunes in to the channel. This correspondence focuses on the design of broadcast schedules that minimize the amount of time the client needs to wait (after it is ready) to obtain the desired update. We refer to this time as the *waiting time*. Constructing such a schedule is a difficult task, due to the diversity of the clients' access patterns. For example, consider the broadcast schedule depicted on Fig. 2(a). Here, and throughout the correspondence, we assume that all updates are carried by packets of equal size and that the client must listen to at least one packet, from beginning to end, to obtain an update. Fig. 2(b) and (c) shows the waiting time of two clients, with different idle times. Note that while the waiting time is low for the first client, the second client needs to wait a significant amount of time, because it misses the first and the third updates.

If some information about the access pattern is available to the server, the server can take advantage of this information in order to design a better schedule. For example, suppose that the idle time of each client is distributed uniformly in the interval $[0, 1]$. In this case, the optimal strategy is to transmit one packet after the other (see Fig. 3(a)). However, this schedule does not behave well in the presence of heterogeneous clients. Indeed, for clients that have a small, but nonzero idle time, the waiting time is very close to one time unit. The goal of this work is to design *universal schedules* that minimize the *worst case* waiting time for any client, regardless of its access pattern. We assume that the server has no information about the idle times of the clients. In this scenario, our goal is to schedule the transmission of the packets in such a way that for any given values of idle times, the maximum waiting time of the client will be as small as possible (here we are maximizing over the first and the subsequent updates). More specifically, the worst case waiting time of the schedule in Fig. 3(a) can be shown to be close to one time unit, thus in this work we seek to design schedules with a significantly lower waiting time. In order to analyze the worst case performance of the schedule, we model the data update process as a game against an *adversarial client*. The adversarial client can select the values of idle times which result in high values of waiting time. Clearly, a schedule that performs well against such a powerful adversary, will perform well for any client.

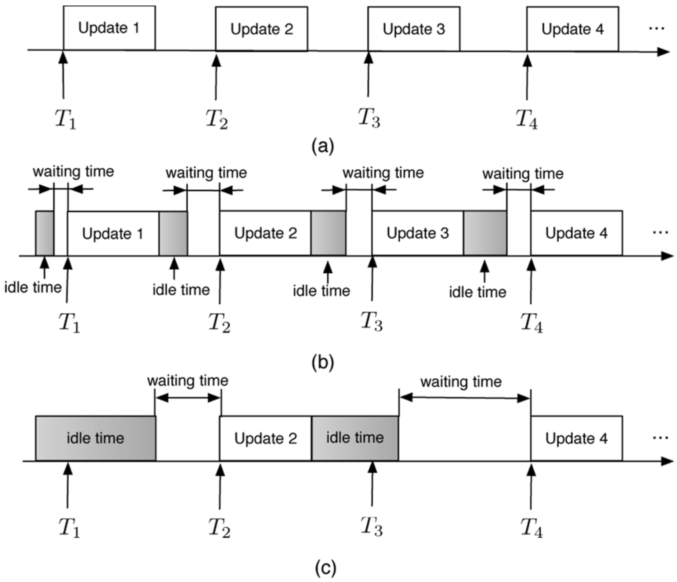

Fig. 2. Examples of possible schedules. The transmission of update i begins at time T_i.

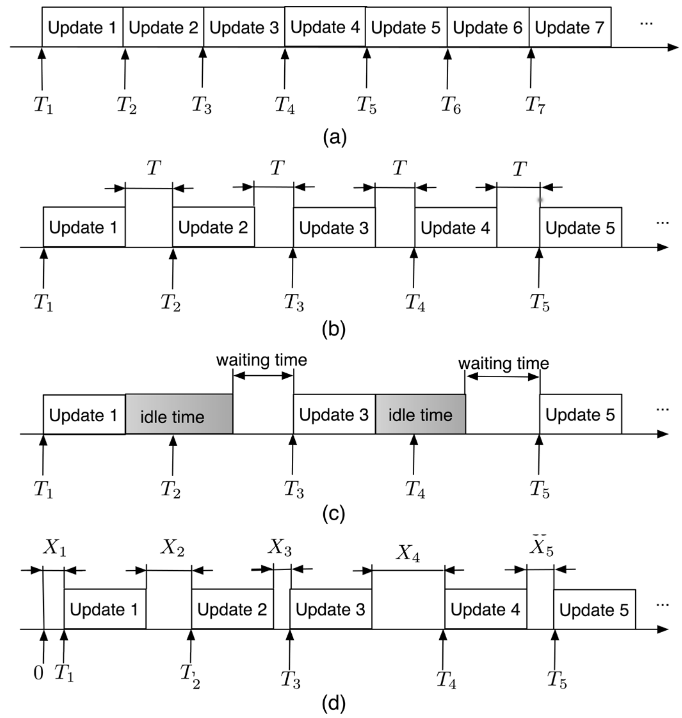

Fig. 3. Examples of possible schedules. The transmission of update i begins at time T_i.

A. Universal Schedules

We begin by observing that *deterministic* schedules do not behave well in the presence of an adversarial client. Indeed, it is easy to verify that for any given deterministic schedule, there exists a series of idle times that yields waiting times arbitrarily close to one time unit. For example, consider the deterministic schedule depicted in Fig. 3(b). This schedule transmits an update every $T+1$ time units for some $T>0$. For this schedule, the optimal strategy of an adversarial client is to select the value of the idle time equal to $T+\varepsilon$, for some small $\varepsilon>0$, resulting in a waiting time of $1+T-\varepsilon$ time units. Indeed, as shown in Fig. 3(c), such a client misses the beginning of the current update and needs to wait an entire period to receive the next update, resulting in a long waiting time.

As our goal is to design schedules in which the worst case waiting time is significantly less than one time unit, we turn to consider *random schedules*, i.e., schedules which are governed by a certain probability distribution. Fig. 3(d) depicts an example of a random schedule, which waits X time units between two subsequent transmissions, where X is a random variable. We note that in random schedules, the waiting time is also a random variable that depends on both the idle time of the client and the probability distribution of the schedule. We assume that the adversarial client knows the probability distribution of the schedule, but does not have access to the server's random bits. In this model, our goal is to design a random schedule that minimizes the worst case expected waiting time of a client, where the expectation is taken over the probability distribution governing the schedule.

We observe that in order to discriminate between different schedules we need to curtail the power of the adversarial client. Indeed, even random schedules cannot guarantee a good performance against an adversarial client that is highly adaptive to the schedule. For such clients, the optimal strategy is to wait until the beginning of the transmission of the next packet, which begins say at time t. Then, the adversarial client can select the value of the idle time in such a way that it starts listening to the channel at time $t+\varepsilon$, where $\varepsilon>0$ is a small constant. With this strategy, the worst case waiting time of the client is again close to one time unit. Thus, we need to find a meaningful way to limit the power of the adversarial client so that our goal of a significantly lower (worst case expected) waiting time can be achieved. We note that such a problem arises frequently in the design of online algorithms [21].

A natural approach to limit the power of the adversarial client is to bound its *adaptivity* to the schedule. This is done by forcing the adversarial client to fix the length of the idle time some period of time in advance. Namely, we assume that the adversarial client must fix the value of the idle time at least one time unit before the end of the idle period. With this assumption, the behavior of the client at time t can only depend on the history of the schedule at time $t-1$. In other words, the adversarial client we consider has a limited degree of adaptivity to the schedule, which can be quantified by one time unit. We also show that our approach can be extended to adversarial clients with a higher degree of adaptivity.

Fig. 4 compares universal schedules and the standard average-case approach for data broadcast. The standard model assumes that the clients' idle times are random and the goal is to construct a deterministic schedule that minimizes the expected waiting time, where the expectation is taken over the distribution of clients' idle times. With this approach, some clients may experience long waiting times. Universal schedules make no assumption on the distribution of clients' idle times. The goal of universal schedules is to minimize the worst case expected waiting time experienced by a client, where the expectation is taken over the probability distribution of the schedule. This approach guarantees low expected waiting time for any client, regardless of its behavior.

B. Related Work

Developing efficient data broadcast methods attracted a large body of research. The first methods were developed for Teletext and Videotex systems [2], [3], [24]. Cheriton [9] proposed wireless data broadcast as a means for addressing the inherent asymmetry between the uplink and downlink channels in wireless networks. The large body of research on data broadcast systems focuses on delivering multiple static items over a shared broadcast channel, assuming that data requests are distributed uniformly over time (see, e.g., [3], [4], [13]–[18], [22], and references therein). In particular, Imielinski *et al.* [19] has developed data

	Average-case Approach	Universal Schedules
Assumptions	Idle times are distributed uniformly over time.	No assumption on the distribution of clients' idle times.
Objective	Minimize the expected waiting time. The expectation is taken over the distribution of requests.	Minimize the worst case expected waiting time. The expectation is taken over the probability distribution governing the schedule.
Waiting time	The worst case waiting time is high for some clients.	The expected worst case waiting time is low for all clients.

Fig. 4. Comparison between the standard approach and universal schedules.

broadcast methods for providing fast and low-power access to data, observing that wireless data broadcasting can be viewed as storage on the air—an extension of the server's memory. Acharya *et al.* [1] proposed a new architecture, referred to as *broadcast disks*, for distributing information in wireless and mobile environments. Currently, there is a significant amount of interest in broadcasting real-time data and providing quality of service (QoS) for user applications [25]. In this correspondence, we consider a special case of a data broadcast systems that distribute a series of dynamic data updates from a single information source to a large number of mobile clients.

Several studies [5]–[8], [10], [11] have focused on minimizing the worst case waiting time for arbitrary client requests. In particular, [5]–[7] compared several streaming techniques for media-on-demand delivery under different assumptions of client request patterns and analyzed the tradeoff between bandwidth and waiting time. The basic assumption in [5]–[7] is that the server is aware of client requests and can respond to them. This differs from our work in which we assume that the server is completely oblivious of clients requests. References [10] and [11] consider server scheduling strategies to minimize the clients' response time. Specifically, they focus on the "*pull-based*" approach in which the clients send their requests to the server. In contrast, this work focuses on the "*push-based*" approach, where the server proactively transmits the data over the broadcast channel. Finally, [8] focused on minimizing the startup delay for an uninterrupted playback for media-on-demand systems that utilize multiple broadcast channels (in which the server is unaware of the requests made by clients). While [8] has a flavor which resembles our work, its model differs significantly from that considered in our study.

C. Our Results

In this correspondence, we present (for the restricted adversarial clients discussed above) a universal broadcast schedule that guarantees a worst case expected waiting time of $1/\sqrt{2} \simeq 0.7$ time units, regardless of the clients' access patterns. We show that our schedule is optimal for settings in which the transmission of every update takes one time unit. Namely, we prove that, in such settings, achieving a worst case expected waiting time less than $1/\sqrt{2}$ is not possible. Our work mainly addresses adversarial clients with limited adaptively to the schedule that can be quantified by one time unit. The techniques developed in this work can be used to address stronger adversarial clients as well (namely, those with higher adaptively).

One of the important characteristics of the schedule is the number of updates it sends over a period of time. This is important for users that have no idle time and want to continuously monitor the updates. Accordingly, we consider the design of universal schedules subject to a given rate constraint. While the optimal schedule we present has a relatively high rate of $r = \frac{2}{1+\sqrt{2}} \simeq 0.82$, in some settings a higher rate may be required. We present, for any $r \geq \frac{2}{1+\sqrt{2}}$, a schedule of rate r whose worst case expected waiting time is bounded by $\frac{2-r-\sqrt{2-2r}}{r}$ time units. We show that, under certain restrictions on the server, this is the best possible schedule. The tradeoff between the transmission rate and worst case expected waiting time is depicted in Fig. 5.

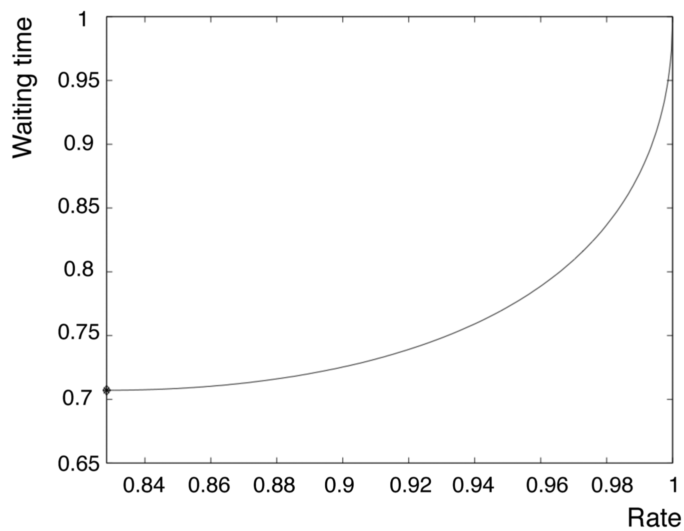

Fig. 5. The tradeoff between the transmission rate and worst case expected waiting time. The asterisk at point $\left(\frac{2}{1+\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ represents the overall optimal schedule with respect to the universal setting.

D. Organization

The remainder of this correspondence is organized as follows. In Section II, we present our model. In Section III, we focus on clients with unit adaptivity and prove our main results. In Section IV, we briefly discuss schedules for highly adaptive clients. Finally, conclusions are presented in Section V.

II. Model and Definitions

A. Random Schedules

As mentioned in the Introduction, we are interested in designing universal schedules for delivering a series of data updates from a single information source over a broadcast channel. The data is delivered in the form of equal size packets, each packet is transmitted over a time interval of length one time unit. The length of the interval is chosen without loss of generality, as our techniques (with an appropriate scaling) can be applied for time intervals of an arbitrary length.

We define a schedule $\mathcal{S}$ by specifying, for each packet i, the amount of time X_i that passes between the end of the transmission of packet $i-1$ and the beginning of the transmission of packet i (for clarity, we assume that the transmission of packet 0 ends at time 0). We refer to X_i as the *interleaving time* for packet i (see Fig. 3(d)). In a random schedule, the interleaving times are random variables.

Definition 1 (Random Schedule $\mathcal{S}$): A *random schedule* is a sequence of random variables $\{X_1, X_2, \ldots\}$ such that X_i is the interleaving time for packet i.

We denote by $[T_i, T_i + 1]$, $i \geq 1$, the transmission interval of packet i. We refer to

$$T_i = i - 1 + \sum_{j=1}^{i} X_j \tag{1}$$

as the starting time of packet i and to $\{T_1, T_2, \ldots\}$ as the *transmission sequence* of schedule $\mathcal{S}$. Note that the starting times T_i, $i \geq 1$ are also random variables. Fig. 3(d) depicts an example of a random schedule $\{X_1, X_2, \ldots\}$ and the corresponding transmission sequence $\{T_1, T_2, \ldots\}$.

Let $\mathcal{S}$ be a schedule, and suppose that the client's idle period ends at time t, that is, the client starts to listen to the channel at time t. We say that in the end of the idle period, a client places a *request* to the channel. We define the client's waiting time as the length of the time interval between t and the beginning of the transmission of the next packet.

Definition 2 (Waiting Time, $WT(\mathcal{S}, t)$): The *Waiting Time* for a request at time t with a schedule $\mathcal{S}$ is defined to be $WT(\mathcal{S}, t) = T_i - t$, where i is the first packet for which it holds that $T_i \geq t$.

Note that $WT(\mathcal{S}, t)$ is a random variable. We denote the expectation of $WT(\mathcal{S}, t)$ by $EWT(\mathcal{S}, t) = E[WT(\mathcal{S}, t)]$. This is the expected waiting time experienced by *oblivious* clients, i.e., clients whose idle times are fixed in advance and independent of the history of the schedule.

A *history* $V_t(x_1, \ldots, x_\ell)$ of schedule $\mathcal{S}$ observed at time t is the event in which:

1) for $n \leq \ell$, the random variables X_n are equal to x_n, i.e., for all $n \leq \ell$ it holds $X_n = x_n$;
2) the number of (partial) packets broadcasted until time t is at least ℓ, i.e., $\ell - 1 + \sum_{i=1}^{\ell} X_i \leq t$;
3) transmission of $\ell + 1$th packet has not been completed at time t, i.e., $\ell + \sum_{i=1}^{\ell+1} X_i > t$.

We denote by $\mathcal{V}(\mathcal{S}, t)$ the set of possible histories of $\mathcal{S}$ at time t. Finally, for any $V \in \mathcal{V}(\mathcal{S}, t)$ let $\mathcal{S}|V$ be the schedule distribution obtained by conditioning $\mathcal{S}$ on the event V.

B. Adversarial Clients

Our goal is to design schedules that perform well for any client, regardless of the viewed history of the schedule. An adversarial client has complete control on the values of idle times, and, as a result, may place a request at any time t. The latter, for example, can be obtained by setting the initial idle time (before the first update) to be equal to t. In the remainder of this work, we focus on bounding the worst case expected waiting time experienced by an adversarial client placing a request at time t. This allows to simplify the analysis, without any loss of generality.

In general, the adversarial clients may have different degrees of adaptivity.

Definition 3 (Degree of Adaptivity, ω): We say that an adversarial client is ω-adaptive if its actions at time t depend only on the history $V \in \mathcal{V}(\mathcal{S}, t - \omega)$ of the schedule $\mathcal{S}$ at time $t - \omega$.

The expected waiting time $EWT_V(\mathcal{S}, t)$ of an ω-adaptive adversarial client with viewed history $V \in \mathcal{V}(\mathcal{S}, t - \omega)$ that places a request at time t is defined as

$$EWT_V(\mathcal{S}, t) = E[WT(\mathcal{S}|V, t)]. \tag{2}$$

Note that in this definition, the random schedule $\mathcal{S}$ is conditioned on the event V. This captures the client's knowledge of the schedule $\mathcal{S}$ at time $t - \omega$.

The worst case expected waiting time of the schedule $\mathcal{S}$ for ω-adaptive adversarial clients, $W(\mathcal{S}, \omega)$, is defined as

$$W(\mathcal{S}, \omega) = \max_{t \geq 0} \left(\sup_{0 \leq t < \omega} EWT(\mathcal{S}, t), \sup_{t \geq \omega} \sup_{V \in \mathcal{V}(\mathcal{S}, t-\omega)} EWT_V(\mathcal{S}, t) \right).$$

Namely, $W(\mathcal{S}, \omega)$ is the maximum expected waiting time of the schedule over all request times t and over all possible histories of the schedule up to the time $t - \omega$. Note that we need to take the maximum over any $t \geq 0$. Indeed, for any $t \geq 0$, there exists a client that can generate a request at this time. Note also that the expression $\sup_{0 \leq t < \omega} EWT(\mathcal{S}, t)$ bounds the maximum waiting time for requests placed at times $t < \omega$, when the client does not have any knowledge of the schedule's history.

We note that a random schedule with independent and identically distributed (i.i.d.) random variables X_i is closely related to *renewal* processes (e.g., [12]). In this context, the random variable $WT(\mathcal{S}, t)$ is referred to as the forward recurrence time or *excess* time. Moreover, the mean and limiting distribution of $WT(\mathcal{S}, t)$ are well studied. In our model, this corresponds to the expected waiting time for oblivious clients. The goal of this work, however, is to minimize the expected waiting time for adaptive clients. To the best of our knowledge, this problem has not been addressed in the literature.

For a real random variable X, we denote by $F_X(t) = \Pr[X < t]$ its cumulative distribution function, by $\mu_X = \int_0^\infty (1 - F_X(x))dx$ its expectation, and by f_X its probability density function (when exists). When clear from the context, we omit X from the above notation.

C. Transmission Rate

The *transmission rate* of a schedule $\mathcal{S} = \{X_1, X_2, \ldots\}$ is defined to be the expected fraction of the time the channel is in use.

Definition 4 (Transmission Rate r): Let $\mathcal{S} = \{X_1, X_2, \ldots\}$ be a random schedule and let R_t be the expected number of packets sent in $\mathcal{S}$ up to time t. The transmission rate of $\mathcal{S}$ is defined to be

$$r = \lim_{t \to \infty} \frac{R_t}{t} \tag{3}$$

if such a limit exists.

III. Universal Schedules for $\omega = 1$

In this section, we study the design of optimal universal schedules for adversarial clients with a degree of adaptivity of one time unit (i.e., $\omega = 1$). Namely, we present a family of schedules, one for each rate value r, that guarantee a minimum worst case expected waiting time subject to the rate constraint r. For $r = \frac{2}{1+\sqrt{2}}$, the schedule we present has an (overall) optimal worst case expected waiting time of value $1/\sqrt{2}$, i.e., any other schedule $\mathcal{S}$ has a corresponding waiting time $W(\mathcal{S}, 1)$ of value at least $1/\sqrt{2}$.

Our schedules are defined by a single random variable X. That is, all interleaving times X_i in our schedules are independent and have the same distribution as X.

A. Optimal Schedules

Each random schedule in the family we present is associated with a parameter μ, which is equal to the expected value of X, i.e., $\mu = E[X]$. The random variable X has the following simple structure. Let Z be a "random" variable equal to 0 with probability 1 and let $U[0, s]$ be the random variable with uniform distribution in the interval $[0, s]$. Finally, let $p = 1 - \sqrt{\frac{2\mu}{\mu+1}}$, and $s = \sqrt{2\mu(\mu + 1)}$. Then, X is defined as follows:

$$X = \begin{cases} Z, & \text{with probability } p \\ U[0, s], & \text{with probability } 1 - p. \end{cases} \tag{4}$$

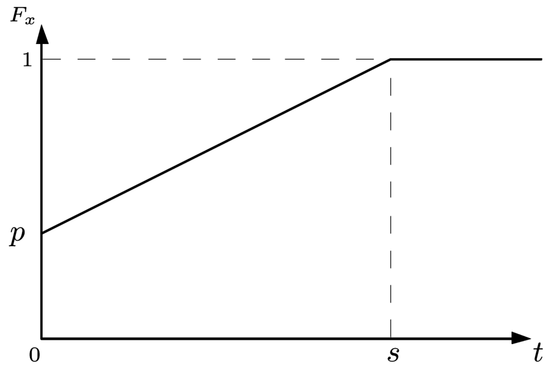

Fig. 6. The cumulative distribution function F_X of X for a schedule with parameter μ. Here $s = \sqrt{2\mu(\mu+1)}$ and $p = 1 - \sqrt{\frac{2\mu}{\mu+1}}$.

The cumulative distribution function $F_X(t)$ of X is

$$F_X(t) = \begin{cases} 0, & \text{if } t = 0 \\ p + \frac{1-p}{s} \cdot t, & \text{if } t \in (0, s] \\ 1, & \text{if } t > s \end{cases} \tag{5}$$

or equivalently

$$F_X(t) = \begin{cases} 0, & \text{if } t = 0 \\ 1 - \sqrt{\frac{2\mu}{\mu+1}} + \frac{t}{\mu+1}, & \text{if } t \in (0, s] \\ 1, & \text{if } t > s. \end{cases} \tag{6}$$

The function $F_X(t)$ is depicted in Fig. 6.

It is not hard to verify that $E[X] = \mu$ and that the rate of the schedule defined by X is $\frac{1}{1+\mu}$ (e.g., [12]). In Theorem 2 below we analyze (for a range of values μ) the worst case expected waiting time of the proposed schedule. Moreover, we show that for the schedule $\mathcal{S}$ that corresponds to $\mu = \frac{\sqrt{2}-1}{2}$ it holds that $W(\mathcal{S}, 1) = 1/\sqrt{2}$.

We begin by proving the following lemma.

Lemma 1: Let $\mathcal{S} = \{X_1, X_2, \ldots\}$ be a random schedule in which X_1 and X_2 are independent. Then, for $t \leq 1$ it holds that

$$EWT(\mathcal{S}, t) = \mu_{X_1} + (\mu_{X_2} + 1)F_{X_1}(t) - t. \tag{7}$$

Proof: Let A be the event that $X_1 < t$ and let $\bar{A}$ be the complement of A. Then, $\Pr[A] = F_{X_1}(t)$. We denote by $X_1|A$ the random variable X_1 conditioned on event A and by $X_1|\bar{A}$ the random variable X_1 conditioned on event $\bar{A}$. We consider the following two cases.

1) If event A happens, then the client misses the transmission of the first packet. In this case, the client must wait $1 - t + X_1|A$ time units until the transmission of the first packet completes and, in addition, X_2 time units until the beginning of the second packet.
2) If event $\bar{A}$ happens, the client needs to wait $X_1|\bar{A} - t$ time units until the transmission of the first packet begins.

Thus, $EWT(\mathcal{S}, t)$ is equal to

$$\begin{aligned} &= \Pr[A]E[1 - t + X_2 + X_1|A] + \Pr[\bar{A}]E[X_1|\bar{A} - t] \\ &= F_{X_1}(t)(1 - t + \mu_{X_2}) + \Pr[A]E[X_1|A] \\ &\quad + \Pr[\bar{A}]E[X_1|\bar{A}] - (1 - F_{X_1}(t))t \\ &= \mu_{X_1} + (\mu_{X_2} + 1)F_{X_1}(t) - t. \end{aligned}$$

□

Theorem 2: For any $\mu \in \left[0, \frac{\sqrt{2}-1}{2}\right]$ the worst case expected waiting time of the corresponding schedule $\mathcal{S}$ is $W(\mathcal{S}, 1) = 1 + 2\mu - \sqrt{2\mu(\mu+1)}$. In particular, the worst case expected waiting time $W(\mathcal{S}, 1)$ of the schedule $\mathcal{S}$ that corresponds to $\mu = \frac{\sqrt{2}-1}{2}$ is $1/\sqrt{2}$.

Proof: To bound the value of $W(\mathcal{S}, 1)$ we must bound both the expressions

$$\sup_{t \in [0,1)} EWT(\mathcal{S}, t) \quad \text{and} \quad \sup_{t \geq 1} \sup_{V \in \mathcal{V}(\mathcal{S}, t-1)} EWT_V(\mathcal{S}, t).$$

We start by computing $EWT(\mathcal{S}, t)$ for $t \in [0, 1)$. The value of $EWT(\mathcal{S}, 0)$ (i.e., the expected waiting time for the request at time 0) is exactly μ. Lemma 1 implies that for $t \in (0, 1)$, $EWT(\mathcal{S}, t) = \mu + (\mu + 1)F_X(t) - t$. Substituting $F_X(t)$ from (6) yields

$$EWT(\mathcal{S}, t) = \begin{cases} \mu, & \text{if } t = 0 \\ 1 + 2\mu - \sqrt{2\mu(\mu+1)}, & \text{if } t \in (0, s] \\ 1 + 2\mu - t, & \text{if } t > s. \end{cases}$$

This implies, in turn, that

$$\sup_{t \in [0,1)} EWT(\mathcal{S}, t) = 1 + 2\mu - \sqrt{2\mu(\mu+1)}.$$

Now consider any $t \geq 1$, and any history $V_{t-1}(x_1, \ldots, x_\ell) \in \mathcal{V}(\mathcal{S}, t-1)$. We would like to analyze the waiting time $EWT_{V_{t-1}}(\mathcal{S}, t)$. Let $t_\ell = \ell - 1 + \sum_{i=1}^{\ell} x_i$ be the beginning of the transmission interval for the ℓth packet. By the fact that $V_{t-1} \in \mathcal{V}(\mathcal{S}, t-1)$ we have that $t_\ell \leq t - 1$. Furthermore, as X is bounded by s, we conclude that $t_\ell > t - 2 - s$. We denote by t^* the time that passed between the end of the transmission interval for packet ℓ and the request time t, i.e., $t^* = t - t_\ell - 1$ (see Fig. 7). We consider two cases.

a) Case 1, $t^* \in [0, 1)$: In this case, $t - 1$ belongs to the transmission interval of packet ℓ, i.e., $t_\ell \leq t - 1 < t_\ell + 1$. This case is demonstrated in Fig. 7(a). We show that $EWT_{V_{t-1}}(\mathcal{S}, t) = EWT(\mathcal{S}, t^*)$. Note that in this case, the client has no knowledge of the value of the interleaving time $X_{\ell+1}$. Thus, $\mathcal{S}$ conditioned on the event $V_{t-1}(x_1, \ldots, x_\ell)$ (i.e., $\mathcal{S}|V_{t-1}$) is a random schedule $\{x_1, x_2, \ldots, x_\ell, X_{\ell+1}, X_{\ell+2}, \ldots\}$, where $x_1, x_2, \ldots, x_\ell$ are fixed values and $X_{\ell+1}, X_{\ell+2}, \ldots$ are random variables identical to X. Hence, the schedule $\mathcal{S}|V_{t-1}$ can be viewed as a sequence of ℓ packets with interleaving times $\{x_1, \ldots, x_\ell\}$ followed by the schedule $\mathcal{S}$ (which now starts at time $t_\ell + 1$). We conclude that for $z > 0$ and any $t \geq t_\ell + 1$ the probability that a request placed at time t in $\mathcal{S}|V_{t-1}$ will have waiting time z is equal to the probability that a request placed at time t^* in $\mathcal{S}$ has waiting time z. This implies that in this case

$$EWT_{V_{t-1}}(\mathcal{S}, t) \leq 1 + 2\mu - \sqrt{2\mu(\mu+1)}.$$

b) Case 2, $t^* \in [1, 1+s)$: In this case, at time $t-1$, the transmission of packet ℓ has already been completed. This case is demonstrated in Fig. 7(b). Similar to the previous case, $\mathcal{S}|V_{t-1}$ is a random schedule defined by the interleaving times $\{x_1, x_2, \ldots, x_\ell, X_{\ell+1}, X_{\ell+2}, \ldots\}$. We note that in this case the client knows that $X_{\ell+1} > t^* - 1$, hence, X_ℓ is equal to random variable X conditioned on this event, i.e., $X_{\ell+1} = X|\{X > t^* - 1\}$. All other random variables $X_{\ell+2}, X_{\ell+3}, \ldots$ are identical to X.

Using an argument similar to that used in Case 1, we now claim that

$$EWT_{V_{t-1}}(\mathcal{S}, t) = EWT(\mathcal{S}', 1)$$

where $\mathcal{S}' = \{X'_{\ell+1}, X_{\ell+2}, X_{\ell+3}, \ldots\}$, $X'_{\ell+1} = X|\{X > t^* - 1\} - (t^* - 1)$, and $X_{\ell+2}, X_{\ell+3}, \ldots$ are identical to X. The cumulative distribution function $F_{X'_{\ell+1}}(t)$ of $X'_{\ell+1}$ is

$$F_{X'_{\ell+1}}(t) = \begin{cases} 0, & \text{if } t \leq 0 \\ \frac{t}{s - t^* + 1}, & \text{if } 0 < t \leq s - (t^* - 1) \\ 1, & \text{if } t > s - (t^* - 1). \end{cases} \tag{8}$$

The expected value of $X'_{\ell+1}$ is $(s - t^* + 1)/2$.

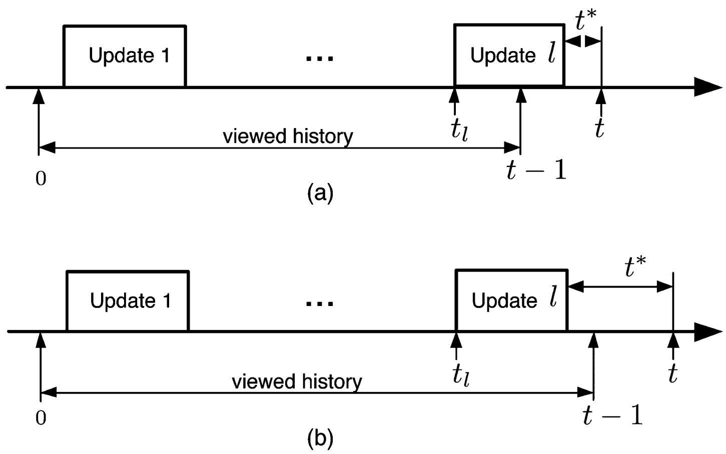

Fig. 7. Two possible cases for the proof of Theorem 2. (a) Case 1. (b) Case 2.

By Lemma 1 we have

$$EWT(\mathcal{S}',1) = \mu_{X'_{\ell+1}} + (\mu+1)F_{X'_{\ell+1}}(1) - 1 = \mu + \frac{s-t^*+1}{2} \le \mu + \frac{s}{2} \tag{9}$$

where we use the fact that $t^* \ge 1$ and that for $\mu \in \left[0, \frac{\sqrt{2}-1}{2}\right]$ it holds that $s < 1$. Substituting s yields

$$EWT(\mathcal{S}',1) \le \mu + \sqrt{\mu(\mu+1)/2} \le 1 + 2\mu - \sqrt{2\mu(\mu+1)}.$$

The last inequality follows from the fact that $\mu \le 2/7$ for our choice of μ. □

B. Proof of Optimality

We proceed to prove that the worst case expected waiting time of any schedule $\mathcal{S} = \{X_1, X_2, \ldots\}$ is greater or equal to $1/\sqrt{2}$. This is done in two steps. We start by proving the assertion for a schedule $\mathcal{S}$ defined by i.i.d. random variables X. Then, we address the general case, in which each random variable X_n in $\mathcal{S}$ may be arbitrarily distributed and may depend on X_i for $i < n$.

Theorem 3: Let X be a random variable and let $\mathcal{S} = \{X_1, X_2, \ldots\}$ be a random schedule where $X_i = X$ for all i. Then, there exists $t \in [0,1)$ such that $EWT(\mathcal{S},t) \ge 1/\sqrt{2}$, which, in turn, implies that $W(\mathcal{S},1) \ge 1/\sqrt{2}$. Moreover, for all $\mu = E[X] \in \left[0, \frac{\sqrt{2}-1}{2}\right]$ it holds that $W(\mathcal{S},1) \ge 1 + 2\mu - \sqrt{2\mu(\mu+1)}$.

Proof: We denote by F the cumulative distribution function of X and by μ the expected value of X. We begin by considering the case in which $\mu \in \left[0, \frac{\sqrt{3}-1}{2}\right]$. Let $\mathcal{S}' = \{X'_1, X'_2, \ldots\}$ be the schedule in which each X'_i is an independent random variable defined by X' as in (4), i.e.,

$$X' = \begin{cases} Z, & \text{with probability } p \\ U[0,s], & \text{with probability } 1-p. \end{cases}$$

Where Z is a random variable equal to 0 with probability 1, $U[0,s]$ is the random variable with uniform distribution in the interval $[0,s]$, $p = 1 - \sqrt{\frac{2\mu}{\mu+1}}$, and $s = \sqrt{2\mu(\mu+1)}$.

We show that

$$\sup_{t\in[0,1)} EWT(\mathcal{S},t) \ge \sup_{t\in[0,1)} EWT(\mathcal{S}',t). \tag{10}$$

As shown in Theorem 2

$$\sup_{t\in[0,1)} EWT(\mathcal{S}',t) = 1 + 2\mu - \sqrt{2\mu(\mu+1)}.$$

Since $1 + 2\mu - \sqrt{2\mu(\mu+1)} \ge 1/\sqrt{2}$, this is sufficient to prove the assertion.

By Lemma 1, for $t \in [0,1]$ it holds that $EWT(\mathcal{S},t) = \mu + (\mu+1)F(t) - t$. Assume, by the way of contradiction, that

$$\sup_{t\in[0,1)} EWT(\mathcal{S},t) < \sup_{t\in[0,1)} EWT(\mathcal{S}',t).$$

This implies that for each $t \in [0,1)$ it holds that

$$EWT(\mathcal{S},t) = \mu + (\mu+1)F(t) - t < 1 + 2\mu - \sqrt{2\mu(\mu+1)}.$$

Thus, $1 - F(t) > \sqrt{\frac{2\mu}{\mu+1}} - \frac{t}{\mu+1}$. However, this implies that the expectation of X is greater than μ

$$\int_0^s (1-F(x))\mathrm{d}x > \int_0^s \left(\sqrt{\frac{2\mu}{\mu+1}} - \frac{x}{\mu+1}\right)\mathrm{d}x = \mu$$

resulting in a contraction. Note that the last equation follows form the fact that $s \le 1$.

Next, we consider the case in which $\mu \in \left(\frac{\sqrt{3}-1}{2}, 1/\sqrt{2}\right)$. For this case, we use a similar argument. In particular, we define a random schedule $\mathcal{S}' = \{X'_1, X'_2, \ldots\}$ in which each X'_i is an independent random variable defined by

$$X' = \begin{cases} Z, & \text{with probability } p_1 \\ U[0,1], & \text{with probability } 1-p_1-p_2 \\ Z', & \text{with probability } p_2. \end{cases}$$

Here, Z is a random variable equal to 0 with probability 1, Z' is a random variable which has the value 1 with probability 1, $U[0,1]$ is the random variable with uniform distribution in the interval $[0,1]$, $p_1 = \frac{1-2\mu^2}{2(1+\mu)}$, and $p_2 = \frac{2\mu^2+2\mu-1}{2(1+\mu)}$.

It is not hard to verify that $E[X'] = \mu$, $\sup_{t\in[0,1)} EWT(S',t) = \frac{1}{2} + \mu - \mu^2$, and that $\frac{1}{2} + \mu - \mu^2 \ge 1/\sqrt{2}$ for any $\mu \in \left(\frac{\sqrt{3}-1}{2}, 1/\sqrt{2}\right)$.

Assume, by the way of contradiction, that

$$\sup_{t\in[0,1)} EWT(\mathcal{S},t) < \sup_{t\in[0,1)} EWT(\mathcal{S}',t).$$

This implies that for $t \in [0,1)$ it holds that

$$EWT(\mathcal{S},t) = \mu + (\mu+1)F(t) - t < \frac{1}{2} + \mu - \mu^2.$$

Thus

$$1 - F(t) > \frac{1/2 + \mu + \mu^2}{1 + \mu} - \frac{t}{1 + \mu}.$$

This implies that the expectation $E[X]$ of X is greater than μ

$$\int_0^1 (1 - F(x))\mathrm{d}x > \int_0^1 \left(\frac{1/2 + \mu + \mu^2}{1 + \mu} - \frac{t}{1 + \mu}\right)\mathrm{d}x = \mu$$

resulting in a contradiction.

Finally, for $\mu \geq 1/\sqrt{2}$, we notice that $EWT(\mathcal{S}, 0) = \mu \geq 1/\sqrt{2}$, which suffices to prove our assertion. □

We now use Theorem 3 to prove the lower bound on $W(\mathcal{S}, 1)$ for general schedules $\mathcal{S}$.

Corollary 4: Let $\mathcal{S} = \{X_1, X_2, \ldots\}$ be a schedule in which each random variable X_n may be arbitrarily distributed and may depend on X_i for $i < n$. Then, $W(\mathcal{S}, 1) \geq 1/\sqrt{2}$.

Proof: Let $\varepsilon > 0$ be an arbitrarily small constant. We construct a random variable X' such that the schedule $\mathcal{S}' = \{X'_1, X'_2, \ldots\}$ in which each X'_i, $i \geq 1$ is independent and identical to X' satisfies $W(\mathcal{S}, 1) \geq EWT(\mathcal{S}', t) - \varepsilon$ for $t \in [0, 1)$. This suffices to prove our assertion. Indeed, by Theorem 3, there exists $t \in [0, 1)$ such that $EWT(\mathcal{S}', t) \geq 1/\sqrt{2}$, which implies that $W(\mathcal{S}, 1) \geq 1/\sqrt{2} - \varepsilon$.

Let

$$\mu = \inf_{\ell \in N} \inf_{\{x_1, \ldots, x_\ell\} \in (R^+)^\ell} E[X_{\ell+1} | X_1 = x_1, X_2 = x_2, \ldots, X_\ell = x_\ell].$$

Here N is the set of natural numbers, and R^+ is a set of nonnegative reals. Let $\ell \in N$, $\{x_1, \ldots, x_\ell\} \in (R^+)^\ell$ that satisfy

$$\mu_{\ell+1} = E[X_{\ell+1} | X_1 = x_1, X_2 = x_2, \ldots, X_\ell = x_\ell] \leq \mu + \varepsilon. \quad (11)$$

We denote $t_0 = \sum_{i=1}^{\ell}(x_i + 1)$ and consider the history $V_{t_0 - 1}(x_1, \ldots, x_\ell)$. Let X' be the random variable equal to $X_{\ell+1} | V_{t_0-1}(x_1, \ldots, x_\ell)$. We now claim that for any $t \in [0, 1)$, it holds that

$$EWT(\mathcal{S} | V_{t_0-1}(x_1, \ldots, x_\ell), t_0 + t) \geq EWT(\mathcal{S}', t) - \varepsilon.$$

Let $\mu_{\ell+2}(x_{\ell+1})$ be the expectation of $X_{\ell+2}$ conditioned on the event $X_1 = x_1, X_2 = x_2, \ldots, X_\ell = x_\ell$ and $X_{\ell+1} = x_{\ell+1}$. Notice that given the history $V_{t_0-1}(x_1, \ldots, x_\ell)$, $\mu_{\ell+2}$ is a random variable that may depend on $X_{\ell+1}$. Let $t \in [0, 1)$. Let A be the event $V_{t_0-1}(x_1, \ldots, x_\ell) \cap (X_{\ell+1} < t)$. Let F be the distribution function of $X_{\ell+1} | V_{t_0-1}(x_1, \ldots, x_\ell)$. We note that $\Pr[A] = F(t)$. Let $\bar{A}$ be the event $V_{t_0-1}(x_1, \ldots, x_\ell) \cap (X_{\ell+1} \geq t)$. Similarly to the proof of Lemma 1, we have that $EWT(\mathcal{S} | V_{t_0-1}(x_1, \ldots, x_\ell), t_0 + t)$ is equal to

$$\begin{aligned}
&\Pr[A]E[1 + X_{\ell+1} | A - t + X_{\ell+2} | A] + Pr[\bar{A}]E[X_{\ell+1} | \bar{A} - t] \\
&\geq \Pr[A]E[1 + X_{\ell+1} | A - t + \mu] + \Pr[\bar{A}]E[X_{\ell+1} | \bar{A} - t] \\
&\geq \mu_{\ell+1} + (\mu_{\ell+1} + 1)F(t) - t - \varepsilon = EWT(\mathcal{S}', t) - \varepsilon.
\end{aligned}$$

In the preceding calculation, all expectations are taken over the schedule $\mathcal{S} | V_{t_0-1}(x_1, \ldots, x_\ell)$. The first inequality follows from the fact that $E[X_{\ell+2} | A] \geq \mu$, while the second inequality is due to the facts that $\mu \geq \mu_{\ell+1} - \varepsilon$ and $F(t) \leq 1$ for any $t \in [0, 1)$. □

C. *Optimal Schedules for Large Rates*

We now summarize our results for schedules of rate $r \in \left[\frac{2}{1+\sqrt{2}}, 1\right]$.

Corollary 5: For any $r \in \left[\frac{2}{1+\sqrt{2}}, 1\right]$, the family presented in Section III-A contains a schedule $\mathcal{S} = \{X_1, X_2, \ldots\}$ with rate r and worst case expected waiting time $W(\mathcal{S}, 1) = \frac{2 - r - \sqrt{2 - 2r}}{r}$. Moreover, any schedule $\mathcal{S}' = \{X'_1, X'_2, \ldots\}$ in which X'_n are i.i.d with rate r satisfies

$$W(\mathcal{S}', 1) \geq \frac{2 - r - \sqrt{2 - 2r}}{r}.$$

Proof: Follows from Theorems 2 and 3 (recall that in the schedules we define in Section III-A the rate r is $\frac{1}{1+\mu}$). □

The tradeoff between the transmission rate of our schedules and the worst case waiting time is depicted in Fig. 5.

Two remarks are in place . First, note that in order to reduce the worst case expected waiting time of a client, we need to introduce a certain amount of randomness to the schedule, which results in rate decrease. The optimal waiting time of $1/\sqrt{2}$ is achieved at rate $r = \frac{2}{1+\sqrt{2}}$. For rate constraints $r \geq \frac{2}{1+\sqrt{2}}$, the achievable worst case expected waiting time is an increasing function in r. This is explained by the fact that the high rate requirement does not provide sufficient slackness needed to minimize the waiting time. In contrast, reducing the rate constraint r below $\frac{2}{1+\sqrt{2}}$ does not lead to a decrease in the waiting time. This corresponds to the situation where excessive slackness results in a high waiting time. Indeed, for the extreme case of $r = 0$ (i.e., no packets are transmitted), the waiting time is unbounded. Second, notice that the tradeoff curve we present has a *knee* phenomenon. That is, increasing the rate beyond $\frac{2}{1+\sqrt{2}} \simeq 0.82$ has little effect on the worst case expected waiting time until r reached a value of approximately 0.95. From this point on, any increase in rate significantly affects the waiting time. The point $(0.94, 0.76)$ is a reasonable tradeoff between the rate and the worst case expected waiting time.

IV. Highly Adaptive Adversarial Clients

In this section, we discuss schedules for adversarial clients with very small degree of adaptivity ω. For this extreme case we give a tight analysis of $W(\mathcal{S}, \omega)$.

Theorem 6: Let n be a sufficiently large integer. There exists a constant c independent of n for which it holds that

$$1 - \frac{1}{n} \leq W(\mathcal{S}, \frac{1}{n}) \leq 1 - \frac{c}{n}. \quad (12)$$

Proof: The lower bound follows easily by the fact that a client request can be placed $\frac{1}{n}$ time units after the transmission of a packet, resulting in a waiting time of at least $1 - \frac{1}{n}$. For the upper bound, consider the schedule $\mathcal{S} = \{X_1, X_2, \ldots\}$ in which X_i are i.i.d. according to X which is 0 with probability $\frac{3}{4}$ and $\frac{1}{2n}$ with probability $\frac{1}{4}$. For a client's request at time $t \leq \frac{1}{n}$, Lemma 1 implies that the expected waiting time is bounded by $1 - \frac{1}{4n}$ (for n large enough). For $t > \frac{1}{n}$ and any viewed history $V_{t-\frac{1}{n}}$, it is not hard to verify (using techniques similar to those used in Theorem 2) that the expected waiting time is also bounded by $1 - \frac{1}{4n}$. This suffices to prove the theorem. □

V. Conclusion

In this correspondence, we defined the notion of *universal schedules* that guarantee low waiting time for any client, regardless of its access pattern. We studied the performance characteristics and the design of universal broadcast schedules, focusing on adversarial clients whose adaptivity is bounded by one time unit. For such clients, we presented a schedule that guarantees a worst case expected waiting time of at most $1/\sqrt{2}$, for any request and for any history of the broadcast channel. Moreover, we have shown that this is the best possible schedule. Our schedule has a transmission rate of $r \simeq 0.82$. For larger values of r we have presented a tight analysis of the tradeoff between the transmission rate and the minimum worst case expected waiting time. Our analysis applies to all random schedules with i.i.d. interleaving times.

It is natural to consider extending our proof techniques to clients with different degrees of adaptivity ω. The major difficulty in extending our techniques to a wider range of values for ω lies in the expressions that arise when computing the value of the expected waiting time. For smaller values of ω, these expressions do not differ significantly from those appearing in this work and an analysis of similar nature may be performed. For larger values of ω, one needs to take into consideration that several packets may be transmitted in the time interval $[t-\omega, t]$. This work is a first step toward the design of worst case efficient schedules. As a future research, it would be interesting to establish tight bounds on the worst case expected waiting time for every value of ω.

References

[1] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast disks: Data management for asymmetric communication environments," in *Proc. ACM SIGMOD*, San Jose, CA, May 1995, pp. 199–210.

[2] M. H. Ammar and J. W. Wong, "The design of Teletext broadcast cycles," *Perform. Eval.*, vol. 5, no. 4, pp. 235–242, 1985.

[3] M. H. Ammar, "Response time in a Teletext system: An individual user's perspective," *IEEE Trans. Commun.*, vol. COM-35, no. 11, pp. 1159–1170, Nov. 1987.

[4] A. Bar-Noy, R. Bhatia, J. Naor, and B. Schieber, "Minimizing service and operation costs of periodic scheduling," *Math. Operations Res.*, vol. 27, no. 3, pp. 518–544, 2002.

[5] A. Bar-Noy, J. Goshi, and R. E. Ladner, "Off-line and on-line guaranteed start-up delay for media-on-demand with stream merging," *J. Discr. Algorithms*, vol. 4, no. 1, pp. 72–105, 2006.

[6] A. Bar-Noy, J. Goshi, R. E. Ladner, and K. Tam, "Comparison of stream merging algorithms for media-on-demand," *Multimedia Syst.*, vol. 9, no. 5, pp. 411–423, 2004.

[7] A. Bar-Noy and R. E. Ladner, "Efficient algorithms for optimal stream merging for media-on-demand," *SIAM J. Comput.*, vol. 33, no. 5, pp. 1011–1034, 2004.

[8] A. Bar-Noy, R. E. Ladner, and T. Tamir, "Optimal delay for media-on-demand with pre-loading and pre-buffering," *Theor. Comp. Sci.*, vol. 399, no. 1-2, pp. 3–31, 2008.

[9] D. Cheriton, Dissemination-Oriented Communication Systems Stanford Univ., Stanford, CA, 1992.

[10] J. Edmonds and K. Pruhs, "Multicast pull scheduling: When fairness is fine," *Algorithmica*, vol. 36, no. 3, pp. 315–330, 2003.

[11] J. Edmonds and K. Pruhs, "A maiden analysis of longest wait first," *ACM Trans. Algorithms*, vol. 1, no. 1, pp. 14–32, 2005.

[12] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York: Wiley, 1966.

[13] K. Foltz and J. Bruck, "Time division is better than frequency division for periodic internet broadcast of dynamic data," in *Proc. IEEE Int. Symp. Information Theory*, Washington, DC, Jun./Jul. 2001, p. 158.

[14] K. Foltz and J. Bruck, "Robustness of time-division schedules for internet broadcast," in *Proc. IEEE Int. Symp. Information Theory)*, Lausanns, Switzerland, Jun./Jul. 2002, p. 383.

[15] K. Foltz and J. Bruck, "Splitting schedules for internet broadcast communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 345–358, Feb. 2002.

[16] K. Foltz, L. Xu, and J. Bruck, "Coding and scheduling for efficient loss-resilient data broadcasting," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun./Jul. 2003, p. 413.

[17] K. Foltz, L. Xu, and J. Bruck, "Scheduling for efficient data broadcast over two channels," in *Proc. IEEE Int. Symp. Information Theory* , Chicago, IL, Jun./Jul. 2004, p. 113.

[18] T. Imielinski, S. Viswanathan, and B. R. Badrinath, "Data on air: Organization and access," *IEEE Trans. Knowl. Data Eng.*, vol. 9, no. 3, pp. 353–372, May 1997.

[19] T. Imielinski, S. Viswanathan, and B. R. Badrinath, "Data on air: Organization and access," *IEEE Trans. Knowledge Data Eng.*, vol. 9, no. 3, pp. 353–372, May 1997.

[20] S. H. Kang, S. Choi, S. J. Choi, G. Lee, J. Lew, and J. Lee, "Scheduling data broadcast based on multi-frequency in mobile interactive broadcasting," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 405–411, Mar. 2007.

[21] R. M. Karp, E. Koutsoupias, C. H. Papadimitriou, and S. Shenker, "Optimization problems in congestion control," in *Proc. IEEE Symp. Foundations of Computer Science*, Redondo Beach, CA, Nov. 2000, pp. 66–74.

[22] C. Kenyon, N. Schabanel, and N. E. Young, "Polynomial-time approximation scheme for data broadcast," in *Proc. Symp. Theory of Computing (STOC)*, Portland, OR, May 2000, pp. 659–666.

[23] N. H. Vaidya and S. Hameed, "Scheduling data broadcast in asymmetric communication environments," *Wireless Netw.*, vol. 5, no. 3, pp. 171–182, 1999.

[24] J. W. Wong and M. H. Ammar, "Analysis of broadcast delivery in a Videotex system," *IEEE Trans. Comput.*, vol. 34, no. 9, pp. 863–866, Sep. 1985.

[25] J. Xu, X. Tang, and W.-C. Lee, "Time-critical on-demand data broadcast: Algorithms, analysis, and performance evaluation," *IEEE Trans. Parallel and Distributed Syst.*, vol. 17, no. 1, pp. 3–14, Jan. 2006.

The Capacity Region of a Class of Discrete Degraded Interference Channels

Nan Liu, *Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

***Abstract*—We provide a single-letter characterization for the capacity region of a class of discrete degraded interference channels (DDICs). The class of DDICs considered includes the DADIC studied by Benzel in 1979. We show that for the class of DDICs studied, encoder cooperation does not enlarge the capacity region, and therefore, the capacity region of the class of DDICs is the same as the capacity region of the corresponding degraded broadcast channel.**

***Index Terms*—Capacity region, degradedness, interference channel.**

I. Introduction

In wireless communications, where multiple transmitter and receiver pairs share the same communication medium, interference is unavoidable. How to best manage interference coming from other users and how not to cause too much interference to other users while maintaining the quality of communication is a challenging question and of a great deal of practical interest.

To be able to understand the effect of interference on communications better, an interference channel (IC) has been introduced in [2]. The IC is a simple network consisting of two pairs of transmitters and receivers. Each pair wishes to communicate at a certain rate with negligible probability of error. However, the two communications interfere with each other. To best understand the management of interference, we need to find the capacity region of the IC. However, the problem of finding the capacity region of the IC is essentially open except in

Manuscript received October 26, 2006; revised March 13, 2008. Published August 27, 2008 (projected). This work was supported by the National Science Foundation under Grants CCR 03-11311, CCF 04-47613, and CCF 05-14846. The material in this correspondence was presented in part at the 44th Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, September 2006.

N. Liu was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA. She is now with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: nanliu@stanford.edu).

S. Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ulukus@umd.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Color version of Figure 1 in this correspondence is available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.928267