

# Delay Guarantees for Throughput-optimal Wireless Link Scheduling

Koushik Kar  
Rensselaer Polytechnic Institute  
Troy, NY 12180  
Email: koushik@ecse.rpi.edu

Xiang Luo  
Rensselaer Polytechnic Institute  
Troy, NY 12180  
Email: luox3@rpi.edu

Saswati Sarkar  
University of Pennsylvania  
Philadelphia, PA 19104  
Email: swati@seas.upenn.edu

**Abstract**—We consider the question of obtaining tight delay guarantees for throughput-optimal link scheduling in arbitrary topology wireless ad-hoc networks. We consider two classes of scheduling policies: 1) a maximum queue-length weighted independent set scheduling policy, and 2) a randomized independent set scheduling policy where the independent set scheduling probabilities are selected optimally. Both policies stabilize all queues for any set of feasible packet arrival rates, and are therefore throughput-optimal. For these policies and i.i.d. packet arrivals, we show that the average packet delay is bounded by a constant that depends on the *chromatic number* of the interference graph, and the overall load on the network. We also prove that this upper bound is asymptotically tight in the sense that there exist classes of topologies where the expected delay attained by any scheduling policy is lower bounded by the same constant. Through simulations we examine the scaling of the average packet delay with respect to the overall load on the network, and the chromatic number of the link interference graph.

## I. INTRODUCTION

Recent proliferation of commercial wireless services has created large scale demands for transmission of traffic like multimedia, voice and video that require stringent quality-of-service (throughput, delay etc.) guarantees. Intelligent scheduling of wireless links is imperative for providing such guarantees. The main challenge in scheduling wireless links is that multiple links in a vicinity can not successfully transmit simultaneously. Efficient resolution of scheduling constraints is the main bottleneck in providing analytical performance guarantees.

In a seminal work, Tassiulas *et al.* [25] obtained a link scheduling policy that attains the maximum possible throughput in presence of arbitrary scheduling constraints, by scheduling in each time slot an independent set (in the link interference or conflict graph) that has the maximum aggregate queue length. This policy, referred to as *Maximum Weighted Scheduling (MWS)* henceforth, schedules at any given time instant (a) the set of links that can be simultaneously scheduled while satisfying the scheduling constraints, and (b) has the maximum sum of queue lengths among all such sets.

Obtaining delay guarantees is substantially more difficult than obtaining throughput guarantees, which is itself a challenging problem, due to the following reasons. Throughput guarantees can be obtained by any scheduling policy as long as it ensures that the expected time intervals between successive

instants in which the system is empty is finite. However, obtaining delay guarantees is contingent upon ensuring that the above expected duration is low. Specifically, consider a family of variants of MWS which does not schedule *any* link in the system if the queue length of every link is below a certain threshold, say  $L$ . Any such variant which has a finite value of  $L$  ensures that the above expected duration is finite, and therefore maximizes the throughput. Yet, the above expected duration, and therefore the delay, attained by any such variant monotonically increases with increase in  $L$ , and thus no such variant minimizes the delay as long as  $L$  is a positive integer.

Some recent insightful results have advanced our understanding in this area. Neely *et al.* [19] considered a specific scheduling policy, maximal scheduling, and showed that if the arrival traffic is in the stability region of maximal scheduling, the expected delay under maximal scheduling is  $O(\log N)$  where  $N$  is the number of links in the network. Maximal scheduling however provides poor throughput guarantees, as depending on the network topology, the stability region of maximal scheduling can become arbitrary small as compared to the stability region of throughput-optimal scheduling policies. Shah *et al.* [12] and Sarkar *et al.* [22] have shown that an  $O(1)$  expected queue length per link is attainable in special classes of networks. Asymptotic guarantees on queue lengths do not imply similar guarantees on delay, and more importantly, the above guarantees do not apply for arbitrary network topologies.

Since both throughput and delay are important performance metrics in wireless networks, we seek to obtain provable guarantees on expected delay using policies that maximize throughput. We focus on the following two policies: 1) MWS, 2) a *Randomized Scheduling (RS)* policy that schedules (independent sets of) links with a fixed probability irrespective of the queue length of the links. MWS is guaranteed to attain maximum throughput [25], has been empirically observed to attain low delay, and does not use any information about the arrival statistics in the scheduling process. RS is also guaranteed to attain maximum throughput for appropriate choice of scheduling probabilities, can be implemented without any knowledge of global network state, but requires the knowledge of arrival statistics for enabling the computation of the optimum scheduling probabilities. We prove that in any network  $\mathcal{N}$  the expected delay attained by both MWS and

RS is  $O(C(\mathcal{N})/\beta)$ , where  $C(\mathcal{N})$  is the *chromatic number of the link interference graph* for network  $\mathcal{N}$ , and  $1 - \beta$  ( $0 < \beta < 1$ ) is a measure of the *load on the network* (Section III). More precisely,  $C(\mathcal{N})$  represents the minimum number of independent sets (“colors”) into which the link interference graph of network  $\mathcal{N}$  can be partitioned, and  $\beta$ , henceforth referred to as the *arrival slack* in the system, is an appropriately defined measure of the distance between the arrival rate vector and the boundary of the network stability region.

Subsequently, we prove that there exists a class of network topologies where the expected delay attained by *any* scheduling policy is  $\Omega(C(\mathcal{N}))$  (Section IV). Thus, for constant  $\beta$ , the delay guarantees attained by MWS and RS are asymptotically tight. Note that one may intuitively expect that the MWS policy, which determines the link schedules based on queue lengths, will attain substantially lower delay than the RS policy. Thus, our results are somewhat counter-intuitive as they show that the static, possibly idling RS policy is able to attain the same delay guarantees as the dynamic, non-idling MWS policy, and that there exist classes of topologies where given  $\beta$ , the expected delay attained by RS differs from that of any other policy (including MWS) only by a constant factor. In Section V, we compare the delay performance of MWS and RS, and examine the scaling of the delay with respect to  $C(\mathcal{N})$  and  $\beta$ . We provide a brief survey of related literature in Section VI and conclude in Section VII.

## II. SYSTEM MODEL

We consider scheduling at the Medium Access Control (MAC) layer in a wireless network. We assume that time is slotted. A wireless network topology can be modeled as a directed graph  $G = (V, E)$ , where  $V$  and  $E$  respectively denote the sets of nodes and links, and  $|E| = N$ . A link exists from a node  $u$  to another node  $v$  if and only if  $v$  can receive  $u$ 's signals. The link set  $E$  depends on the transmission power levels of nodes and the propagation conditions in different directions.

*Definition 1:* A link  $i$  *interferes* with a link  $j$  if  $j$  can not successfully transmit a packet when  $i$  is transmitting.

A wireless network  $\mathcal{N}$  can be described by the topology  $G = (V, E)$  and the pairwise interference relations between the links.

*Definition 2:* The *interference set* of a link  $i$ ,  $S_i$ , is the set of links  $j$  such that either  $i$  interferes with  $j$  or  $j$  interferes with  $i$ .

Note that if  $j \in S_i$ , then  $i \in S_j$ .

*Definition 3:* The *interference graph*  $I^{\mathcal{N}} = (V_I^{\mathcal{N}}, E_I^{\mathcal{N}})$  of a network  $\mathcal{N}$  is an undirected graph in which the vertex set  $V_I^{\mathcal{N}}$  corresponds to the set of links in  $\mathcal{N}$  and there is an edge between two vertices  $i$  and  $j$  if  $j \in S_i$ .

*Definition 4:* An *independent set* in a graph is a subset of its vertices such that there does not exist an edge between any two vertices in the subset. Let  $J_1, \dots, J_M$  be the independent sets of  $I^{\mathcal{N}}$ , and let  $\vec{J}^i$  be the indicator vector representing any independent set  $J_i$ . Let  $\mathcal{J} = \{J_1, \dots, J_M\}$ .

*Definition 5:* A *coloring* of a graph is allocation of colors to vertices of the graph such that no two vertices that have an edge between them is assigned the same color. The *chromatic number* of a graph is the minimum number of colors required for coloring the graph. Equivalently, it is the minimum number of independent sets of a graph that can partition its vertex set.

Let  $\mathcal{C}(\mathcal{N}) = \{V_1, \dots, V_{C(\mathcal{N})}\}$  represent a minimum coloring of the link interference graph of network  $\mathcal{N}$ ,  $I^{\mathcal{N}}$ , where  $V_1, \dots, V_{C(\mathcal{N})}$  are the subsets of the vertices of  $I^{\mathcal{N}}$  that have been assigned the same color. Clearly,  $\mathcal{C}(\mathcal{N}) \subseteq \mathcal{J}$ , and  $C(\mathcal{N}) = |\mathcal{C}(\mathcal{N})|$  represents the chromatic number of  $I^{\mathcal{N}}$ .

At the MAC layer, each packet flow (session) can be assumed to span a single link. In the following discussion, therefore, we only consider single-hop flows. We now describe the arrival process for the single-hop flows (links). Let  $A_i(t)$  be the number of packets that link  $i$  generates in interval  $(t, t + 1]$ ,  $i = 1, \dots, N$ . We assume that for each  $i$ ,  $A_i(t)$  is i.i.d. across slots  $n$ , and  $\mathbf{E}(A_i(t)) = \lambda_i$ , where  $\lambda_i$  is referred to as the *arrival rate* of link  $i$ . We also assume that  $\mathbf{E}(A_i^2(t)) \leq \gamma \mathbf{E}(A_i(t))$ , where  $\gamma$  is a constant that depends on the distribution of the arrival process. A sufficient (but not necessary) condition for this to hold is that the maximum number of packets that arrive in a slot is upper bounded by a constant,  $\gamma$ .

*Definition 6:* The *arrival rate vector*  $\vec{\lambda}$  is an  $N$ -dimensional vector whose components are the arrival rates.

*Definition 7:* A *scheduling policy* is an algorithm that decides in each slot the subset of links that would transmit packets in the slot. Clearly, a scheduling policy must select an element of  $\mathcal{J}$  in each slot.

Every packet has a transmission time of one slot. Thus, if a backlogged link is scheduled in a slot, it transmits a packet in the slot. We assume that any packet arriving in a slot may be transmitted in the next slot.

Let  $D_i(t)$  be the number of packets that link  $i$  transmits in interval  $(t, t + 1]$ ,  $i = 1, \dots, N$ . Clearly the transmissions depend on the scheduling policy. Let  $Q_i(t)$  be the number of packets that are waiting for transmission in link  $i$  at the beginning of slot  $t$ . Let  $\vec{Q}(t), \vec{A}(t), \vec{D}(t)$  be the queue length, arrival and departure vectors respectively, with components,  $Q_i(t), A_i(t), D_i(t)$  respectively. Then,

$$Q_i(t + 1) = Q_i(t) + A_i(t) - D_i(t). \quad (1)$$

We now describe two scheduling policies that we analyze in this paper: 1) maximum weighted scheduling (MWS), and 2) randomized scheduling (RS( $\vec{p}$ )). MWS considers the weight of an independent set  $J_i$  as the sum of the queue lengths of the links in  $J_i$ , and in each slot  $t$  schedules the independent set that has the maximum weight among all independent sets in  $\mathcal{J}$ . In each slot  $t$ , RS( $\vec{p}$ ), schedules independent set  $J_i$  with probability  $p_i$  irrespective of the queue lengths of the links, and the schedules selected in different slots are mutually independent. Here,  $\vec{p}$  is a  $M$ -dimensional probability vector such that  $\sum_{i=1}^M p_i = 1$  and  $p_i \geq 0$  for each  $i$ . Note that since RS schedules independent sets at random without

considering queue length information, it can possibly schedule an independent set with all empty queues, while there are other independent sets with backlogged queues in the system. In contrast, MWS is non-idling, in the sense and it will never serve an independent set with all empty queues as long as there are backlogged queues in the system.

*Definition 8:* The network is said to be *stable* if  $\lim_{T \rightarrow \infty} \sum_{n=1}^T Q_i(t)/T$  is finite.

*Definition 9:* The *stability region* of a scheduling policy is the set of arrival rate vectors for which the network is stable when the policy is used. An arrival rate vector  $\vec{\lambda}$  is said to be *feasible* if it is in the stability region of some scheduling policy. The *network stability region*  $\Lambda$  is the set of all feasible arrival rate vectors.

Note that  $\Lambda$  depends on the network  $\mathcal{N}$ . Define

$$\Lambda_\beta = \left\{ \vec{\lambda} : \vec{\lambda} = \sum_{i=1}^M w_i \vec{J}^i, \text{ for some } w_1, \dots, w_M \right. \\ \left. \text{such that } w_i \geq 0, \sum_{i=1}^M w_i = 1 - \beta \right\}.$$

Tassiulas *et al.* [25] has shown that  $\Lambda_\beta \subset \Lambda$  for each  $\beta \in (0, 1]$  and  $\Lambda \subseteq \Lambda_\beta$  when  $\beta = 0$ . From the above definition, an arrival rate vector  $\vec{\lambda}$  is in  $\Lambda_\beta$  if  $(1 - \sum_{i=1}^M w_i) = \beta$ . In other words, if an arrival rate vector  $\vec{\lambda}$  is in  $\Lambda_\beta$  then its “distance” from the boundary of the network stability region is  $\beta$ . We thus refer to  $\beta$  as the *arrival slack* in the system when the arrival vector  $\vec{\lambda}$  is in  $\Lambda_\beta$ .

*Definition 10:* The *expected delay* in a network is the expected number of time slots that elapse between the arrival and departure of a packet.

In the next section, we upper bound the expected delays of MWS and RS( $\vec{p}$ ) for appropriate choice of  $\vec{p}$ .

### III. UPPER BOUNDS ON EXPECTED DELAY

The maximum weighted scheduling policy is known to attain the the network stability region, and thereby maximizes network throughput. With an appropriate choice of scheduling probabilities, the randomized scheduling policy can be easily shown to attain maximum throughput as well. We now show that for any given network  $\mathcal{N}$ , when the arrival slack is  $\beta$ , the expected delays attained by these policies are  $O(C(\mathcal{N})/\beta)$  (Theorems 1, 2).

*Theorem 1:* Consider a network  $\mathcal{N}$ , and a  $\vec{\lambda} \in \Lambda_\beta$  where  $\beta \in (0, 1)$ . Then, the expected delay attained by MWS in  $\mathcal{N}$  is at most  $\frac{\gamma+1}{2} \frac{C(\mathcal{N})}{\beta}$ .

Recall that  $\gamma$  is an upper bound on the ratio of the second and first moments of the arrival distribution.

*Proof:* Consider a quadratic Lyapunov function  $V : R^N \rightarrow R$  such that  $V(\vec{x}) = \sum_{i=1}^N x_i^2$ . Using (1), we get

$$\begin{aligned} & V(\vec{Q}(t+1)) - V(\vec{Q}(t)) \\ &= \left( \vec{Q}(t+1) - \vec{Q}(t) \right)^T \left( \vec{Q}(t+1) + \vec{Q}(t) \right) \\ &= \left( \vec{A}(t) - \vec{D}(t) \right)^T \left( 2\vec{Q}(t) + \vec{A}(t) - \vec{D}(t) \right) \end{aligned}$$

$$\begin{aligned} &= 2 \left( \vec{A}(t) - \vec{D}(t) \right)^T \vec{Q}(t) \\ &\quad + \left( \vec{A}(t) - \vec{D}(t) \right)^T \left( \vec{A}(t) - \vec{D}(t) \right) \\ &\leq 2 \left( \vec{A}(t) - \vec{D}(t) \right)^T \vec{Q}(t) + \left( \vec{A}(t) \right)^T \left( \vec{A}(t) \right) \\ &\quad + \left( \vec{D}(t) \right)^T \left( \vec{D}(t) \right). \end{aligned} \quad (2)$$

Now, since  $\vec{\lambda} \in \Lambda_\beta$ ,  $\beta \in (0, 1)$ ,  $(\vec{Q}(t), \vec{A}(t), \vec{D}(t))$  constitutes a positive recurrent Markov chain [25]. We take expectations of both sides of (2) under the stationary distribution of this Markov chain.

First consider the left hand side, and note that

$$\mathbf{E} \left( V(\vec{Q}(t+1)) - V(\vec{Q}(t)) \right) = 0. \quad (3)$$

Now, consider the right hand side. Since  $\vec{\lambda} \in \Lambda_\beta$ , there exists  $w_1, \dots, w_M$  such that  $\sum_{i=1}^M w_i = 1 - \beta$  and  $\vec{\lambda} = \sum_{k=1}^M w_k \vec{J}^k$ . Thus,

$$\begin{aligned} \mathbf{E} \left( \left( \vec{A}(t) \right)^T \vec{Q}(t) | \vec{Q}(t) \right) &= \vec{\lambda}^T \vec{Q}(t) \\ &= \sum_{k=1}^M w_k (\vec{J}^k)^T \vec{Q}(t) \\ &\leq \left( \sum_i w_i \right) \max_{k=1}^M (\vec{J}^k)^T \vec{Q}(t) \\ &= (1 - \beta) \max_{k=1}^M (\vec{J}^k)^T \vec{Q}(t). \end{aligned}$$

Thus,

$$\mathbf{E} \left( \left( \vec{A}(t) \right)^T \vec{Q}(t) \right) \leq (1 - \beta) \mathbf{E} \max_{k=1}^M (\vec{J}^k)^T \vec{Q}(t). \quad (4)$$

Using the properties of MWS scheduling,  $\left( \vec{D}(t) \right)^T \vec{Q}(t) = \max_{k=1}^M (\vec{J}^k)^T \vec{Q}(t)$ . Thus,

$$\mathbf{E} \left( \vec{A}(t) - \vec{D}(t) \right)^T \vec{Q}(t) \leq -\beta \mathbf{E} \max_{k=1}^M (\vec{J}^k)^T \vec{Q}(t). \quad (5)$$

Since  $\mathbf{E} \left( A_i^2(t) \right) \leq \gamma \mathbf{E} \left( A_i(t) \right)$ ,

$$\mathbf{E} \left( \vec{A}(t) \right)^T \left( \vec{A}(t) \right) \leq \gamma \sum_{i=1}^N \lambda_i. \quad (6)$$

The last equality follows since  $\vec{A}(t)$  is independent of  $\vec{Q}(t), \vec{D}(t)$ .

Since the components of the departure vector are either 0 or 1,  $D_i^2(t) = D_i(t)$ . Thus,

$$\mathbf{E} \left( \vec{D}(t) \right)^T \left( \vec{D}(t) \right) = \mathbf{E} \left( \sum_{i=1}^N D_i(t) \right) = \sum_{i=1}^N \lambda_i. \quad (7)$$

We now justify the last equality. Since  $\vec{\lambda} \in \Lambda_\beta$  where  $\beta \in (0, 1)$  and MWS is used, the system is stable [25]. Thus, since the time average of the queue length in each link is finite, and the Markov chain  $(\vec{Q}(t), \vec{A}(t), \vec{D}(t))$  is positive recurrent,  $\mathbf{E} \left( \vec{D}(t) \right) = \vec{\lambda}$ . The equality follows.

Thus, taking expectations w.r.t. the steady state distribution of  $\vec{Q}(t)$  on both sides of (2), it follows from (3) to (7),

$$0 \leq -2\beta \mathbf{E} \max_{k=1}^M \left( \vec{J}^k \right)^T \vec{Q}(t) + (\gamma + 1) \sum_{i=1}^N \lambda_i.$$

Thus,

$$\mathbf{E} \max_{k=1}^M \left( \vec{J}^k \right)^T \vec{Q}(t) \leq \frac{(\gamma + 1) \sum_{i=1}^N \lambda_i}{2\beta}.$$

Thus,

$$\mathbf{E} \sum_{k \in J} Q_k(t) \leq \frac{(\gamma + 1) \sum_{i=1}^N \lambda_i}{2\beta} \text{ for any } J \in \mathcal{J}. \quad (8)$$

Now,

$$\begin{aligned} & \mathbf{E} \left( \sum_{k=1}^N Q_k(t) \right) \\ &= \mathbf{E} \left( \sum_{j=1}^{C(\mathcal{N})} \sum_{k \in V_j} Q_k(t) \right) \\ &\leq \sum_{j=1}^{C(\mathcal{N})} \frac{(\gamma + 1) \sum_{i=1}^N \lambda_i}{2\beta} \text{ (from (8) and since } V_j \in \mathcal{J}) \\ &= C(\mathcal{N})(\gamma + 1) \frac{\sum_i \lambda_i}{2\beta}. \end{aligned}$$

Since the expected delay is  $\mathbf{E} \left( \sum_{k=1}^N Q_k(t) \right) / \left( \sum_{i=1}^N \lambda_i \right)$ , the result follows. ■

*Theorem 2:* Consider a network  $\mathcal{N}$ , and  $\vec{\lambda} \in \Lambda_\beta$  where  $\beta \in (0, 1)$ . Assume that the arrival process is i.i.d. Bernoulli. Then, there exists a  $M$ -dimensional probability vector  $\vec{p}$ , such that the expected delay attained by RS( $\vec{p}$ ) in  $\mathcal{N}$  is at most  $\frac{C(\mathcal{N})}{\beta}$ .

*Proof:* Since  $\vec{\lambda} \in \Lambda_\beta$ , there exists  $w_1, \dots, w_M$  such that  $\sum_{i=1}^M w_i = 1 - \beta$  and  $\vec{\lambda} = \sum_{k=1}^M w_k \vec{J}^k$ . Recall that  $\mathcal{C}(\mathcal{N}) \subseteq \mathcal{J}$  represents a collection of independent sets that constitutes a minimum coloring of the link interference graph. Let  $p_k = w_k$  if  $J_k \in \mathcal{J} \setminus \mathcal{C}(\mathcal{N})$ , and  $p_k = w_k + \beta/C(\mathcal{N})$ , if  $J_k \in \mathcal{C}(\mathcal{N})$ . Note that  $p_k \geq 0$ , and  $\sum_{k=1}^M p_k = \sum_{k=1}^M w_k + |\mathcal{C}(\mathcal{N})|\beta/C(\mathcal{N})$ . Since  $|\mathcal{C}(\mathcal{N})| = C(\mathcal{N})$ ,  $\sum_{k=1}^M p_k = \sum_{k=1}^M w_k + \beta = 1$ . Thus,  $\vec{p}$  is a  $M$ -dimensional probability vector.

Now, when an independent set in  $\mathcal{J}$  is scheduled, all the links of  $\mathcal{N}$  that constitute the independent set are scheduled. Also, each link  $l$  is in one independent set in  $\mathcal{C}(\mathcal{N})$ . Thus, each link  $l$  is scheduled with probability  $\sum_{i=1}^M w_i \vec{J}_l^i + \beta/C(\mathcal{N})$ , where  $\vec{J}_l^i$  is the component corresponding to link  $l$  in  $\vec{J}^i$ , the indicator vector representing the independent set  $J_i$ . Since  $\lambda_l = \sum_{i=1}^M w_i \vec{J}_l^i$ ,  $l$  is scheduled with probability  $\lambda_l + \beta/C(\mathcal{N})$ . Since the arrival process is Bernoulli, from standard queueing analysis, the system is stable and expected delay for link  $l$  is  $\frac{1 - \lambda_l}{\lambda_l + \beta/C(\mathcal{N}) - \lambda_l}$  which is at most  $C(\mathcal{N})/\beta$ . The result follows. ■

Note that unlike the other results stated in the paper, Theorem 2 relies on an additional assumption that the arrival

process is Bernoulli. Computing the probability vector  $\vec{p}$  that attains the above delay guarantee requires finding the weights  $\{w_k\}$  from the arrival vector  $\vec{\lambda}$ , and finding the chromatic number of the interference graph. The rate decomposition problem to calculate the independent set weights  $\{w_k\}$  is not known to be polynomially solvable in the general case; in presence of primary interference constraints only (i.e., if two links can transmit together successfully as long as they do not have any common end node), this problem can be solved in polynomial time [10]. Finding the chromatic number requires solving a graph coloring problem, which is NP-hard [3]. It should be noted however that these complex calculations need not be done on a per slot basis – it can be done once at the very beginning, and the scheduling probabilities  $\{p_k\}$  thus computed can be used thereafter, until the network topology changes. In contrast, MWS requires solving the NP-hard maximum-weighted independent set problem at each scheduling instant.

Thus, for a given  $\beta$ , the upper bound on the expected delay of MWS and RS depends on the network topology only through the chromatic number of the interference graph. Therefore, if the size of the network is increased without increasing the chromatic number, the upper bound does not change. Furthermore,  $C(\mathcal{N})$  is significantly less than  $N$  in a large class of topologies. For example, consider a network where  $n$  nodes are uniformly deployed in a square of side one unit so as to constitute a random geometric graph, i.e., there exists a link between two nodes if and only if the distance between them is less than a given number  $D$ , and two links interfere if and only if an end node of one is within a distance  $D$  of that of another. It is well known that if  $D^2 \sim O(\frac{\log n}{n})$ , then for large  $n$ , the network is connected and the throughput is maximized with probability one [9]. Then, for large  $n$ , the degree of each link in the interference graph is  $O(\log n)$  with probability one. Since the chromatic number in any graph can exceed the maximum degree of its nodes by at most 1 [11], for a given  $\beta$ , the expected delay attained by the above policies is  $O(\log n)$ , or equivalently  $O(\log N)$ , in such networks.

#### IV. LOWER BOUNDS ON EXPECTED DELAY

We now obtain a lower bound on the expected delay of an arbitrary policy. Specifically, we prove the following theorem.

*Theorem 3:* For any real numbers  $\beta \in (0, 1)$ ,  $\epsilon \in (0, 1)$ , and any positive integer  $\tilde{C}$ , there exists a network  $\mathcal{N}$  with  $C(\mathcal{N}) = \tilde{C}$  and an arrival rate vector in  $\Lambda_\beta$ , such that the expected delay attained in  $\mathcal{N}$  by any scheduling policy is at least  $\frac{(1-\epsilon)^2 \tilde{C}}{(1+\epsilon)^2}$ .

The above lower bound holds irrespective of the arrival slack  $\beta$ , and its value does not depend on  $\beta$ . An important question therefore is whether a tighter lower bound can be obtained by exploiting the relation between the expected delay and  $\beta$ . We show that this is indeed the case for any randomized scheduling. Specifically, we show the following result.

*Theorem 4:* For any real number  $\beta \in (0, 1)$ , any probability vector  $\vec{p}$ , and any positive integer  $\tilde{C}$ , there exists a network  $\mathcal{N}$

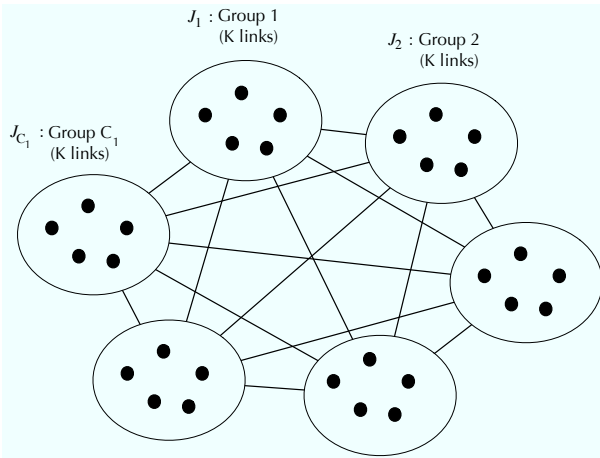


Fig. 1. Group based interference graph topology. There is no interference between links in the same group. An edge between two groups implies that a link in any one group interferes with every link in the other group.

with  $C(\mathcal{N}) = \tilde{C}$  and an arrival rate vector in  $\Lambda_\beta$ , such that the expected delay attained in  $\mathcal{N}$  by  $\text{RS}(\bar{p})$  is at least  $\frac{\tilde{C}-1}{\beta} + 1$ .

Note that for large  $\tilde{C}$ , the lower bound  $\frac{\tilde{C}-1}{\beta} + 1 \approx \frac{\tilde{C}}{\beta}$ . Proving a similar lower bound that involves the arrival slack  $\beta$  for an arbitrary scheduling policy remains open.

We now prove Theorem 3.

*Proof:* We first describe a network  $\mathcal{N}$  with  $C(\mathcal{N}) = \tilde{C}$ . The network consists of  $\tilde{C}$  disjoint groups (sets) of links,  $J_1, \dots, J_{\tilde{C}}$ , each of size  $K$ . (The value for  $K$  will be specified later.) Thus each  $J_i$  consists of  $K$  links,  $J_i \cap J_l = \phi$  if  $i \neq l$ . Here, links  $i, l$  can be served simultaneously if and only if  $i, l \in J_w$  for some  $w$ . The interference graph of this network has been shown in in Figure 1 (set  $C_1 = \tilde{C}$ ). Sets  $J_1, \dots, J_{\tilde{C}}$  and their subsets constitute the independent sets of the interference graph of  $\mathcal{N}$ . Clearly, the colors assigned to links in  $J_i$  and  $J_l$  must be different if  $i \neq l$ . Thus, the chromatic number of the interference graph is at least  $\tilde{C}$ . Note that the interference graph can be colored using color  $i$  for links in  $J_i$ ,  $i = 1, \dots, \tilde{C}$ . Thus, the chromatic number of the interference graph is at most, and hence equal to,  $\tilde{C}$ .

We now consider a large class of arrival processes for which the arrival rate vector is in  $\Lambda_\beta$  of  $\mathcal{N}$ . Let  $\alpha = (1 - \beta)/\tilde{C}$ . Clearly  $\alpha \in (0, 1)$ . Consider an arbitrary probability distribution  $F$  on non-negative integers with expectation  $\alpha$  that satisfies the following technical condition: the moment generating function  $Z(\tau)$  of the distribution  $F$  (i.e.,  $\mathbf{E}(\exp(\tau X))$  where  $X$  is a random variable with distribution  $F$ ) is finite in some neighborhood of  $\tau = 0$ . Note that  $F$  can be selected from a large class of probability distributions which consists of, but is not limited to, Bernoulli( $\alpha$ ), Poisson( $\alpha$ ), Binomial( $x, y$ ) with  $xy = \alpha$ , etc. Let packets arrive in each link as per temporally and mutually independent random processes with distribution  $F$  each. Note that the corresponding arrival rate vector  $\vec{\lambda} = (\alpha, \dots, \alpha)$  can be expressed as  $\vec{\lambda} = \sum_i w_i \vec{J}_i$ , where  $w_i = \alpha$ ,  $i = 1, \dots, \tilde{C}$ , and  $w_i = 0$  for  $i > \tilde{C}$ . Since  $\sum_i w_i = \alpha \tilde{C} = 1 - \beta$ ,  $\vec{\lambda} \in \Lambda_\beta$ .

Consider an arbitrary slot  $t$  and the packets that arrive in this slot. Let the delay of any such packet be  $\bar{D}$ , and the

expectation of  $\bar{D}$  be  $\mathbf{E}\bar{D}$ . Then, the expected delay of any packet in the system is  $\mathbf{E}\bar{D}$ . We therefore need to show that  $\mathbf{E}\bar{D} \geq (1 - \epsilon)^2 \tilde{C} / (2(1 + \epsilon))$ . Note that  $\bar{D}$  is minimized if the system schedules the transmissions so as to minimize the delay experienced by the packets that arrive in this slot. Such a schedule will not serve a link  $i$  if link  $i$  does not have any of the above packets and another link  $j$  has at least one of these packets. We also assume that the first time a link is served after  $t$  all packets that arrived in  $t$  in the link are served from it; the delay can only increase otherwise (which may happen when a link receives more than one packet in  $t$ ).

Let  $X_i$  be the total number of packets received by links in  $J_i$  in  $t$ , for  $i = 1, \dots, \tilde{C}$ . Consider an event  $A$  in which  $K\alpha(1 - \epsilon) \leq X_i \leq K\alpha(1 + \epsilon)$  for each  $i$ . Clearly,  $\mathbf{E}\bar{D} \geq \Pr(A)\mathbf{E}(\bar{D}|A)$ . Now, without loss of generality, let links in  $J_1$  be served in slot  $t+1$ , links in  $J_2$  be served in slot  $t+2$  and so on, until all links that have received packets in slot  $t$  have been served. Then,  $\mathbf{E}(\bar{D}|A) \geq \sum_{i=1}^{\tilde{C}} i \mathbf{E}(\frac{X_i}{\sum_{i=1}^{\tilde{C}} X_i} | A)$ . Now,  $\mathbf{E}(\frac{X_i}{\sum_{i=1}^{\tilde{C}} X_i} | A) \geq \frac{K\alpha(1-\epsilon)}{CK\alpha(1+\epsilon)} = \frac{1-\epsilon}{C(1+\epsilon)}$ . Thus,  $\mathbf{E}(\bar{D}|A) \geq \frac{1-\epsilon}{C(1+\epsilon)} \sum_{i=1}^{\tilde{C}} i \geq \frac{\tilde{C}}{2} \frac{1-\epsilon}{1+\epsilon}$ .

We now evaluate  $\Pr(A)$ . From large deviation results, for all large enough  $K$ , the probability that  $X_i \notin [K\alpha(1 - \epsilon), K\alpha(1 + \epsilon)]$  is at most  $3\exp(-K\nu)$  where  $\nu$  is a positive constant that depends on the distribution  $F$  and  $\alpha, \epsilon$  (Section 5.11 [8]). Using union bound, for all large enough  $K$ , the probability that  $X_i \notin [K\alpha(1 - \epsilon), K\alpha(1 + \epsilon)]$  for at least one  $i$  is at most  $3\tilde{C}\exp(-K\nu)$ , which is at most  $\epsilon$  for all large enough  $K$ . Thus,  $\Pr(A) \geq 1 - \epsilon$  for all large enough  $K$ . The result follows. ■

The proof reveals that the lower bound on delay holds for a large class of arrival distributions. Thus, intuitively, there are no “good” arrival processes for which the above lower bound on the expected delay attained by any policy does not apply. The lower bound however applies only for a class of topologies. Thus, there may be “good” topologies where some policy may attain delays lower than the above bound. For example, consider the case of a full mesh link interference graph (i.e., any two links interfere with each other), where the chromatic number is equal to the number of links,  $N$ . In this case, our lower bound is roughly  $N/2$ , and the upper bound on the expected delay obtained for MWS in the last section is  $O(N/\beta)$ . However, any work-conserving scheduling policy (including MWS) attains an expected delay of  $(1/\beta)$  in this case for Poisson arrivals. This shows that there exist topology classes where the lower bound may not apply, or our upper bounds derived in Section III may be loose, for certain arrival patterns. An interesting direction for future research is to identify topologies, or rather characteristic properties of topologies, that allow at least some scheduling policies to attain low delays.

We prove Theorem 4 next.

*Proof:* We first describe a network  $\mathcal{N}$  with  $C(\mathcal{N}) = \tilde{C}$ . The network consists of  $\tilde{C}$  links such that any two links in the network interfere with each other. Thus, the link interference

graph is a clique of size  $\tilde{C}$ , and hence its chromatic number is  $\tilde{C}$ . The independent sets of the interference graph consist of  $J_0, J_1, \dots, J_{\tilde{C}}$ , where  $J_0 = \phi$ , and  $J_i$  consists only of link  $i$  when  $i \geq 1$ .

We now consider a Bernoulli arrival process for which the arrival rate vector is in  $\Lambda_\beta$  of  $\mathcal{N}$ . Let  $\alpha = (1 - \beta)/\tilde{C}$ . Clearly,  $\alpha \in (0, 1)$ . Let the arrival process at each link be Bernoulli( $\alpha$ ), independent of the arrival processes at other links. The arrival processes in different slots are also independent. Note that the corresponding arrival rate vector  $\vec{\lambda} = (\alpha, \dots, \alpha)$  can be expressed as  $\vec{\lambda} = \sum_{i=0}^{\tilde{C}} w_i \vec{J}^i$ , where  $w_i = \alpha$ ,  $i = 1, \dots, \tilde{C}$ , and  $w_0 = 0$ . Since  $\sum_i w_i = \alpha|\tilde{C}| = 1 - \beta$ ,  $\vec{\lambda} \in \Lambda_\beta$ .

Consider  $\text{RS}(\vec{p})$  for an arbitrary  $(\tilde{C} + 1)$ -dimensional probability vector  $\vec{p}$ . Then, in each slot,  $\text{RS}(\vec{p})$  serves link  $i$  with probability  $p_i$  for  $i \geq 1$ , and the service opportunities are temporally independent. From standard queueing analysis, the expected delay in link  $i$  is  $\frac{1-\alpha}{p_i-\alpha}$ . Since all links have equal arrival rates, the overall expected delay is  $\frac{1}{\tilde{C}} \sum_{i=1}^{\tilde{C}} \frac{1-\alpha}{p_i-\alpha}$ . Thus, if  $\vec{p}^*$  minimizes the expected delay for  $\text{RS}(\vec{p})$  among all possible choices for  $\vec{p}$ , then  $\vec{p}^*$  must be the optimum solution of the following convex optimization problem:

$$\begin{aligned} & \text{minimize} && \frac{1}{\tilde{C}} \sum_{i=1}^{\tilde{C}} \frac{1-\alpha}{p_i-\alpha}, \\ & \text{subject to:} && p_i \geq 0, \quad i = 1, \dots, \tilde{C}, \\ & && \sum_{i=1}^{\tilde{C}} p_i \leq 1. \end{aligned}$$

It is easy to show that the optimum solution to the above symmetric convex optimization problem is given by  $p_i^* = 1/\tilde{C}$ . Since  $\alpha = (1 - \beta)/\tilde{C}$ , the expected delay under  $\text{RS}(\vec{p}^*)$  is  $\frac{(1-\frac{1-\beta}{\tilde{C}})\tilde{C}}{\beta} = \frac{\tilde{C}-(1-\beta)}{\beta} = \frac{\tilde{C}-1}{\beta} + 1$ . ■

## V. SIMULATION STUDY

In this section, we compare the delay performance of the two scheduling algorithms, MWS and RS, in a few representative network topologies. We also evaluate through simulations how the average delay attained by these algorithms varies as a function of the quantities  $C(\mathcal{N})$  and  $\beta$ .

One major difficulty in evaluating the performance of MWS and RS in large scale, arbitrary topology networks (like networks formed by placing nodes at random in an area) is the computational complexity of the scheduling algorithms. Both MWS and RS require solving NP-hard problems – maximum weighted independent set for MWS, and graph coloring for RS. In our simulation study, therefore, we focus on small topologies like “star” and “line” link interference graphs, where these problems can be solved in polynomial time. To study the scaling of delay with chromatic number, we consider large topologies, but those that exhibit a regularity so that maximum weighted independent set and chromatic number can be calculated easily.

The topologies that we consider are parametrized by a constant  $K$ , which roughly represents the size of the maximum

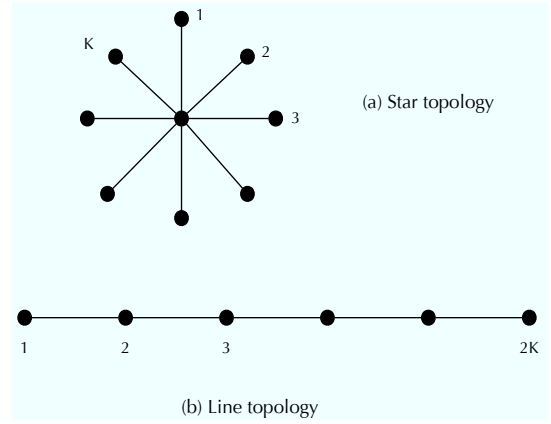


Fig. 2. Star and line link interference graph topologies.

independent set in the link interference graph. The first two of these topologies as shown in Figure 2 are the star and the line topologies. Note that these topologies correspond to the link interference graphs (not the actual networks), where each node represents a link and each edge represents a (symmetric) interference relationship between the corresponding links. The chromatic number for both these topologies is 2. For the star topology, note that there are only two maximal independent sets, with sizes one (the central link) and  $K$  (the peripheral links). For the line network, while there are many *maximal* independent sets, there are only two *maximum-size* independent sets, each with size  $K$  (the set of odd nodes, and the set of even nodes). For this line topology, the maximum weight independent set calculation, as required by MWS, is performed in polynomial time using a dynamic programming procedure.

The third topology we consider is the one in Figure 1, where the link interference graph can be viewed as a collection of  $C_1$  groups of links, each of size  $K$ , where a link does not interfere with any other link in its group, but interferes with all links in the other  $C_1 - 1$  groups. This interference graph topology is the same as the one considered in Section IV to obtain a lower bound on the delay performance of any scheduling policy. Note that there are  $C_1$  maximal independent sets in this topology, one corresponding to each of the  $C_1$  groups, and each of these sets consists of  $K$  links. Also note that at any time, only one of the  $C_1$  maximal independent sets can be scheduled, and  $C_1$  represents the chromatic number of this link interference graph. The last topology we consider is similar to the earlier one, but with the difference that the  $K$  links in group  $i$  ( $1 \leq i \leq C_1$ ) do not interfere with links in group  $(i + 1) \bmod C_1$ . Note that while the third topology consists of disjoint maximal independent sets, the fourth topology introduces overlaps between some of the independent sets, which is more likely to be the case in practice. Note that the chromatic number of this interference graph is  $\lceil C_1/2 \rceil$ . We use these topologies to study how average delay scales with the chromatic number, by varying  $C_1$ . Note that the difference in the chromatic numbers between the last two topologies (for the same  $C_1$ ,  $K$ , and therefore same number of links) is caused due to lower interference in the latter topology. We assume that packet arrivals occur at each queue according to a Bernoulli



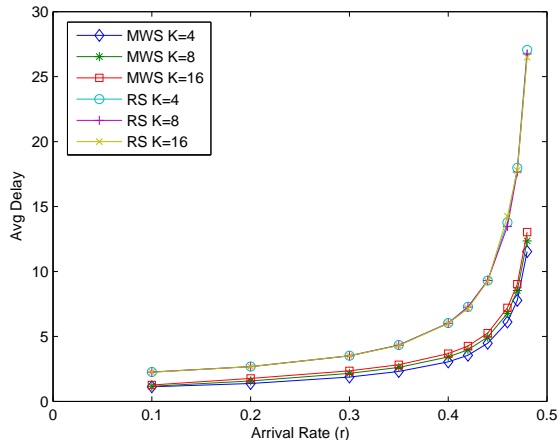


Fig. 3. Average delay vs network load, for the star interference graph topology.

process with probability  $r$ .

Figures 3 and 4 plot the average delay attained by MWS and RS with the arrival rate, for different values of  $K$ . The delay curves look similar to those of an  $M/M/1$  queue, and sharply increases as the arrival rate approaches the boundary of the stability region (which can be easily computed as  $r = 0.5$  for both topologies). We observe that the delay attained by MWS is significantly better than the delay for RS at the same network load. This is intuitively expected, since unlike MWS, RS is a “static” policy which does not take into account queue lengths, and can even be non work-conserving. The results show that the average delay for RS is insensitive to  $K$  (which corresponds to the maximum independent set size in both topologies): this happens since each queue behaves independently of the others, owing to random arrivals and random scheduling. Also the delay for MWS increases as  $K$  increases, for the same arrival rate, as we would expect from intuition.

Figures 5 and 6 replot the results in Figures 3 and 4, so that the y-axis corresponds to the logarithm of the average delay, and the x-axis corresponds to  $\log(1/\beta)$ , where  $\beta$  is the arrival slack. We observe an almost linear relationship; furthermore, the slopes for the curves for MWS and RS are roughly the same. This is in agreement with the upper bounds derived in Section III which show a scaling of  $O(1/\beta)$  for the average delay for a fixed  $C(\mathcal{N})$ .

The average delay vs arrival rate (network load) plots for the group based interference graph topologies are very similar to those of the star and line topologies considered above. Thus we only show how the average delay scales with the chromatic number in these cases. Figures 7 and 8 show the results for the two topologies, when the network load is kept at 80% of the maximum (or  $\beta = 0.2$ ). This implies  $r = 0.8/C$  for the interference graph topology with disjoint maximal independent sets, and  $r = (2 \times 0.8)/C$  for the interference graph topology with overlapping maximal independent sets. The figures show that the average delay scales linearly with the chromatic number when network load is kept constant.

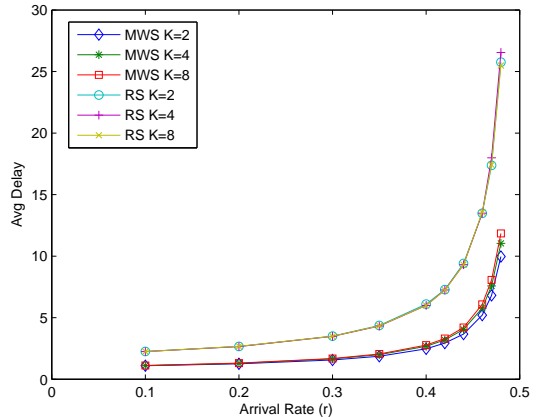


Fig. 4. Average delay vs network load, for the line interference graph topology. Note that the total number of links in the network is  $2K$ .

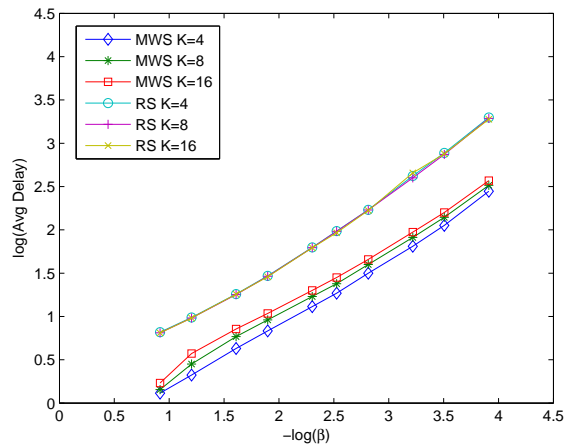


Fig. 5. Logarithm of average delay vs logarithm of arrival slack ( $\beta$ ), for the star interference graph topology.

## VI. RELATED WORK

Tassiulas *et al.* have characterized the maximum throughput region and also provided a scheduling strategy that attains this throughput region in any given wireless network [25]. Subsequently, several policies have been shown to attain (in some cases, for specific topology classes and interference constraints), either the maximum throughput region [1], [5], [18], [21], [23], [24] or a guaranteed fraction thereof [2], [4], [13], [15], [26], while requiring lower computation time. However, whether these algorithms are able to provide low delay guarantees remains largely unknown.

There has been some recent interesting work on characterizing the delay performance of scheduling algorithms, or formulating scheduling and routing policies in random networks that attain order-optimal delay [6], [7], [20], [14], [16], [17]. Most of these works do not consider queuing delay, and focus on attaining order-optimal packet delivery delay in presence of node mobility. Scheduling policies that exploit node mobility as in these works are in general unlikely to attain low delay guarantees for the static network setting that

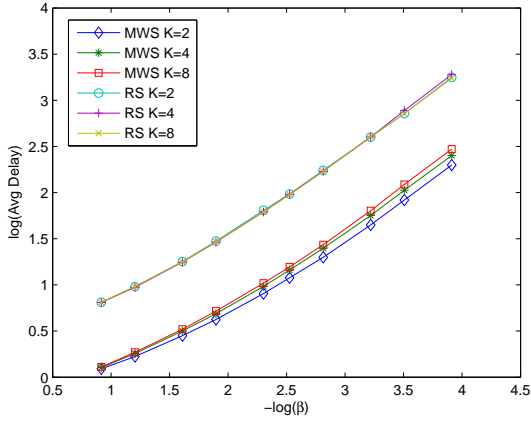


Fig. 6. Logarithm of average delay vs logarithm of arrival slack ( $\beta$ ), for the line interference graph topology. Note that the total number of links in the network is  $2K$ .

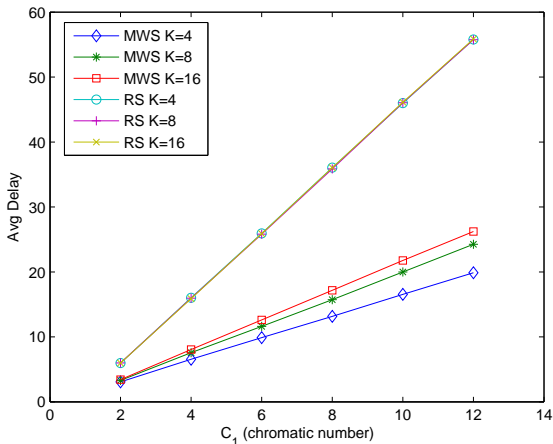


Fig. 7. Average delay vs  $C_1$ , for the group based interference graph topology shown in Figure 1. The chromatic number for this network topology is  $C_1$ .

we consider in this paper.

Analysis of scheduling algorithms in terms of expected queuing delay or expected queue lengths is considered in [19], [12], [22]. In [19], the author obtains a  $O(\log N)$  upper bound on the expected delay, but only for maximal scheduling policies which (depending on network topology and interference relationships) can attain only a small fraction of the overall stability region. Specifically, the author shows that if the arrival rate vector is such that for each link  $l$ , the sum of the arrival rates of  $l$  and the links that interfere with  $l$  is less than or equal to 1 (which ensures that maximal scheduling stabilizes the network), then maximal scheduling attains  $O(\log N)$  expected delay. The set of arrival rate vectors that satisfy the above sufficient condition can however be significantly smaller than the stability region of the network. Specifically, in the topology considered in the proof of Theorem 3 (Figure 1), the arrival rate vector  $\vec{\lambda} = (\alpha, \dots, \alpha)$  is in the network stability region provided  $\tilde{C}\alpha < 1$ . However,  $\vec{\lambda}$  satisfies the above sufficiency condition for maximal scheduling only if  $\alpha (K(\tilde{C} - 1) + 1) < 1$ . Thus,

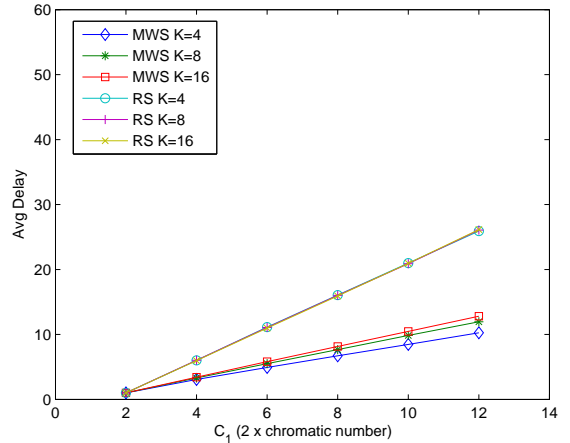


Fig. 8. Average delay vs  $C_1$ , for the group based interference graph topology shown in Figure 1, but without interference between links in adjacent groups. The chromatic number for this network topology is  $C_1/2$ .

this sufficiency condition reduces the stability region by a factor of at least  $K$  (approximately) and  $K$  can be arbitrarily large. Therefore the delay guarantee obtained for maximal scheduling only applies if the arrival rate vector belongs to a stability region which is possibly much smaller than the maximum stability region. While our lower bound on the delay obtained by any policy (Theorem 3) may appear to contradict the  $O(\log N)$  upper bound on delay for maximal scheduling, there is no contradiction in reality. We prove Theorem 3 by choosing an arrival rate vector which does not satisfy the above sufficiency condition (and therefore the above upper bound does not apply to this case), but is nevertheless in the stability region of the network.

In [12], [22], the authors show an  $O(1)$  upper bound on expected queue length per link for scheduling policies they propose, but their results apply only to the special class of non-expanding graphs (which include random geometric interference graphs). In [12], the authors also show that there exist arrival processes for which  $\Omega(1)$  queue length is necessary. Note that a bound on the queue length does not necessarily translate to that on the expected delay (since expected delay is expected queue length divided by the expected arrival rate), without additional assumptions on how the arrival rate scales with increase in the size of the topology. In contrast, (a) we obtain bounds on the expected delay, and (b) our results apply to general network graph models.

## VII. CONCLUDING REMARKS

This paper considers the problem of attaining tight delay guarantees through link scheduling in arbitrary wireless networks. While significant progress has been made in recent years on the topic of throughput-optimal scheduling, the question of characterizing the delay properties of such scheduling policies, or that of obtaining delay optimal schedules, has remained largely open. In this work, we take an important step towards addressing this issue by characterizing the worst-case performance of the widely known maximum weighted scheduling (MWS) algorithm [25], which has been a precursor



to a lot of scheduling algorithms developed in this context. We show that this scheduling policy attains a delay guarantee that depends on the chromatic number of the interference graph of the network, and an appropriate measure of the load on the network. Interestingly, we also show that it is possible to attain the same asymptotic delay guarantee by a throughput-optimal “static” randomized scheduling (RS) policy which precomputes the scheduling probabilities (based on the average arrival rates) and does not take into account the dynamics of the queue length or arrival processes. We observe that the delay guarantee attained by these two scheduling policies is asymptotically tight in the sense that there exist networks (for any given chromatic number) for which these bounds cannot be improved (beyond a constant factor) by any scheduling policy. Our simulation results on a few representative topologies demonstrate that the average packet delay indeed scales linearly with the chromatic number of the link interference graph, and inversely with the arrival slack in the system, in agreement with the scaling laws obtained analytically.

Both MWS and RS require obtaining solutions to NP-hard problems. However, since RS requires us to solve the complex rate decomposition and graph coloring problems only when the network topology changes, or the arrival rates change significantly, its “per-slot” complexity is very low, unlike that of MWS. The reduced complexity of RS comes at increased delay, although the average delays of RS and MWS differ asymptotically by only a constant factor. Development of low-complexity dynamic scheduling algorithms that can match the performance of MWS, and low message-complexity distributed adaptations of RS, remain interesting topics for future research.

#### ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation through grants CNS-0435141, CNS-0448316, NCR-0238340, CNS-0721308 and ECS-0622176.

#### REFERENCES

- [1] A. Brzezinski, G. Zussman, and E. Modiano. Distributed throughput maximization in wireless mesh networks - a partitioning approach. In *Proceedings of ACM MOBICOM*, Los Angeles, CA, September 2006.
- [2] P. Chaporkar, K. Kar, and S. Sarkar. Throughput guarantees through maximal scheduling in multihop wireless networks. In *Proceedings of 43d Annual Allerton Conference on Communication, Control and Computing*, Allerton, Monticello, Illinois, September 28-30 2005.
- [3] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. MIT Press, 2000.
- [4] J. Dai and B. Prabhakar. The throughput of data switches with and without speedup. In *Proceedings of INFOCOM*, pages 556–564, Tel Aviv, Israel, Mar 2000.
- [5] A. Dimakis and J. Walrand. Sufficient conditions for stability of longest queue first scheduling: second order properties using fluid limits. *Advances of applied Probability*, 38(2):505–521, June 2006.
- [6] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput and delay in wireless networks – part i: Fluid case. *IEEE Transaction on Information Theory and IEEE/ACM Transaction on Networking (joint issue)*, 52(6):2568–2592, June 2006.
- [7] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput and delay in wireless networks – part ii: Constant-size packet. *IEEE Transaction on Information Theory*, 52(11):5111–5116, 2006.
- [8] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition, 2001.
- [9] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transaction on Information Theory*, 46(2):388–404, March 2000.
- [10] B. Hajek and G. Sasaki. Link scheduling in polynomial time. *IEEE Transactions on Information Theory*, 34(5):910–917, Sep 1988.
- [11] F. Harary. *Graph Theory*. Addison-Wesley, 1969.
- [12] K. Jung and D. Shah. Low delay scheduling in wireless networks. In *Proc. IEEE ISIT*, France, 2007. Full version available at <http://web.mit.edu/devavrat/www/delay.pdf>.
- [13] X. Lin and S. Rasool. Constant-time distributed scheduling policies for ad hoc wireless networks. In *Proceedings of IEEE CDC-ECC’05*, San Diego, CA, Dec 2006.
- [14] X. Lin, G. Sharma, R. Mazumdar, and N. B. Shroff. Degenerate delay-capacity trade-offs in ad hoc networks with brownian mobility. *IEEE Transaction on Information Theory and IEEE/ACM Transaction on Networking (joint issue)*, 52(6):2777–2784, June 2006.
- [15] X. Lin and N. Shroff. The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks. In *Proceedings of INFOCOM*, Miami, FL, Mar 2005.
- [16] X. Lin and N. B. Shroff. The fundamental capacity-delay tradeoff in large mobile ad hoc networks. In *Proc. Third Annual Mediterranean Ad Hoc Networking Workshop*, Bodrum, Turkey, June 2004.
- [17] J. Mammen and D. Shah. Throughput and delay in random wireless networks with restricted mobility. *IEEE Transaction on Information Theory*, 53(3):1108–1116, March 2007.
- [18] E. Modiano, D. Shah, and G. Zussman. Maximizing throughput in wireless networks via gossiping. In *Proc. ACM SIGMETRICS / IFIP Performance’06*, June 2006.
- [19] M. Neely. Delay analysis for maximal scheduling in wireless networks with bursty traffic. In *Proc. IEEE INFOCOM*, Phoenix, AZ, April 2008.
- [20] M. J. Neely and E. Modiano. Capacity and delay tradeoffs for ad-hoc mobile networks. *IEEE Transaction on Information Theory*, 51(6):1917–1937, June 2005.
- [21] T. Salonidis and L. Tassiulas. Distributed dynamic scheduling for end-to-end rate guarantees in wireless ad hoc networks. In *Proceedings of ACM MOBIHOC*, 2005.
- [22] S. Sarkar and S. Ray. Arbitrary throughput versus complexity tradeoffs in wireless networks using graph partitioning. *IEEE Transaction on Automatic Control*, November 2008 (To appear). Available at <http://www.seas.upenn.edu/~swati/arbitraryclosepoly.ps>.
- [23] D. Shah, P. Giaccone, and B. Prabhakar. An efficient randomized algorithm for input-queued switch scheduling. *IEEE Micro*, 22(1):19–25, Jan-Feb 2002.
- [24] L. Tassiulas. Linear complexity algorithms for maximum throughput in radio networks and input queued switches. In *Proceedings of INFOCOM*, pages 533–539, 1998.
- [25] L. Tassiulas and A. Ephremidis. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1948, Dec 1992.
- [26] X. Wu and R. Srikant. Regulated maximal matching: a distributed scheduling algorithm for multihop wireless networks with node-exclusive spectrum sharing. In *Proceedings of IEEE CDC-ECC’05*, Seville, Spain, Dec 2005.