

# Scheduling in the Queuing System with Asynchronously Varying Service Rates

Shyamsundar R<sup>#1</sup>, Geofrin Shirly. S<sup>#2</sup>

<sup>#1</sup>M.E. Student, Computer Science & Engg.,  
School of Engineering, Vels University, Chennai, India.

<sup>#2</sup>Asst. Prof., Dept. of Computer Science & Engg.,  
School of Engineering, Vels University, Chennai, India.

**Abstract** - In this paper, we target on the problem of scheduling for multichannel wireless networks, for e.g., the downlink of OFDM-based cellular networks for single cell in fourth-generation (4G). Our purpose is to build a practical scheduling policy that will achieve provably good performance in terms of both throughput and delay, at a low complexity. While a class of  $O(n^{2.5} \log n)$ -complexity hybrid scheduling policy is developed to approve both rate-function delay optimality and throughput optimality, and their complexity is typically high. To fix this issue, we build a simple greedy policy called Delay-based Server-Side-Greedy (D-SSG) with a lower complexity of  $2n^2+2n$ , and accurately prove that D-SSG not only achieves throughput optimality, also guarantees near-optimal asymptotic delay performance. Exactly, the rate-function of the delay-violation probability attained by D-SSG for any fixed integer threshold of delay  $b > 0$  is no smaller than the maximum achievable rate-function by any scheduling policy for threshold  $b-1$ . Finally, we simulate to validate our theoretical results in various scenarios. The simulation results show that in all scenarios we consider, D-SSG not only guarantees a near-optimal rate-function, but also anatically has a similar delay performance to the rate-function delay-optimal policies.

## I. INTRODUCTION

In this paper, we target the scheduling problem in a multichannel wireless network, where the system or network has a large bandwidth that can be splitted equally into multiple orthogonal subbands or channels. In such a multichannel system, a key challenge is to design efficient scheduling policy that can simultaneously achieve high throughput and low delay. This will lead to the problem of extremely critical in OFDM systems that are expected to meet the increasing demands from multimedia applications with more stringent quality-of-service (QoS) requirements, and thus look for new ways to achieve higher data rates, lower latencies, and a much better user experience. Even huge challenge is how to design high-performance scheduling policies at a low complexity. On the other hand, there are several orthogonal channels that are need to be allocated to several users. Hence, the scheduling decision has to be made within a short scheduling cycle.

Similarly, another line of work proposed on delay-based scheduling policies and directly focused on the delay performance rather than the queue-length performance. The performance of delay is often harder to identify because the delay in a queueing system does not admit a Markovian representation. The problem becomes even complicated in a multiuser system with fading channels and interference

constraints, where the service rate for individual queues becomes more unpredictable. The scheduling policy called Delay Weighted Matching (DWM), which maximizes the sum of the delay of the scheduled packets in each time-slot, DWM is not only throughput-optimal, but also rate-function delay-optimal, we used the derived rate-function of DWM to develop a simple threshold policy for admission control when the number of users scales nearly with the number of channels in the system. However, DWM meet with a high complexity, which renders it to impractical for modern OFDM systems with many channels and users. We proposed a class of hybrid scheduling policies with a much lower complexity, while still guaranteeing both throughput optimality and rate-function delay optimality. However, the practical complexity of the hybrid policies is still high as the constant factor hidden in the notation is typically large due to the required two-stage scheduling operations and the operation of computing a maximum-weight matching in the first stage. Hence, scheduling policies with an even lower both theoretical and practical complexity are needed in the multiuser multichannel systems.

## II. RELATED WORK

We processed a practical and low-complexity greedy scheduling policy (D-SSG) that not achieves throughput optimality, but also guarantees a near-optimal delay rate-function, for multichannel wireless networks.

### A. Delay Based Server

It been shown that a class of two-stage hybrid policy can achieve both throughput optimality and rate-function delay optimality at a lower complexity compared to of DWM. The hybrid policy is built by combining both certain throughput-optimal policy with a rate-function delay-optimal policy DWM- where is the number of users or channels, which in each time-slot maximizes the sum of the delay of the scheduled packets among the oldest packets in the system.

### B. File Handling

Before we identify the detailed operations of D-SSG, we like to remark on the D-MWS policy in our multichannel system, due to the similarity between D-MWS and D-SSG. Under D-MWS, in each server chooses to serve a queue that has the largest HOL delay among all the server connected to this queues.

### C. Throughput Optimality

We establish the throughput optimality of D-SSG in non asymptotic settings with any fixed value of  $\rho$ . We will examine the delay performance of D-SSG in the asymptotic regime, where  $\rho$  goes to infinity. Hence, even if the convergence rate of the delay rate-function is fast, the throughput performance may still be poor for small to moderate values of  $\rho$ . For a fixed rate-function delay-optimal policy, may not even be throughput-optimal. This ends the focus on studying the throughput performance of D-SSG in general nonasymptotic settings. We remark that the throughput performance of scheduling policy had been extensively studied in various settings, including the multi channel systems. Specifically, for such multichannel systems, we propose a class of Maximum Weight in the Fluid limit (MWF) policies and proved throughput-optimality of the MWF policies in very general settings. The key insight is that to achieve throughput-optimality in such multi channel systems, it is sufficient for each server to choose a connected queue with a large enough weight such that this queue has the largest weight in the fluid limit. Next, we prove that D-SSG is throughput-optimal in general nonasymptotic settings by showing that D-SSG is an MWF policy.

### III. NEAR-OPTIMAL ASYMPTOTIC DELAY PERFORMANCE

In this section, we consider our main result on the near-optimal rate-function. We define the near-optimal rate-function, and then evaluate the delay performance of D-SSG. A policy is said to be achieve near-optimal rate-function if the delay rate-function attained by policy for any fixed integer threshold is no smaller than the threshold for optimal rate-function. The best result of the paper is defined in the following theorem, which states that D-SSG achieves a near-optimal rate-function.

Theorem proved by the following strategy: 1) a key property of D-SSG is motivated, we propose the Greedy Frame-Based Scheduling (G-FBS) policy, which is a different of the FBS policy that has been proved to be rate-function delay-optimal in some cases; 2) show that G-FBS achieves a near-optimal rate-function; 3) prove a dominance property of D-SSG over G-FBS. We monitor that for any given sample path, by the end of each time slot, D-SSG has served every packet that G-FBS has served.

### IV. PERFORMANCE METRICS

We next compare D-SSG to a new policy called Greedy Frame-Based Scheduling. In G-FBS is only for assisting our analysis and will not be used any actual scheduling algorithm. We fix a correctly chosen parameter  $h > 0$ . In the G-FBS policy, packets are grouped into frames satisfying the following requirements.

- 1) No two packets in the same frame have a delay difference larger than time-slots. This guarantees the frame, not more than  $H = Lh$  packets from the same queue can be queued into a single frame.
- 2) Each frame has a capacity of  $n_0 = n - H$  packets, i.e., at most packets can be queued into a frame.
- 3) As packets enter to the time-slot in each system, the

frames are created by queuing the packets respectively. Specifically, packets that arrive first are queued into the frame with a higher priority, and packets from queues with a smaller index are filled with a higher priority when multiple packets arrive in the same system in same time-slot.

Once any one of the above criteria is violated, the current frame will be closed and a new frame will be open. We assume that there is a “leftover” frame, called L-frame, with a capacity  $H\sqrt{n}$  of packets. The L-frame is storing the packets that were not queued in the previous time-slot and were carried over to the next or new frame which open currently in time-slot.

At the starting of each time-slot, we put together the HOL frame and the L-frame into a “super” frame, called S-frame, with a capacity of  $n$  packets. It is easy to follow that in the S-frame, no more than  $2H$  packets are from the same queue/frame. If there are fewer packets in the S-frame, we can manually queue some dummy packets with a delay of zero at the end of the S-frame, so that the S-frame is fully occupied, but still we need to guarantee that not more than  $2H$  packets from the same queue can be queued. In each time-slot, G-FBS runs the D-SSG policy, but restricted to only the packets of the S-frame. We say it a success, if D-SSG can schedule at least  $n_0$  packets, including the oldest  $f(n)$  packets, where  $f(n) < n/2$ , is any function that satisfies that  $f(n) \in o(n/\log^2 n)$  and  $f(n)$ . In each time-slot, if a success does not occur, then no packets will be queued.

#### A. Simulation Result

In this section, we simulate to compare scheduling performance of our proposed D-SSG policy to DWM, hybrid DWM-n-MWS, D-MWS, and Q-SSG. We simulate the policy in Java and compare the empirical probabilities that the largest HOL delay in the system in any given time-slot exceeds an integer threshold  $b$ , i.e.  $P(W(0) > b)$ .

We consider the arrivals that are driven by a two-state in Markov chain and that are correlated over time. For different user, there are five packet-arrivals when the Markov chain is in state 1, and none arrivals when it is in state 2. The Markov chain transition probability is given by the matrix  $\begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix}$ , and the transitions occur at the end of each time-slot. The entries for each user are correlated over time, but they are independent across users. For the channel mode, we first assume i.i.d . ON – OFF channels with unit and set. We later consider the more general scenarios with heterogeneous users and channels that are correlated over time. We run simulations for a system with servers and users, where  $n \in \{10, 20, \dots, 100\}$ . The simulation period for  $10^7$  time-slots for each policy and each system.

We evaluate the scheduling performance of different policies in more realistic scenarios, where users are heterogeneous and channels are correlated over time. Specifically, we consider the channels that can be modeled as a two-state Markov chain, where the channel is “ON” when it is in state 1, and is “OFF” when it is in state 2. We assume that there are two classes of users: users with an odd index are called near-users, and users with an even

index are called far-users. Different classes of users see different channel conditions: Near-users see better channel condition, and far-users see worse channel condition. We consider that the transition matrices of channels probability for far-users and near-users are  $[0.5, 0.5; 0.167, 0.833]$  and  $[0.833, 0.167; 0.5, 0.5]$ , respectively.

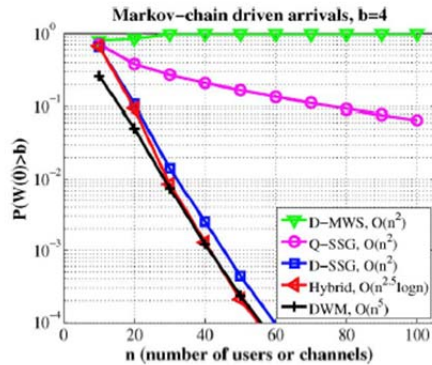


Fig. 1: Comparison performance of different scheduling policies in the case with Markov-chain driven heterogeneous channels, for delay threshold.

We supervise the related results with homogeneous users and i.i.d. Channels in time. In these particular channels, D-SSG exhibits a rate-function that is related to that of DWM and Hybrid, although its delay performance is slightly worse. In this fact, a rate-function delay-optimal policy is not known yet. Hence, for future work, it would be impressive to understand how to create a rate-function delay-optimal or near-optimal policies in general scenarios.

## V. CONCLUSION

As a result, we created a practical and low-complexity greedy scheduling policy (D-SSG) that not only achieves throughput optimality, but also guarantees a near-optimal delay rate-function, for multichannel wireless networks. Our study says that the throughput optimality is relatively easier to achieve in such multichannel networks also, while there exists an explicit tradeoff between complexity and delay performance. If one can deliver a minimal drop in the delay performance, lower-complexity scheduling policies can be advantaged.

## REFERENCES

- [1]. B. Ji, C. Joo, and N. B. Shroff, "Delay-based back-pressure scheduling in multi hop wireless networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1539–1552, Oct. 2013.
- [2]. B. Ji, G. R. Gupta, X. Lin, and N. B. Shroff, "Low-complexity scheduling policies for achieving throughput and asymptotic delay optimality in multi-channel wireless networks," *IEEE/ACM Trans. Netw.*, 2013, accepted for publication.
- [3]. M. Neely, "Delay-based network utility maximization," in *Proc. 29th IEEE INFOCOM*, 2010, pp. 1–9.
- [4]. S. Bodas and T. Javidi, "Scheduling for multi-channel wireless networks: Small delay with polynomial complexity," in *Proc. WiOpt*, 2011, pp. 78–85.
- [5]. S. Bodas, S. Shakkottai, L. Ying, and R. Srikant, "Low-complexity scheduling algorithms for multi-channel downlink wireless networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.