2018 ICSEE International Conference on the Science of Electrical Engineering

Scheduling For 5G Cellular Networks With Priority And Deadline Constraints

Ido Hadar
Faculty of Engineering
Bar-Ilan University
Ramat Gan 52900, Israel

Li-on Raviv
Faculty of Engineering
Bar-Ilan University
Ramat Gan 52900, Israel

Amir Leshem
Faculty of Engineering
Bar-Ilan University
Ramat Gan 52900, Israel

Abstract—**Packet scheduling in 5G networks can significantly affect the performance of beamforming techniques since the allocation of multiple users to the same time-frequency block causes interference between users. A combination of beamforming and scheduling can thus improve the performance of multi-user MIMO systems. Furthermore, in realistic conditions, data packets have both priority and deadlines beyond which they become obsolete. In this paper we propose a simple scheduling algorithm which takes priorities and deadlines into account and allocats users to resource blocks and spatial beams. We demonstrate the merits of the proposed technique compared to other state-of-the-art scheduling methods through simulations.**

Index Terms—**scheduling, 5G networks, cellular communication, resource allocation, EDF**

I. Introduction

In recent years there has been constant growth of media-rich applications for mobile devices. These applications require higher bandwidth from each mobile device. Similarly, the number of devices could reach the tens or even the hundreds of billions when including Internet-of-Things devices that go beyond those serving personal communication [1], [2]. This growth will lead to very increasingly bandwidth requirements in the near future. According to Cisco's Networking Visual Index report [3], data traffic is expected to grow at a compound annual rate of 57% by 2019. Today's 4G mobile network infrastructure cannot support this mushrooming data traffic demand.

The introduction of 5G networks attempts to respond to the growing demand for bandwidth. Andrews et al. [4] defined the three technological approaches that allow 5G networks to meet expectations. 5G networks use extreme densification and offloading. This densification improves the area spectral efficiency, or put differently, more active nodes per unit area and Hz. In addition, 5G increases bandwidth, primarily by moving towards and into the mmWave spectrum but also by making better use of WiFi's unlicensed spectrum in the 5 GHz band. Finally to increase the bandwidth 5G relies on increasing the spectral efficiency, primarily through advances in MIMO, to support more bits/s/Hz per node. All 5G network designs incorporate these three features.

New critical applications such as autonomous cars demand a very low latency for packet transmission. Enforcing low latency means that each packet has a deadline that it needs to meet. In hard real time systems, if a packet fails to be delivered before its deadline expires, it is considered to be lost. The hard real time systems problem has been widely discussed in queuing theory [5]–[8]. The Earliest Deadline First (EDF) scheduling policy is one of the most common methods to schedule packets in a hard real time environment [7]. The EDF is optimal in many queuing models [7], [9]–[13]; however, in the presence of prioritized packets it might be sub-optimal [14].

In the current 4G network environment the latency is not sufficient to support these new applications. The 5G networks aims to cope with this challenge. The problem is handled by the three tiers of the 5G technology: the first tier requires a low-latency core network architecture, the second tier requires a flexible MAC layer, and the last tier handles the congestion control as presented in [15].

Applications differ in terms of their importance. Application priority normally reflects their importance. The priority is attached to the application's packet. Priority becomes a reward upon successful delivery of a packet. The reward is considered to benefit the network if the packet is delivered on time [16]. Scheduling mechanisms that consider both rewards and deadlines are presented in [17]. Another scheduling policy is based on the $c\mu_k/\theta_k$-Rule presented in [18], [19]. This scheduling policy implements a cost function which is similar to a reward summation. The $c\mu_k/\theta_k$-Rule is based on having a finite number of queues each of which presents the packet's priority. The c presents a cost function which is commonly used to indicate the number of packets in the queue. μ_k is the arrival rate of packets and θ_k is the abandon rate of packets that are waiting in the k^{th} queue. In 4G and 5G networks several algorithms have been proposed to handle scheduling these algorithms and aim to support real-time and non-real time traffic [20]–[22]. Similar approaches can be found in WLAN standard 802.11e [23].

Although the scheduling of packets along with rewards and deadlines has been studied [5]–[8], [14], [17]–[19], [24], the problem of jointly scheduling packets and allocating RF resources is still unresolved when the physical layer applies beamforming techniques. In this paper we present three methods to schedule packet transmission and order the usage of resource elements to achieve maximal rewards. The first method is the well known Earliest Deadline First (EDF). The second method is $c\mu_k$-Rule with ordering of the queues in

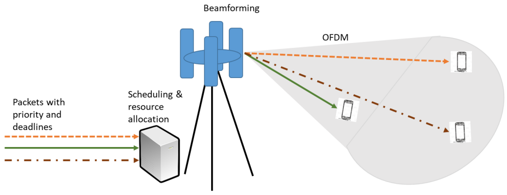
Beamforming
OFDM
Packets with priority and deadlines
Scheduling & resource allocation

Fig. 1. System Architecture

EDF. The third method is Maximal Utility with Dropping [14]. In all cases, the techniques are modified to cope with multiple servers.

The full paper is organized as follows. Section II presents the system model including the arrival process and RF bearer assumptions. Section III describes the proposed system architecture. This description includes both a state-of-the-art algorithm and the new algorithm. The complete article will present simulations testing the proposed system against state-of-the-art algorithms.

II. System Model

We examine a sub-system in the cellular network that includes a Next Generation Node B (gNB), user equipment (UE) and the radio bearer between them. Figure 1 depicts the system model. The model assumes that packets arrive at the gNB from the core network. The packets' arrival discipline is described in sub-section II-A. The radio bearer model is defined in sub-section II-B. The system model assumes that:

- Each packet is addressed to a specific UE.
- The UEs are distributed randomly around the gNB.
- The gNB uses static beam forming. This assumption covers the case of sectorization of the gNB using an antenna array.
- It is assumed that each UE can be served by a single optimal beam.

A. Arrival Process

Packet arrival process is a renewal reward process [16]. Each data packet has an arrival time, a destination or a UE, a packet size, priority and deadline. Let J_i be the packet that arrives at the cell at time t_i. Let A_i be the inter-arrival time of the renewal-reward process $\mathcal{A}$, $t_i = t_{i-1}+A_i = \sum_{j=1}^{i} A_j$. Let B_i be the size of packet J_i. Let D_i be the deadline of J_i. The time is measured from the arrival time to the end of a successful delivery. Let W_i be the priority of packet J_i. Upon successful delivery the priority becomes the reward and let L_i be the identity of the destination UE. It is assumed that the numbers of UEs are finite. Let the tuple $J_i =< a_i, b_i, d_i, w_i, e_i, l_i >$ represent a packet i with its random parameters. Let a_i, b_i, d_i and w_i be realizations of A_i, B_i, D_i and W_i and let $e_i = t_i + d_i$. Let $\hat{S}_t^\pi$ be the set of processed jobs by policy π up to time t.

By definition, the renewal reward process provides a mechanism to analyze the performance of the system. The cumulative rewards function is a simple way to compare the performance of different algorithms.

Definition 1. *The cumulative reward function for time t and policy π is:*

$$U_t^\pi = \sum_{J_i \in S_t^\pi} w_i \qquad (1)$$

Let $t = \hat{t}_n$ and $t' = \hat{t}_{n-1}$ then, the reward difference function is:

$$\Delta U_t^\pi = U_t^\pi - U_{t'}^\pi \qquad (2)$$

The objective is to find a policy π that maximizes the cumulative reward function. If the rewards are deterministic and $W_i = 1$ the cumulative reward function measures the number of jobs that received service.

B. Resource Blocks

In OFDMA networks the transmitted data in the downlink is divided into resource blocks (RB). The resource block is a space time element i.e., part of the spectrum is allocated at a given time interval to a specific set of UEs. The length of the RB in time is called the Transmission Time Interval (TTI). All frequency blocks at a given TTI are called a sub-frame. For example in LTE networks, a TTI has a duration of 1 ms. During this period 14 or 12 resource elements are transmitted depending on whether a normal or extended cyclic prefix is used. A resource element is an OFDM symbol. In the frequency domain, 180Khz sub-channels are allocated to a resource block. The resource block 180Khz spacing is divided into 12 sub-carriers with 15Khz spacing. The resource block period and spacing enable feedback on channel quality and allow the downlink scheduler to optimize the channel utility. The resource block is a data unit dedicated to a specific UE. The resource block scheduler is responsible for allocating resource blocks to connections between the cell and the UEs. The scheduler allocates the resource blocks while trying to maximize the bearer utilization [20], [25], [26].

The cell's antenna system consists of an array of N_a elements which create N_b static beams. In order to simplify the system model, the 12 sub-carriers used in one resource block are assumed to have similar characteristics and are considered in this paper to be a single sub-channel.

- Let u be the number of UEs, k be the number of sub-channels and N_a be the number of antennas in the beam former.
- Let $\mathbf{h}_{u,k} = [\ h_1^{(u,k)} \ \cdots \ h_{N_a}^{(u,k)}\ ]^{\mathrm{T}} \in C^{N_a \times 1}$ be the gain vector.
- Let $\mathbf{w}_{u,k} = [\ h_1^{(u,k)} \ \cdots \ h_{N_a}^{(u,k)}\ ]^{\mathrm{T}} \in C^{N_a \times 1}$ be the steering vector.
- Let $s_{u,k} \in C$ be the information signal and let $P_{u,k} = E[|s_{u,k}|^2]$ be its the transmission power.
- Let $z_{u,k} \in C$ be the additive noise. It is assumed that the additive noise is white Gaussian with variance N_0.

It is assumed that the channel coefficients of different UEs are independent. The signal received by UE u on subcarrier

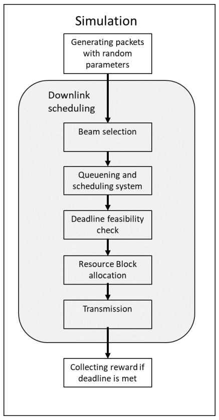
Simulation
Generating packets with random parameters
Downlink scheduling
Beam selection
Queuening and scheduling system
Deadline feasibility check
Resource Block allocation
Transmission
Collecting reward if deadline is met

Fig. 2. System Modules

k, $y_{u,k} \in C$, is given by

$$y_{u,k} = \mathbf{h}_{u,k}^{\mathrm{H}} \mathbf{w}_{u,k} s_{u,k} + \sum_{m \neq u} \mathbf{h}_{u,k}^{\mathrm{H}} \mathbf{w}_{m,k} s_{m,k} + z_{u,k} \quad (3)$$

The SINR of the UE is given by:

$$SINR_{u,k} = \frac{\left| \mathbf{h}_{u,k}^{\mathrm{H}} \mathbf{w}_{u,k} \right|^2 P_u}{\sum\limits_{m \neq u} \left| \mathbf{h}_{u,k}^{\mathrm{H}} \mathbf{w}_{m,k} \right|^2 P_m + N_0} \quad (4)$$

The achievable bit rate is given by:

$$R_{u,k} = \log_2(1 + SINR_{u,k}) \quad (5)$$

The attenuation matrix $\{\mathbf{A}_u\}_{u=1}^{U} \in R^{K \times N_a}$ is computed as follows,

$$\mathbf{A}_u = \begin{bmatrix} \left| \mathbf{h}_{u,1}^{\mathrm{H}} \mathbf{w}_1 \right|^2 & \cdots & \left| \mathbf{h}_{u,1}^{\mathrm{H}} \mathbf{w}_{N_a} \right|^2 \\ \vdots & \ddots & \vdots \\ \left| \mathbf{h}_{u,K}^{\mathrm{H}} \mathbf{w}_1 \right|^2 & \cdots & \left| \mathbf{h}_{u,K}^{\mathrm{H}} \mathbf{w}_{N_a} \right|^2 \end{bmatrix} \quad (6)$$

The beam selection is determined according to the attenuation matrix for each user.

III. Packet Scheduling Architecture

This scheduling architecture is composed of the following elements as depicted in Figure 2:

- Beam selection - The system selects the optimal beam or sector that can serve the UE.
- Queue management - The system uses a queuing scheduling policy that chooses the next packet to transmit.
- Deadline feasibility check - The system decides whether to transmit the packet or not.
- RB allocation - The system allocates RBs to the packet for transmission.
- Packet Transmission

A. Queue Management

This section presents state of the art scheduling policies that are designed to schedule packets with deadlines and rewards. The EDF, $c\mu/\theta$ and $c\mu_k/\theta_k$, EDF version scheduling policies are described here as references for the new MUD policy described in subsection III-A2.

1) State-of-the-art scheduling policies: We define the EDF as follows. Let t be the current time and assume that the queue is not empty and the server is ready to process a job.

Algorithm 1 Earliest Deadline First policy

```
1: $J_i := \underset{J_j \in Q_t^{\pi}}{\text{argmin}}(e_j)$ .
2: If $e_i < t$ then drop $J_i$
3: Else provide service to $J_i$
4: Return to state 1
```

The EDF was shown to be optimal under different metrics [7], [9], [10], [12]. For this reason the EDF has become the standard policy in queuing models with deadlines.

Another policy that can serve as a benchmark for purposes of comparison is the $c\mu/\theta$ scheduling policy [18]. The $c\mu/\theta$ policy assumes that there are Q queues in the system and it needs to select the queue to be served. We assume that there are Q levels of rewards (or Q groups of rewards) each against a queue.

Algorithm 2 $c\mu_k/\theta_k$

```
1: $k_0 := \underset{k=1..K}{\text{argmax}}(\sum c_k \mu_k/\theta_k)$
2: If $(e_i < t)$ then drop $J_i$
3: Else provide service to $J_i$
4: Return to state 1
```

In our model we assume that the service times, deadlines and rewards are known upon arrival. The EDF assumes that there is no information about the service times, whereas $c\mu/\theta$ the deadline is known statistically in terms of the probability of abandonment. This information has advantages in the case of a non-deterministic service time. We present a simple example showing the advantages of using knowledge of the service time upon arrival in the EDF case and present a new version of $c\mu/\theta$ that exploits knowledge of deadline information. In the case of the $c\mu/\theta$ policy, knowing deadlines upon arrival allows us to modify the queue order to use EDF instead of first comes first served as was proposed in [17]. However, as shown below, our proposed technique outperforms this variation as well.

Algorithm 3 $c\mu_k/\theta_k$, EDF version

```
1: $k_0 := \underset{k=1..K}{\text{argmax}}(\sum c_k \mu_k/\theta_k)$
2: $P_i = \underset{J_j \in \hat{Q}_t^{k_0}}{\text{argmin}}(e_j)$
3: If $(e_i < t)$ then drop $J_i$
4: Else provide service to $J_i$
5: Return to state 1
```

2) Maximum Utility with Dropping scheduling policy : We next present the Maximum Utility with Dropping (MUD) scheduling policy as described by the authors of this paper in [14]. The MUD policy combines both scheduling and dropping mechanisms. The scheduling algorithm is based on the following mechanism. Upon arrival, the policy inserts the new job into the queue while keeping the EDF order. If the insertion causes a job to miss its deadline, the policy drops the packet with a minimal throughput ratio ($\frac{w_i}{b_i}$) from the queue. The selected job is the one with the highest throughput ratio as long as it does not cause the second job to miss its deadline. Let $o(J_i,t)$ and $s(J_i,t)$ be functions that characterize the order of J_i in the queue potential at time t.

- $o(J_i,t)$ is the index of job J_i in Q_t^M. $o(J_i,t)=1$ means that J_i is at the head of the queue.
- $s(J_i,t) := \sum\limits_{J_k:o(J_k,t)<o(J_i,t)} b_k$ is the time a job waits before it is processed assuming that no new jobs arrive until it starts processing.

Below we describe how the MUD policy handles a job arrival. Let J_i be the new job which reaches the queue $Q_{t_{i-1}}^M$ at time t_i.

Algorithm 4 MUD

```
1: Wait for the arrival of a new job ($J_i$)
2: If $Q_{t_i}^M = \emptyset$ and the server is idle then process $J_i$ and go
   to statement 1. Add $J_i$ to the queue according to shortest
   time to expiry order. If there are already jobs with the
   same expiration time, order them in the descending order
   of their rewards.
3: Find the first job ($J_k$) which will miss its deadline due to
   the insertion of $J_i$ into the queue $(o(J_k,t_i) \geq o(J_i,t_i))$
   $o_k := \begin{cases} \min\limits_{e_j<s(J_j,t_i)} (o(J_j,t_i)) & \exists j : e_j \leq s(J_j,t_i) \\ \infty & \text{otherwise.} \end{cases}$
4: If $o_k := \infty$ then go to statement 1.
5: Find the job ($J_l$) with the minimum reward per service
   time. If there are several, pick the one with the shortest
   time to expiry. $J_l := \underset{o(J_j,t_i)\leq o_k}{\text{argmin}} (\frac{w_j}{b_j})$
6: Drop job $J_l$ from the queue
7: Go to statement 1
```

Note that $o(J_k,t_i)$ and $s(J_k,t_i)$ values change after adding a new job J_i at time t_i as follows: if $o(J_k,t_i) > o(J_i,t_i)$ then $s(J_k,t_i) := s(J_k,t_{i-1}) + b_i$ and $o(J_k,t_i) = o(J_k,t_{i-1}) + 1$; otherwise there is no change.

B. *Deadline Feasibility Check*

The system estimates whether a packet meets its deadline after transmission. The idea is that dropping packets that miss their deadline can save RF bearer resources. Before transmission starts, the transmission throughput until the end of transmission is unknown. Hence the feasibility check should be against a throughput estimation. These assumptions may be the maximal interference by nearby transmission beams, the highest theoretical throughput, or the average throughput measured in the system. Each approach has advantages and limitations since in the case of a false positive the system drops a packet that could reach the UE on time or in the opposite case the packet consumes resources that are wasted.

C. *Resource Block Allocation*

The system is required to select one of the beams that is used to transmit the packet. In this paper the beams are static and each beam has one or more queues that store the packets that are supposed to be transmitted. The selection mechanism chooses the beam that provides the best channel performance. The metric to define the best beam is the beam that provides the maximal SINR (as defined in equation 4 of the packet's UE destination. The Resource Blocks Allocator (RBA) allocates available channels to the packet. It starts with the channels with the highest bit rate and moves to the lowest bit rate. The process stops when the UE transmission is allocated to N_{RB} sub-channels. When a packet ends its transmission it frees its resource blocks. These resource blocks are allocated to the packet with the shortest deadline.

Algorithm 5 Resource Block Allocation Algorithm

```
For $k = 1..K$
    If $(\mathbf{I}_{k,\mathbf{w}_j} = 0)\&(|\mathcal{K}_{J_i}| < N_{RB})$
        then $\mathbf{I}_{k,\mathbf{w}_j} = u$
End For
```

This module is responsible for the resource allocation process (described in Alg. 5). Matrix $\mathbf{I}$ describes the resource block allocation; i.e., each element $I_{k,\mathbf{w}_j}$ indicates whether the sub-channel and corresponding beam are available ($I_{k,\mathbf{w}_j} = 0$) or occupied by user u ($I_{k,\mathbf{w}_j} = u$). At each iteration, a resource block is allocated for transmitting packet J_i until N_{RB} sub-channels have been allocated for transmission. $\mathcal{K}_{J_i}$ is the number of sub-channels allocated for packet J_i transmission. The validation procedure tests whether the packet's transmission time is suitable for its deadline or not.

D. *Simulation Results*

The simulation is composed of two parts. The first involves choosing randomly finite number of UEs and their channel parameters. The second part is the simulation of the arriving packets. The simulation assumes a cell with a transmission radius $r = 50[m]$ and $N_u = 30$ randomly distributed UEs inside the cell and $N_s = 12$ sub-channels of $BW = 200[kHz]$. The cell is equipped with $N_b = 8$ one dimensional antenna arrays, such that each array consists of $N_a = 32$ antennas. We consider a downlink channel transmission with beamforming using $N_b = 8$ beams for every antenna array, where $\theta_l = \frac{\pi}{12}$ is the departure angle and $\Delta_t = \lambda/2 = c/(2f_c) = 0.15[m]$ is the distance between the elements in the antenna array. To simulate the channel coefficients of each user we use the 3GPP spatial channel model for the Extended Pedestrian A model (EPA). The power spectral density of the AWGN noise is $-170[dBm/Hz]$. For the allocation process we use our proposed Alg. 5. To overcome the change of rate on each allocation step we consider the worst case scenario in terms of SINR. The RBA allocates 4 sub-channels to every packet transmission. In the following simulation we ran 2500 packets with arrival times distributed exponentially and where the packet destination was uniformly distributed over the

TABLE I
PACKET PARAMETERS

Packet Size	Priority	Deadline
64B (40%)	6-10 (70%), 1-4 (30%)	30 x transmission time of a short packet
1522B (20%)	6-10 (20%, 1-4 (80%))	10 x transmission time of a long packet
64B-1522B (40%)	U[1,10]	10 x transmission time of a long packet

finite number of UEs. The other parameters are distributed as described in Table I.

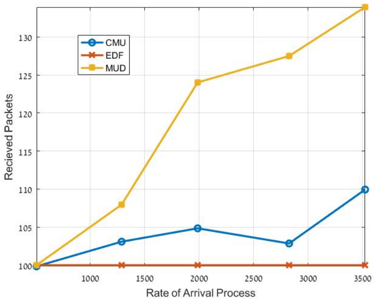

Fig. 3. Number of received packets relative to EDF

We compared the performance of the MUD algorithm against the EDF and $c\mu_k/\theta_k$ algorithms. The number of packets that were received is presented in Figure 3. The y axis in Figure 3 shows the percentage of transmitted packets of the EDF algorithm. The cumulative reward is shown in Figure 4 as the percentage of the EDF algorithm.

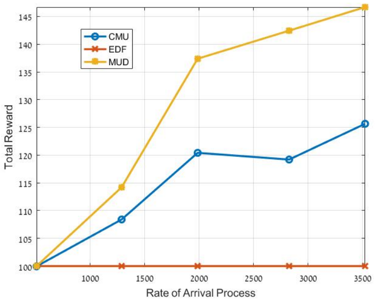

Fig. 4. Collected Rewards relative to EDF

The MUD algorithm thus performs remarkably better, in terms of total reward and received packets, than $c\mu_k/\theta_k$ and EDF algorithms.

ACKNOWLEDGMENT

This paper was partially supported by the Israeli Innovations Authority as part of the HERON 5G Magnet Consortium and ISF-NRF research grant no. 2277/2016.

REFERENCES

[1] M. S. Corson, R. Laroia, J. Li, V. Park, T. Richardson, and G. Tsirtsis, "Toward proximity-aware internetworking," *IEEE Wireless Communications*, vol. 17, no. 6, 2010.

[2] A. Maeder, P. Rost, and D. Staehle, "The challenge of m2m communications for the cellular radio access network," in *Proc. Würzburg Workshop IP, Joint ITG Euro-NF Workshop "Vis. Future Gener. Netw." EuroView*, 2011, pp. 1–2.

[3] V. Cisco, "Cisco visual networking index: Forecast and methodology 2014–2019 white paper," *Cisco, Tech. Rep*, 2015.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[5] D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data networks*. Prentice-Hall International New Jersey, 1992, vol. 2.

[6] J. A. Stankovic, M. Spuri, M. Di Natale, and G. C. Buttazzo, "Implications of classical scheduling results for real-time systems," *Computer*, vol. 28, no. 6, pp. 16–25, 1995.

[7] J. A. Stankovic, M. Spuri, K. Ramamritham, and G. C. Buttazzo, *Deadline scheduling for real-time systems: EDF and related algorithms*. Springer Science & Business Media, 2012, vol. 460.

[8] R. Srikant and L. Ying, *Communication networks: an optimization, control, and stochastic networks perspective*. Cambridge University Press, 2013.

[9] S. S. Panwar, D. Towsley, and J. K. Wolf, "Optimal scheduling policies for a class of queues with customer deadlines to the beginning of service," *Journal of the ACM (JACM)*, vol. 35, no. 4, pp. 832–844, 1988.

[10] P. P. Bhattacharya and A. Ephremides, "Optimal scheduling with strict deadlines," *IEEE Transactions on Automatic Control*, vol. 34, no. 7, pp. 721–728, 1989.

[11] D. Towsley and S. Panwar, *Optimality of the stochastic earliest deadline policy for the G/M/c queue serving customers with deadlines*. Citeseer, 1991.

[12] R. C. L. Katzir, "Scheduling of voice packets in a low-bandwidth shared medium access network," *EEE/ACM Transactions on Networking (TON)*, vol. 15, no. 4, pp. 932–943, 2007.

[13] P. Brucker and P. Brucker, *Scheduling algorithms*. Springer, 2007, vol. 3.

[14] L. Raviv and A. Leshem, "Maximizing Service Reward for Queues with Deadlines," 2018, arXiv:1805.11681v1.

[15] O. N. Yilmaz, Y.-P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5g communication for a factory automation use case," in *Communication Workshop (ICCW), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1190–1195.

[16] G. Grimmett and D. Stirzaker, *Probability and random processes*. Oxford university press, 2001.

[17] J. M. Peha and F. A. Tobagi, "Cost-based scheduling and dropping algorithms to support integrated services," *IEEE Transactions on Communications*, vol. 44, no. 2, pp. 192–202, 1996.

[18] R. Atar, C. Giat, and N. Shimkin, "The cμ/θ rule for many-server queues with abandonment," *Operations Research*, vol. 58, no. 5, pp. 1427–1439, 2010.

[19] Z. Yu, Y. Xu, and L. Tong, "Deadline scheduling as restless bandits," in *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*. IEEE, 2016, pp. 733–737.

[20] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: Key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[21] H.-S. Liao, P.-Y. Chen, and W.-T. Chen, "An efficient downlink radio resource allocation with carrier aggregation in lte-advanced networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2229–2239, 2014.

[22] H. S. B. Abdelmula, M. M. Warip, O. B. Lynn, and N. Yaakob, "An efficient scheduling scheme for heterogeneous services in ofdma based 5g lte-advanced network with carrier aggregation," in *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. IEEE, 2018.

[23] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, and L. Stibor, "Ieee 802.11 e wireless lan for quality of service," in *Proc. European Wireless*, vol. 2, 2002, pp. 32–39.

[24] U. Ayesta, P. Jacko, and V. Novak, "A nearly-optimal index rule for scheduling of users with abandonment," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 2849–2857.

[25] D. Astély, E. Dahlman, A. Furuskär, Y. Jading, M. Lindström, and S. Parkvall, "Lte: the evolution of mobile broadband," *IEEE Communications magazine*, vol. 47, no. 4, 2009.

[26] H. Zarrinkoub, *Understanding LTE with MATLAB: from mathematical modeling to simulation and prototyping*. John Wiley & Sons, 2014.