# A Downlink Resource Scheduling Strategy for URLLC Traffic

Shashi Raj Pandey
Department of Computer Science and Engineering
Kyung Hee University
Republic of Korea
Email: shashiraj@khu.ac.kr

Madyan Alsenwi
Department of Computer Science and Engineering
Kyung Hee University
Republic of Korea
Email: malsenwi@khu.ac.kr

Yan Kyaw Tun
Department of Computer Science and Engineering
Kyung Hee University
Republic of Korea
Email: ykyawtun7@khu.ac.kr

Choong Seon Hong
Department of Computer Science and Engineering
Kyung Hee University
Republic of Korea
Email: cshong@khu.ac.kr

***Abstract*—In this paper, we propose a dynamic resource scheduling of URLLC (Ultra-reliable low latency communication) and eMBB (Enhanced mobile broadband) services for the downlink in 5G networks. The eMBB traffic is characterized with the high bandwidth network services such as web browsing, video streaming, augmented realiy, and the URLLC services require sub-millisecond latency with low error rates (e.g., autonomous driving and remote surgery). Therefore, we aim to formalize a resource scheduling over a time-slot having multiplexed eMBB and URLLC traffic with the objective to protect eMBB users while maintaining URLLC stringent latency requirements. For this purpose, to guarantee minimum latency, we allow puncturing technique for URLLC traffic over the scheduled resources, and study the sequential scheduling scenario for the preempted eMBB users in the next time-slot. The simulations result show the efficacy of proposed dynamic scheduling method, and gain in priority based scheduling.**

***Index Terms*—5G New Radio (NR), resource scheduling, eMBB, URLLC, Gittins Index.**

## I. Introduction

Radio resources are considered an expensive utility whose management has been of a great challenge for the evolution of mobile networks. On the top of that, traffic with strict latency and reliability requirements in upcoming 5G networks requires an efficient resource allocation scheme to make it a success [1]. However, with existing network scenario, there appears a challenge to simultaneously achieve both of such requirements [2], [3]. This is because of the current wireless system architecture where the primary focus is on maximizing the throughput with long packets. On the other hand, the use of short packets dramatically reduces data rate [4], though we can achieve ultra-reliability . These circumstances are exacerbated with diversified services and application for the 5G networks. Specifically, the new features in 5G New Radio (NR) are architectured to support three service catagorizes of traffic with distinguished specifications in terms of applications and requirements i.e., enhanced Mobile Broad Band (eMBB), massive Machine Type Communications (mMTC) and Ultra Reliable Low Latency Communications (URLLC).

eMBB is an internet access service that is suitable for high bandwidth applications such as web browsing and video streaming. It is an extension of Long Term Evolution-Advanced (LTE-A) [5], which we are used to with, primarily focused for maximizing data rate. mMTC is defined with narrowband internet access, typically by IoT devices relevant to sensing, metering and monitoring services. Therefore, the traffic is sporadic in nature for a certain period of active time. URLLC services are categorized for latency sensitive devices with applications like industrial automation, autonomous driving, and remote surgery that requires utmost measure of reliability. As an example, the current 3GPP requirements define reliability metric as $(1-10^{-5})$ success probability while transmitting a PDU of 32 bytes within 1ms, and the user plane latency of 1ms for URLLC traffic assuming for a single user. This hard constraint defines the quality of service (QoS) of URLLC.

To meet this stringent latency and reliability requirements for URLLC, new techniques need to be devised accordingly. Also, the massive growth of IoT networks has brought up numerous challenges while handling such heterogeneous traffic in an effective way [6]. In this paper, we discuss about the coexistence of the services of URLLC and eMBB deployed in the same radio spectrum, associated with 5G NodeB (gNB). In such case, when URLLC packets arrive at the gNB while some ongoing eMBB transmission, due to the hard latency requirements for URLLC, the gNB might suspend the current transmission to free up the radio resources. To maintain the QoS of URLLC, [7] indicates the immediate forwarding of such latency sensitive packets. This technique is referred as the puncturing mechanism [8] where the preempted eMBB users due to puncturing would be rescheduled. The scenario when the eMBB and URLLC services need to be dynamically multiplexed over the network bandwidth for improving the spectral efficiency, the eMBB users need to be protected while meeting the tight URLLC latency requirements i.e., the impact of immediate URLLC packet transmission due to puncturing on the ongoing services need to be addressed. In [9], authors discussed about various models as the implications of eMBB rate loss such as *linear rate loss model* associated with the URLLC puncturing/superposition, and characterize the iterative gradient scheduler for eMBB.

Recent works, as discussed and others related studies in this regards are more focused on system level design. In [10], [11], [12] authors have discussed about the lower layer design entities such as packet sizes, overheads, and control channel structure to ensure lower latency, with improved reliability metrics for the envisioned URLLC services. This measure, however, cannot properly address the dynamic multiplexing situation of the eMBB and URLLC services which is critical to ensure the corresponding QoS requirements. Similarly, under the common domain of system level design metrics, in [10], [13] authors have argued about the physical aspects of URLLC that includes coding and modulation. In this work, we consider a dynamic multiplexing scenario between eMBB and URLLC, and formalize resource scheduling over a time-slot to maintain URLLC specifications while protecting the eMBB users rate. We formulated an optimization problem with the objective to maximize the rate of the eMBB users while consolidating URLLC requirements, and solve its deterministic form with transformations. Furthermore, a priority selection strategy is studied for resource allocation to the preempted eMBB users.

The remainder of this paper is organized as follows. We present the details the system model in Section II. In Section III, we formulate the resource scheduling problem as an optimization problem, and formalize the methodology of priority selection for preempted eMBB users. In Section IV, the evaluation is presented using the proposed mechanism as mentioned in section III. Finally, conclusions are drawn in Section V.

## II. System Model

We follow the radio frame structure specified in 3GPP specification [14], and adapt equally spaced time slots further divided into multi-minislots based upon the configuration of the sub-carriers. Let us consider a set of eMBB users $\mathcal{K} = \{1, 2, \ldots, K\}$ associated with a 5G NodeB (gNB) within a time-slot $T$. Correspondingly, we consider a set of terminals $\mathcal{U} \subseteq \mathcal{K}$ requesting sporadic URLLC traffic multiplexed in the frequency-time slot. If $B$ denotes the effective bandwidth in each slot of 1ms, then we can formalize resource associated with each minislot ($t_m$) as a fraction of $f(t_m) = B/\eta$, for an integer value $\eta > 0$. We will further quantize $f(t_m)$ into orthogonal resource blocks (RBs) of $N$ levels denoted by a set $\mathcal{N} = \{1, 2, \ldots, N\}$. Here, for each associated user $k$, we will define its state at time $t_m$ with the fraction of resource requested by $x_k(t_m) = f(t_m)/N$, where $x_k(t_m) \in \mathcal{N}$.

Fig.(1) summarizes the frame structure with multiplexed eMBB and URLLC traffic. From the illustration, we see that

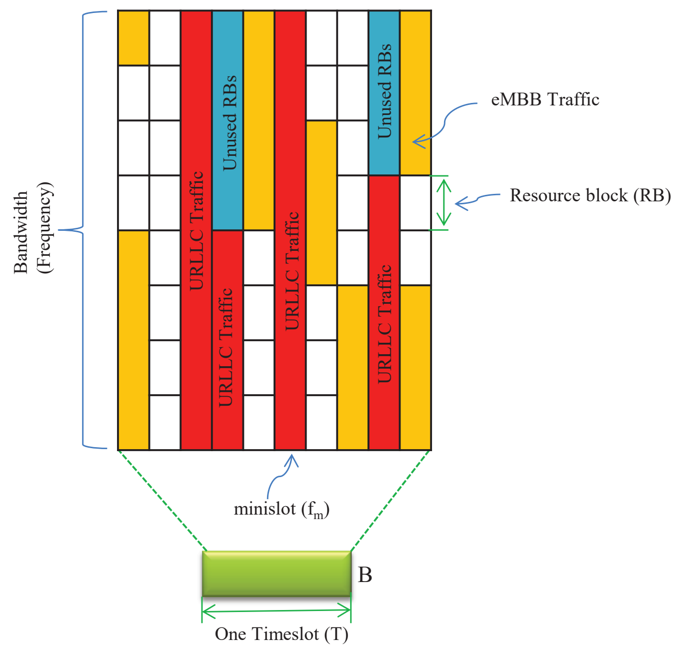


Fig. 1. Illustration of multiplexed eMBB and URLLC traffic: Time is divided into minislots, and resource in each minislot is further divided into Resource Blocks (RBs). URLLC traffic can puncture (or overlap) at any minislot, while eMBB is scheduled at the beginning of slots.

the immediate scheduling of URLLC traffic will create a puncturing effect in the minislot, where already scheduled eMBB users are dropped out of service for the next time slot $T$. Under such circumstances, the order of scheduling such users can be done as per its state information i.e., the level of punctured resources so as to maximize the network utility in terms of resource utilization over the time. If user $k$ in state $x_k(t_m)$ is scheduled for a time slot $T$, the reward value obtained by the gNB in terms of resource utilization is defined as $r_k(x_k(t))$. This way, for a time-slot $T$ a sequential user selection scenario exists for the gNB to schedule the resources to the ordered $\mathcal{Z} = \{1, 2, \ldots, Z\} \subseteq \mathcal{K}$ eMBB users defined by the set $\mathcal{U}_T^z := \{k | x_k(t) \in \mathcal{Z}\}$ affected by the puncturing for URLLC packets. Because the reward distribution is unknown, the gNB can allocate the resource following the solution approach for the multi-armed bandit problem, to maximize its cumulative reward over the time. For this purpose, we can prioritize a node $k$, given its state $x_k(t_m)$ and corresponding reward $r_k(x_k(t))$, and then the gNB can sequentially resolve the resource scheduling strategy in an efficient way. Also, we capture the stochastic nature of downlink traffic, and associate the state transition representation of the user with a transitional probability matrix that is to be updated over the time. In the following section, we will formulate our problem for this scenario.

## III. Problem Formulation

In this section, we formulate the problem of resource scheduling and allocation in two phases: First we formulate

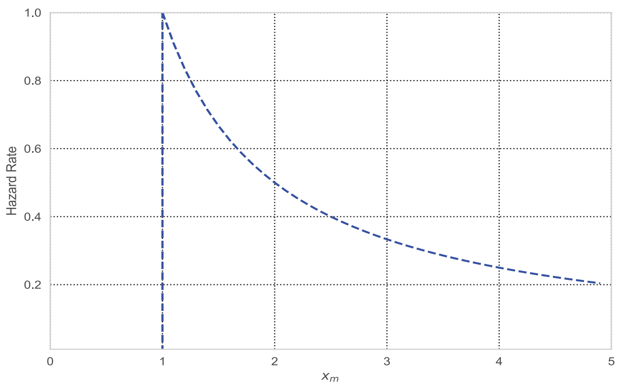


Fig. 2. Function h(x) for the Pareto distribution with parameters $x_m = 1, \alpha = 1$.

the resource scheduling problem between the eMBB users and URLLC traffic as an optimization problem where we maximize the throughput of the eMBB users while maintaing the reliability requirements of the URLLC traffic. Secondly, we formalize the scheduling of punctured eMBB users resource (time-frequency slices) with a bandit problem, and present it as an efficient scheduling strategy to maximize the network utilization over the time.

### A. Resource Scheduling Mechanism

For an eMBB user $k$, the instantaneous rate is characterized as,

$$R_k^e = \eta_k f(t_m) \log_2 \left(1 + \frac{\rho_k |G_k|^2}{\mathcal{N}_k}\right), \forall k \in \mathcal{K}, \quad (1)$$

where $\eta_k f(t_m)$ is the total spectrum allocated to the user $k$, $\rho_k$ is the transmission power of the user $k$, $|G_k|^2$ is the channel gain between user $k$ and the base station, and $\mathcal{N}_k$ is the noise power.

We can extend problem (1), and define the rate of all eMBB users in time slot $T$ as,

$$R_T^e = \sum\nolimits_{k \in \mathcal{K}} \eta_k f(t_m) \log_2 \left(1 + \frac{\rho_k |G_k|^2}{\mathcal{N}_k}\right), \forall k \in \mathcal{K}, \quad (2)$$

such that $\sum_{k \in \mathcal{K}} \eta_k f(t_m) \leq B$. Correspondingly, the expected instantaneous rate $R_T^u, u \in \mathcal{U}$ of URLLC load is defined as,

$$R_T^u = \sum\nolimits_{u \in \mathcal{U}} \eta_u f(t_m) \log_2 \left(1 + \frac{\rho_u |G_u|^2}{\mathcal{N}_u}\right), \forall u \in \mathcal{U}, \quad (3)$$

such that $\sum_{k \in \mathcal{K}; u \in \mathcal{U}} f(t_m)(\eta_k + \eta_u) \leq B$. We characterize URLLC traffic load multiplexed with eMBB users within time slot $T$ as a pareto-distributed random variable $X$ with parameters $\alpha$ and $x_m$ such that its cumulative distribution function (CDF) is,

$$F_X(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^\alpha, & x \geq x_m; \\ 0, & x < x_m. \end{cases}$$

For the values of $\alpha$ and $x_m$, we can analyze the URLLC traffic load defined by Pareto distribution with the metrics of hazard rate, as the function of ratio $f_X(x)$ to $1 - F_X(x)$,

$$h(x) = \begin{cases} 0, & 0 \leq x < x_m; \\ \frac{\alpha}{x}, & x \geq x_m. \end{cases} \quad (4)$$

Fig. (2) illustrates the right-continuous (and decreasing) hazard rate for the Pareto distribution, with the corresponding parameters $x_m = 1$ and $\alpha = 1$. The significance of changing the parametric values upon the resource scheduling problem will be presented in the simulation results later.

Now we can denote the reliability measure of URLLC traffic in terms of the maximum outage probability such as ,

$$P(R_T^u < X) \leq \epsilon, \quad (5)$$

where $\epsilon$ is a small value that captures the confidence level. Using the CDF structure of $X$, $F_X$ we recast the (5) as,

$$\begin{aligned} P(R_T^u < X) \leq \epsilon &\iff 1 - F_X(R_T^u) \leq \epsilon, \\ &\iff F_X(R_T^u) \geq (1 - \epsilon), \\ &\iff R_T^u \geq F_X^{-1}(1 - \epsilon), \end{aligned} \quad (6)$$

where $F_X^{-1}(.)$ is the inverse CDF of random variable $X$ that can be evaluated for the confidence level defined by the value of $\epsilon$. This way plugging the value of $F_X^{-1}(.)$ in (6), we transform (5) into a deterministic form i.e.,

$$R_T^u \geq F_X^{-1}(1 - \epsilon) \iff (R_T^u)^\alpha \geq \frac{(x_m)^\alpha}{\epsilon}. \quad (7)$$

With the aforementioned definitions we can formalize the optimization problem to maximize the rate of eMBB users while consolidating URLLC requirements. For this, we assume that the impact of multiplexed eMBB traffic to punctured resources by URLLC traffic is proportional [9]. Therefore, the optimization problem is defined as,

$$\begin{aligned} \underset{\eta_z}{\text{Max}} \quad & \sum\nolimits_{k \in \mathcal{K}} (\eta_k - \eta_u) f(t_m) \log_2 \left(1 + \frac{\rho_k |G_k|^2}{\mathcal{N}_k}\right), \forall k \in \mathcal{K} \\ \text{s.t} \quad & C_1 : (R_T^u)^\alpha \geq \frac{(x_m)^\alpha}{\epsilon}, \\ & C_2 : \sum\nolimits_{u \in \mathcal{U}} \eta_u f(t_m) \leq B. \end{aligned} \quad (8)$$

Here, constraint $C_1$ characterizes the reliability guarantee with confidence bound value $\epsilon$, and pareto-distribution parameters $\alpha$ and $x_m$. Constraint $C_2$ is the resource budget in one time-slot $T$. For $\alpha = 1$, we can obtain a convex structure of the problem (8) which is easy to solve for optimal URLLC resource level $\eta_u^*$.

### B. Resource Allocation to eMBB Users

As we have discussed about the result of scheduling for the URLLC traffic, there exists a sequential scheduling scenario for the punctured eMBB resources in the next time slot $T$

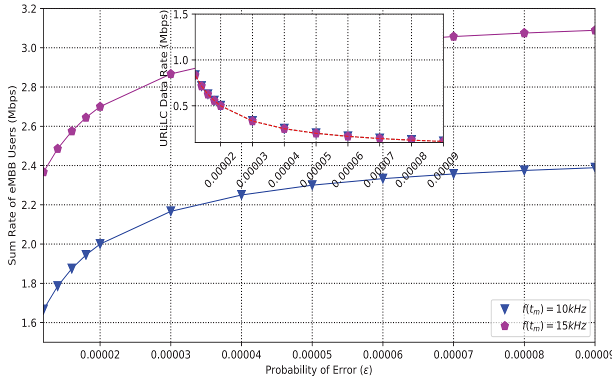


Fig. 3. Sum Rate of eMBB users for different reliability levels of URLLC traffic load; B = 3 MHz, $x_m = 1$.

**Algorithm 1** Priority Selection Strategy for Resource Allocation to eMBB Users

**Input:** $\mathcal{B} = \mathcal{F}, \mathbf{r}, a$.
**Output:** $\mathcal{U}_T^z$.
$\mathcal{U}_T^z = \{\}, \mathcal{I} = \{\}, k = K$;
**while** $k > 0$ **do**
    Calculate the index value, $\nu(F_k)$
    $\mathcal{I} = \mathcal{I} \cup \nu(F_k)$;
    $\mathcal{B} = \mathcal{B} \backslash k$;
    $k = k - 1$;
**end while**
Return $\mathcal{U}_T^z \leftarrow \mathcal{U}_T^z \cup \{\mathcal{I}_{max}^Z\}$ ▷ top $Z$ users

amongst the possible users, which appears to be a $k$-arm bandit and a single player (gNB) scenario, where at each time $T$, the player (gNB) chooses one arm (eMBB user) to play (allocate its scheduled resource) with priority. In our formulation the decision of resource allocation is performed before the initiation of a time slot, so we can extend this process with the sequence of time $T$ and the states of users, $x_k(t_m), \forall k$. We consider a family of bandit process $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ as in [15]. Here, the bandit process $F_k, \forall k$ is defined with the state $x_k(t_m)$, and reward at the state $r_k(x_k(t_m)) > 0$, and is considered to be an exponentially discounted semi-Markov decision process. For the ease of presentation, we assume the a constant duration between the decision times (e.g., say 1).

We adapt the control set $\xi = \{0, 1\}$, where the control 0 freezes the process. That means, there is no change in state and no reward is obtained from the process. Similarly, control 1 is defined as the continuation control that returns an immediate reward $a^t r_k(x_k(t_m)) = e^{-\gamma t_m} r_k(x_k(t_m))$. Here, the parameters $a (0 < a < 1)$ and $\gamma (\gamma > 0)$ are defined as the discount factor and the discount parameter respectively for obtaining a bounded reward value. This way, the presented problem resembles with a discounted-reward Markov decision process whose solution can be found using dynamic programming equations [16]. However, the solution becomes difficult to solve for a $k-$bandit process where the problem grows exponentially. Therefore, using these definitions, we refer [17] which states that for a discrete time Markov decision process, there exists an optimal policy defined as index policy, which is characterized by a real-valued index, $\nu(F_k, x_k(t_m))$, and it is to continue the bandit process having greatest index. With this definition, for a given stopping time $\tau$ we have,

$$\nu(F_k) = \max_{\tau > 0} \nu_\tau(F_k), \tag{9}$$

where, $\tau$ defines the decision time to stop applying the control $u = \{1\}$.

**Lemma 1:** For the expected total discounted reward over $\tau$ steps: $R_\tau(F_k) = E \sum_{t=0}^{\tau-1} a^t r_k(x_k(t_m))$, and the expected total discounted time over $\tau$ steps: $W_\tau(F_k) = \sum_{t=0}^{\tau-1} a_{=}^{t}(1 - a)^{-1} E(1 - a^\tau)$, we can formalize the Gittins Index value using the definition in (9) as,

$$\begin{aligned} \nu(F_k) &= \max_{\tau > 0} \nu_\tau(F_k), \\ &= \frac{R_\tau(F_k)}{W_\tau(F_k)}. \end{aligned} \tag{10}$$

*Proof:*

Let us analyze the formulation in (10) with a simple case scenario of two bandit processes $F_1$ and $F_2$. For a stochastically independent bandit processes $F_1$ and $F_2$, having indices as $\nu(F_1)$ and $\nu(F_2)$, with $\nu(F_1) > \nu(F_2)$, and the stopping time $\tau$ for the process $F_1$, and $\varphi$ as an arbitrary stopping time for $F_2$, we can derive the following relation as,

$$\begin{aligned} \nu(F_1) > \nu(F_2) \Leftrightarrow R_\tau(F_1) + E a^\tau R_\varphi(F_2) \\ > R_\varphi(F_2) + E a^\varphi R_\tau(F_1). \end{aligned} \tag{11}$$

From (11), we can analyze that the return of reward with the choice of control $\xi$ applied for the bandit process will be improved by selecting continuation control on the bandit with the greatest index. Therefore, for the user scheduling problem at time $t$, we can evaluate the index values at the bandit processes given state $x_k(t), \forall k$. Then after, we can choose to apply the continuation control to the bandits with the ordered $Z$ index value, as (11) guarantees for better discounted cumulative reward. The detail implementation for

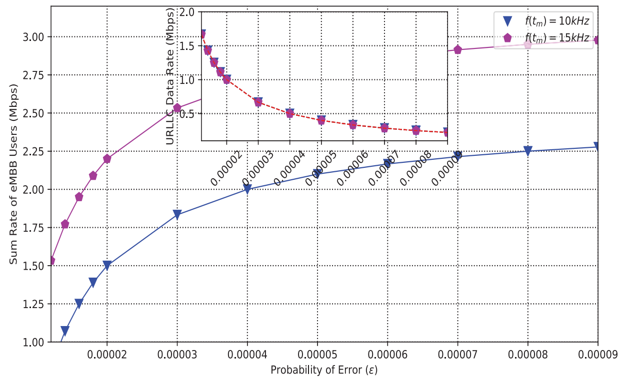


Fig. 4. Sum Rate of eMBB users for different reliability levels of URLLC traffic load; B = 3 MHz, $x_m > 1$.

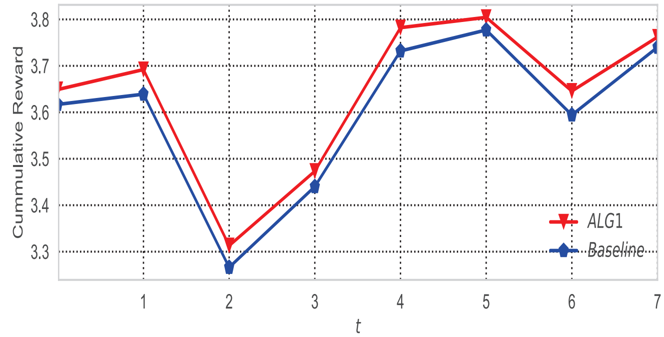


Fig. 5. Comparative analysis on cumulative reward upon priority based scheduling and baseline.

this scenario is presented in Algorithm.1.

## IV. Simulation Results

For the simulation scenario, we quantize the fraction of reserve resource as $|\mathcal{N}| = 12$ levels defined as resource block, with subcarrier spacing 15 kHz. We select system bandwidth to be 3 MHz, with two configuration for $f_m$ for comparision. For system variables, we consider a consistent network environment with parameters defined in [18]. Similarly, the URLLC traffic is generated as a pareto distribution for the value of $\alpha = 1$. To study the scheduling of the punctured eMBB users, each level is defined as the state of the bandit process. We then consider users associated with the BS, and define the transition probability matrix for the family of bandit processes. In each round, the reward $r$ was randomly generated for the states. We use the values of discount factor $a = 0.9$. Fig. 3 shows the sum data rate of all eMBB users over the different values of reliability metric defined as the error probability $\epsilon$ for the URLLC traffic. When $f_m$ = 10kHz, for higher error probability $\epsilon$ upon URLLC traffic, more resources are scheduled to the eMBB users, thus increasing the sum rate. Also, when the value of $f_m$ is increased, large mini-slots are available to schedule for eMBB users. Correspondingly, we can observe the decrease in rate of URLLC for larger values of $\epsilon$. Here, because of sufficient network resource in our configuration, both comparative setting shows similar trend for URLLC rate.

Fig. 4 reflects the impact of parameter $x_m > 1$ that characterizes the pareto distribution of URLLC traffic. The expected system response is highlighted with the hazard rate in Section III-A, Fig. 2. With the increase in the value $x_m$, we observe more strict constraint imposed upon the optimization problem, as sufficiently reflected with the URLLC rates. Because of this, we see lower sum rate of eMBB users for both configurations of $f_m$.

For the study of schdeuling of dropped eMBB users due to puncturing, in Fig. 5 we analyze the priority based scheduling of four eMBB users associated with a transition probability that defines its state change over time-slot. Considering the random puncturing of the users, we compare the proposed methodology Alg.1 with the Baseline, a random schdeuling approach. We see there exists a gain in cumulative reward when we prioritize to schedule the punctured eMBB users; as observed in the results.

## V. Conclusions

In this work, we have studied a dynamic multiplexing scenario between the eMBB and the URLLC, and characterize the solution of the downlink resource scheduling problem as an optimization problem where the objective is to maximize the overall sum rate of eMBB users while we adhere to the stringent latency requirements of the URLLC. We simulated and compared the results to show the efficacy of the proposed mechanism that highlights the impact of QoS constraints upon overall rate of the network. We further studied the priority selection strategy for scheduling the dropped eMBB transmission while fulfilling URLLC requirements. The results show some improvement as the metric of cummulative reward. As a future extension, we intend to consider both uplink and downlink scenario while scheduling for URLLC in 5G NR.

## Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2015-0-00557,Resilient/Fault-Tolerant Autonomic Networking Based on Physicality, Relationship and Service Semantic of IoT Devices) and supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2016R1D1A1B01015320). *Dr. CS Hong is the corresponding author.

## References

[1] M. Bennis, M. Debbah, and H. V. Poor, "Ultra-reliable and low-latency wireless communication: Tail, risk and scale," *arXiv preprint arXiv:1801.01270*, 2018.

[2] B. Soret, P. Mogensen, K. I. Pedersen, and M. C. Aguayo-Torres, “Fundamental tradeoffs among reliability, latency and throughput in cellular networks,” in *Globecom Workshops (GC Wkshps), 2014*. IEEE, 2014, pp. 1391–1396.

[3] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, “Ultra-reliable and low-latency communications in 5g downlink: Physical layer aspects,” *IEEE Wireless Communications*, vol. 25, no. 3, 2018.

[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[5] C. Hoymann, D. Astely, M. Stattin, G. Wikstrom, J.-F. Cheng, A. Hoglund, M. Frenne, R. Blasco, J. Huschke, and F. Gunnarsson, “Lte release 14 outlook,” *IEEE Communications Magazine*, vol. 54, no. 6, pp. 44–49, 2016.

[6] S. Chen and J. Zhao, “The requirements, challenges, and technologies for 5g of terrestrial mobile telecommunication,” *IEEE communications magazine*, vol. 52, no. 5, pp. 36–43, 2014.

[7] Tech. Rep., study on New Radio Access Technology Physical Layer Aspects (Release 14).

[8] K. I. Pedersen, G. Pocovi, J. Steiner, and S. R. Khosravirad, “Punctured scheduling for critical low latency data on a shared channel with mobile broadband,” in *Vehicular Technology Conference (VTC-Fall), 2017 IEEE 86th*. IEEE, 2017, pp. 1–6.

[9] A. Anand, G. de Veciana, and S. Shakkottai, “Joint scheduling of urllc and embb traffic in 5g wireless networks,” *arXiv preprint arXiv:1712.05344*, 2017.

[10] G. Durisi, T. Koch, and P. Popovski, “Toward massive, ultrareliable, and low-latency wireless communication with short packets,” *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.

[11] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, “A flexible 5g frame structure design for frequency-division duplex cases,” *IEEE Communications Magazine*, vol. 54, no. 3, pp. 53–59, 2016.

[12] C.-P. Li, J. Jiang, W. Chen, T. Ji, and J. Smee, “5g ultra-reliable and low-latency systems design,” in *Networks and Communications (EuCNC), 2017 European Conference on*. IEEE, 2017, pp. 1–5.

[13] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, “Short-packet communications over multiple-antenna rayleigh-fading channels,” *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 618–629, 2016.

[14] [Online]. Available: https://www.3gpp.org/ftp/Specs/archive

[15] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.

[16] R. Bellman, *Dynamic programming*. Courier Corporation, 2013.

[17] R. Weber *et al.*, “On the gittins index for multiarmed bandits,” *The Annals of Applied Probability*, vol. 2, no. 4, pp. 1024–1033, 1992.

[18] [Online]. Available: https://sites.google.com/site/lteencyclopedia/lte-radio-link-budgeting-and-rf-planning/lte-link-budget-comparison