Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 5 (2016), pp. 4243–4251

http://www.ripublication.com/gjpam.htm

Queueing Delay Analysis of CDF-based Scheduling with G/M/1 Queue Approximation

Yoora Kim

Department of Mathematics,
University of Ulsan, Ulsan 44610, Korea.

Abstract

In this paper, we consider a cumulative distribution function (CDF)-based opportunistic scheduling for downlink transmission in a cellular network consisting of a base station and multiple mobile stations. Using G/M/1 queue approximation, we present an explicit formula for the average queueing delay of each mobile station, which is applicable to a general arrival process under the assumption that wireless channel varies independently over time-slots. Through numerical studies and simulation, we verify the tightness of our approximation and investigate the delay performance of the CDF-based scheduling for various system parameters.

AMS subject classification:
Keywords: CDF-based scheduling, Multiuser diversity, Queue, Delay, Opportunistic scheduling, G/M/1 queue.

1. Introduction

In wireless networks, each user experiences different channel conditions at different times due to fading of the wireless medium. Opportunistic scheduling can maximize the system throughput both in uplink and downlink by selecting the user who experiences the best channel condition at each time-slot [1]. Such throughput gain is usually achieved with the cost of fairness, since users having higher signal-to-noise ratio (SNR) on average tend to be scheduled more frequently.

A cumulative distribution function (CDF)-based opportunistic scheduling was proposed in [2] to resolve the fairness problem in opportunistic scheduling. It is shown that, by utilizing a set of weight parameters, the CDF-based opportunistic scheduling can provide *precise* control of fairness among users while achieving high system throughput [3].

With these useful features, recent studies have extended the CDF-based scheduling to various networks including multicell coordination, multiuser multiple-input-multiple-output (MIMO), and orthogonal frequency-division multiple access (OFDMA) relay systems (e.g., [4, 3] and references therein). Making the CDF-based scheduling more practical, limited feedback and CDF learning techniques are proposed [5]. While there have been many studies on the CDF-based scheduling, its queueing performance has been less explored. In [6], asymptotic behavior of queueing delay is analyzed for an i.i.d. (independent and identically distributed) type of arrival process using the concept of effective bandwidth and effective capacity.

The main purpose of this paper is as follows:

1. We provide a simple and explicit formula for the average queueing delay of a user under the CDF-based scheduling, so that we can easily check and understand the queueing performance through our analytic formula.

2. The formula for the average queueing delay is applicable to a general arrival process.

To this end, we approximate the queue of each user with G/M/1 queue under the assumption that wireless channel varies randomly and independently over time-slots. This assumption practically holds in a fast fading environment having high Doppler frequency [6]. Through numerical studies and simulation, we verify the tightness and efficacy of our approximation.

The rest of the paper is organized as follows. In Section 2, we describe the system model. In Section 3, we derive an approximation formula for the average queueing delay of a user under the CDF-based scheduling using G/M/1 queue approximation. In Section 4, we present numerical studies, and we conclude the paper in Section 5.

2. System Model and Assumptions

We consider downlink in a cellular network consisting of one base station (BS) and N mobile stations (MSs). At the BS, there are N buffers, denoted as buffer n ($n = 1, 2, \ldots, N$), where buffer n stores packets to be transmitted to MS n.

Let $\gamma_n(t)$ be the SNR of the wireless channel between the BS and MS n at time-slot t. Due to fading, the SNR changes over time. Moreover, channel statistics of MSs can be different from each other depending on their distance from the BS. Thus, we assume that $\gamma_n(t)$ is a random variable following a *general* distribution function $F_n(x) \triangleq \mathrm{P}(\gamma_n(t) \leq x)$. Also, we assume that channel variation processes are independent over time-slots, which is appropriate in modeling a fast fading channel. Such channel variation processes are usually independent across MSs due to independent mobility patterns of MSs.

A CDF-based scheduling operates slot-by-slot by selecting MS $n^*(t)$ at time-slot t as follows:

$$n^*(t) = \underset{1 \leq n \leq N}{\operatorname{argmax}} \left[F_n(\gamma_n(t))\right]^{1/w_n}, \tag{2.1}$$

where w_n (> 0) denotes the weight of MS n and satisfies $\sum_{n=1}^{N} w_n = 1$. Then, the channel access probability of MS n in steady state becomes [2]:

$$\lim_{t\to\infty} \mathrm{P}(n^*(t) = n) = w_n. \tag{2.2}$$

Thus, the CDF-based opportunistic scheduling can provide precise control of fairness among users by adapting the weight parameters $(w_1, w_2, \ldots, w_N)$.

We model the physical layer as an on/off fading channel; i.e., the service rate of buffer n (in packets per time-slot) is given by

$$C_n(t) = \begin{cases} 1 & \text{if } n^*(t) = n \text{ and } \gamma_n(t) \geq l, \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

where the threshold l represents the minimum SNR level required to transmit a packet successfully to a mobile station. Hence, a packet in buffer n can be successfully transmitted to MS n when buffer n is selected by the scheduler and the channel quality between the BS and MS n exceeds a certain level.

Let $A_n(t)$ be the number of packets arriving at buffer n during time-slot t. We assume that the packet inter-arrival times are *generally* distributed with the distribution function $G_n(x) \triangleq \mathrm{P}(T_n \leq x)$. Let $B_n(t)$ be the number of packets in buffer n at time-slot t. Then, the buffer length under the CDF-based scheduling evolves as

$$B_n(t+1) = [B_n(t) - C_n(t)]^+ + A_n(t),$$

where $[x]^+ = \max\{x, 0\}$.

3. Queueing Analysis

In this section, we derive an approximation formula for the average delay in buffer n under the CDF-based scheduling. To this end, we first construct a virtual queue, denoted as queue n, that mimics the dynamics in buffer n as follows.

The arrival process of customers at queue n is identical to the arrival process of packets at buffer n. The customers in queue n are served by one server, denoted as server n, as shown in Fig. 1. In each time-slot, server n is in one of the two different states: S or U. Server n is in state S at time-slot t if $n^*(t) = n$ and $\gamma_n(t) \geq l$ (i.e., server n is in state S at time-slot t if $C_n(t) = 1$ by (2.3)). Otherwise, it is in state U. A customer in server n is served completely only when server n is in state S. After service completion, the customer in server n departs the system, and the head-of-line customer waiting in queue n enters server n.

The above system of queue n and server n is then characterized as follows:

(i) The departure process of customers at server n is the same as the departure process of packets at buffer n. Since the arrival process of customers at queue n is also

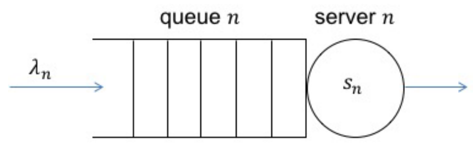

Figure 1: Analytical queueing model for MS n

identical to the arrival process of packets at buffer n, we have

$$L_n^{\text{sys}}(t) = B_n(t), \qquad t = 1, 2, 3, \ldots, \tag{3.4}$$

where $L_n^{\text{sys}}(t)$ is the total number of customers in queue n and server n at time-slot t.

(ii) Let $s_n \triangleq \mathrm{P}(n^*(t) = n, \gamma_n(t) \geq l)$. Then, server n is in state S with probability s_n and in state U with probability $1 - s_n$. Since a customer is served completely when server n is in state S, the service time of the customer (i.e., the time spent in server n until it departs the system) is geometrically distributed with parameter s_n. In the following proposition, we present the closed-form formula for s_n.

Proposition 3.1. The probability s_n that server n is in state S is given by

$$s_n = w_n(1 - [F_n(l)]^{1/w_n}).$$

Proof. By conditioning on $\gamma_n(t)$, we have

$$\begin{aligned} s_n &\triangleq \mathrm{P}(n^*(t) = n, \gamma_n(t) \geq l) \\ &= \int_l^\infty \mathrm{P}(n^*(t) = n \mid \gamma_n(t) = x)\, dF_n(x). \end{aligned} \tag{3.5}$$

According to the policy of the CDF-based opportunistic scheduling given in (2.1), we have

$$\mathrm{P}(n^*(t) = n \mid \gamma_n(t) = x) = \mathrm{P}\big([F_k(\gamma_k(t))]^{1/w_k} < [F_n(x)]^{1/w_n} \ \forall k \neq n\big). \tag{3.6}$$

Note that, for any random variable X having CDF $F(x) = \mathrm{P}(X \leq x)$, $F(X)$ becomes a uniform random variable on the interval [0, 1] [7]. Hence, each random variable $F_k(\gamma_k(t))$ for $k \neq n$ is uniformly distributed on the interval [0, 1], which is independent across k due to the independence of channel variation processes across MSs. Thus, the right-hand side of (3.6) is reduced to

$$\begin{aligned}
\mathrm{P}([F_k(\gamma_k(t))]^{1/w_k} < [F_n(x)]^{1/w_n}\ \forall k \neq n) &= \prod_{k=1,k\neq n}^{N} \mathrm{P}\big([F_k(\gamma_k(t))]^{1/w_k} < [F_n(x)]^{1/w_n}\big) \\
&= \prod_{k=1,k\neq n}^{N} [F_n(x)]^{w_k/w_n} \\
&= [F_n(x)]^{(\sum_{k=1}^{N} w_k - w_n)/w_n} \\
&= [F_n(x)]^{1/w_n - 1}, \qquad (3.7)
\end{aligned}$$

where the last equality follows from the constraint $\sum_{k=1}^{N} w_k = 1$. Combining (3.5)–(3.7) gives

$$\begin{aligned}
s_n &= \int_l^\infty [F_n(x)]^{1/w_n - 1} dF_n(x) \\
&= \lim_{b\to\infty} \left\{ w_n [F_n(b)]^{1/w_n} \right\} - w_n [F_n(l)]^{1/w_n} \\
&= w_n (1 - [F_n(l)]^{1/w_n}).
\end{aligned}$$

This completes the proof. ■

As a geometric random variable is the floor of an exponential random variable, queue n is well approximated by G/M/1 queue having service rate s_n and arrival rate λ_n $(= 1/\mathrm{E}[T_n])$. Hence, from the classical queueing theory on G/M/1 queue [8], we obtain

$$\lim_{t\to\infty} \mathrm{E}[L_n^{\mathrm{sys}}(t)] \approx \frac{\lambda_n}{s_n(1-\sigma_n)}, \qquad (3.8)$$

where σ_n is the unique root of the equation

$$\sigma_n = \int_0^\infty e^{-s_n(1-\sigma_n)x} dG_n(x). \qquad (3.9)$$

Finally, we obtain the average queueing delay in buffer n, as shown in Proposition 3.2.

Proposition 3.2. The average queueing delay of a packet in buffer n under the CDF-based scheduling is approximated by

$$\overline{D}_n \approx \frac{1}{s_n(1-\sigma_n)}.$$

Proof. From Little's law, we have

$$\overline{D}_n = \lim_{t\to\infty} \frac{\mathrm{E}[B_n(t)]}{\lambda_n}. \qquad (3.10)$$

Applying (3.4) and (3.8) sequentially to the right-hand side of (3.10) gives

$$\overline{D}_n = \lim_{t\to\infty} \frac{\mathrm{E}[L_n^{\mathrm{sys}}(t)]}{\lambda_n} \approx \frac{1}{s_n(1-\sigma_n)},$$

which completes the proof. ∎

Example 3.3. Suppose that the packet arrival process at buffer n is Poisson with rate λ_n (> 0), and MS n is subject to Rayleigh fading with the average SNR $\overline{\gamma}_n$ (> 0). That is,

$$G_n(x) = 1 - \exp(-\lambda_n x),$$
$$F_n(x) = 1 - \exp\left(-\frac{x}{\overline{\gamma}_n}\right).$$

From Proposition 3.1, we have

$$s_n = w_n\left(1 - \left[1 - \exp\left(-\frac{l}{\overline{\gamma}_n}\right)\right]^{\frac{1}{w_n}}\right).$$

In addition, (3.9) reduces to

$$\sigma_n = \frac{\lambda_n}{s_n(1-\sigma_n)+\lambda_n},$$

which gives $\sigma_n = \lambda_n/s_n$. Therefore, by Proposition 2, the average queueing delay in buffer n is

$$\overline{D}_n \approx \frac{1}{w_n(1-[1-\exp(-l/\overline{\gamma}_n)]^{1/w_n}) - \lambda_n}.$$

Example 3.4. Suppose that the packet inter-arrival times at buffer n follow a Pareto distribution with parameters α (> 0) and δ (> 0), and MS n is subject to Nakagami-m fading with the average SNR $\overline{\gamma}_n$ (> 0) [9]. That is,

$$G_n(x) = \begin{cases} 1 - \left(\dfrac{\delta}{x}\right)^{\alpha} & \text{if } x \geq \delta, \\ 0 & \text{if } x < \delta, \end{cases}$$
$$F_n(x) = 1 - \frac{1}{\Gamma(m)}\Gamma\left(m, \frac{mx}{\overline{\gamma}_n}\right),$$

where m ($\geq 1/2$) is the Nakagami fading parameter, $\Gamma(m) \triangleq \int_0^\infty x^{m-1}\exp(-x)\,dx$ is the Gamma function, and $\Gamma(m,t) \triangleq \int_t^\infty x^{m-1}\exp(-x)\,dx$ is the upper incomplete Gamma function. From Proposition 3.1, we have

$$s_n = w_n\left(1 - \left[1 - \frac{1}{\Gamma(m)}\Gamma\left(m, \frac{ml}{\overline{\gamma}_n}\right)\right]^{\frac{1}{w_n}}\right).$$

In addition, from (3.9), we have

$$\sigma_n = \alpha\delta^\alpha \int_\delta^\infty e^{-s_n(1-\sigma_n)x} x^{-(\alpha+1)}\, dx$$
$$= \alpha\, [s_n(1-\sigma_n)\delta]^\alpha\, \Gamma(-\alpha, s_n(1-\sigma_n)\delta). \tag{3.11}$$

We can find σ_n numerically by solving (3.11). Then, the average queueing delay in buffer n is

$$\overline{D}_n \approx \frac{1}{w_n(1 - [1 - \Gamma(m, ml/\overline{\gamma}_n)/\Gamma(m)]^{1/w_n})(1-\sigma_n)}.$$

4. Numerical Study

In this section, we examine the delay performance of the CDF-based scheduling through numerical studies. We assume that there are $N = 50$ MSs in the network. We fix $n \in \{1, 2, \ldots, N\}$, and consider the scenario that the packet arrival process at buffer n is a Poisson process with rate λ_n, and MS n is subject to Rayleigh fading with the average SNR $\overline{\gamma}_n$. The threshold l in (2.3) (i.e., the minimum SNR level required to transmit a packet successfully to a mobile station) is set to 4 dB.

Fig. 2 shows the impact of the weight w_n on the average queueing delay of MS n when $\lambda_n = 0.05$ (packets/slot) and $\overline{\gamma}_n \in \{2, 4, 6\}$ dB. From the figure, we observe that the average queueing delay of MS n decreases *convexly* as its weight w_n increases. By (2.2), the weight w_n represents the channel access probability of MS n. Hence, the convexity in Fig. 2 implies that the CDF-based scheduling has the rich-get-richer and the poor-get-poorer characteristics in the delay performance. For example, suppose that the channel access probability of MS n decreases 10% from 0.9 to 0.8. In this case, the average queueing delay of MS n increases 2.04% from 6.558 (slots) to 6.692 (slots) when $\overline{\gamma}_n = 2$ dB, which is less than 10%. However, if the channel access probability of MS n decreases 10% from 0.3 to 0.2, then the average delay of MS n increases 27.55% from 9.063 (slots) to 11.560 (slots), which is greater than 10%. Similar behaviors are observed for $\overline{\gamma}_n \in \{4, 6\}$ dB. That is, a user scheduled with a small channel access probability will experience much longer queueing delay than the other users. In Fig. 3, we investigate the impact of the offered load $\rho_n (= \lambda_n/s_n)$ on the average queueing delay of MS n when $w_n \in \{0.05, 0.10, 0.80\}$ and $\overline{\gamma}_n = 2$ dB. We observe that the average delay $\overline{D}_n$ increases *exponentially* with the offered load ρ_n.

Combining the results in Figs. 2 and 3, we have the following interpretations. The weight w_n is an important parameter to control the delay performance of MS n under the CDF-based scheduling. For a slight change in the wight w_n, the amount of decrease or increase in the average queueing delay of MS n becomes prominent when w_n is small and MS n has large offered load. In addition, our analytic results in Figs. 2 and 3 match the simulation results well, showing the tightness of our approximation.

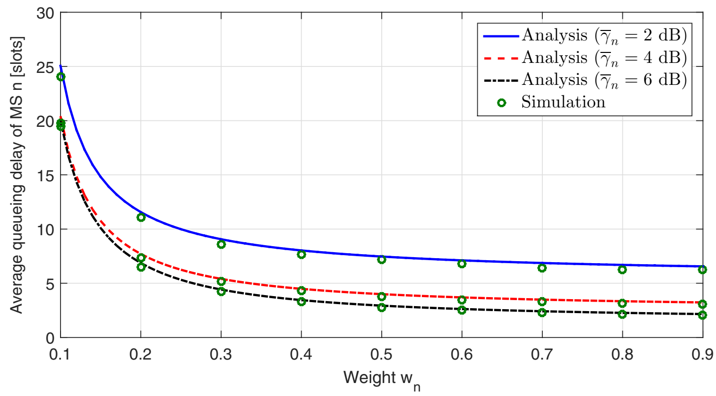

Figure 2: The impact of the weight w_n on the average queueing delay $\overline{D}_n$

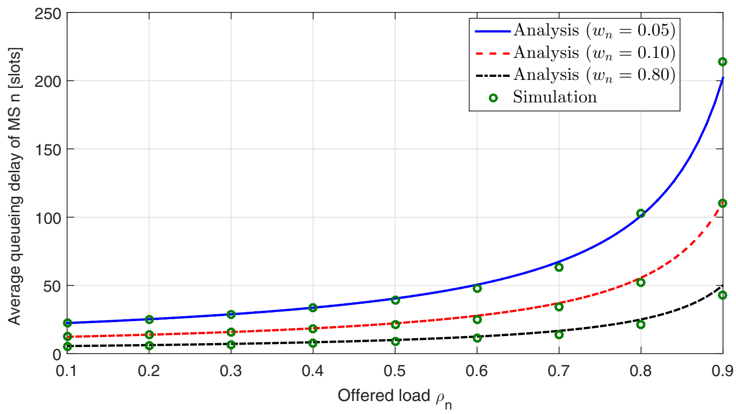

Figure 3: The impact of the offered load ρ_n on the average queueing delay $\overline{D}_n$

5. Conclusion

In this paper, we present an approximation formula for the average queueing delay of a user under the CDF-based scheduling. Our formula can be applied to a general packet arrival process under the assumption that wireless channel varies independently over time-slots. Using the formula, we investigate the delay performance of the CDF-based scheduling for various system parameters. Our future work is to analyze the delay performance of the CDF-based scheduling for time-correlated fading and a general arrival process.

Acknowledgment

This work was supported by the 2014 Research Fund of University of Ulsan.

References

[1] A. Asadi and V. Mancuso, "A survey on opportunistic scheduling in wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1671–1688, 2013.

[2] D. Park, H. Seo, H. Kwon, and B. G. Lee, "Wireless packet scheduling based on the cumulative distribution function of user transmission rates," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1919–1929, 2005.

[3] A. H. Nguyen and B. D. Rao, "CDF scheduling methods for finite rate multiuser systems with limited feedback," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3086–3096, 2015.

[4] H. Jin, B. C. Jung, and V. C. Leung, "Fundamental limits of CDF-based scheduling: Throughput, fairness, and feedback overhead," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 894–907, 2015.

[5] A. H. Nguyen, Y. Huang, and B. D. Rao, "Learning methods for CDF scheduling in multiuser heterogeneous systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3727–3740, 2014.

[6] D. Park and B. G. Lee, "QoS support by using CDF-based wireless packet scheduling in fading channels," *IEEE Transactions on Communications*, vol. 54, no. 11, pp. 2051–2061, 2006.

[7] S. M. Ross, *Stochastic processes*. John Wiley & Sons New York, 1996.

[8] L. Kleinrock, *Queueing systems, volume I: Theory*. Wiley Interscience, 1975.

[9] Y. Kim and S. Kwon, "Capacity analysis of opportunistic scheduling in Nakagami-*m* fading environments," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5379–5384, 2015.