J. Appl. Math. & Informatics Vol. 26(2008), No. 1 - 2, pp. 389 - 393
Website: http://www.kcam.biz

GENERAL FORMULAS OF SOME VACATION MODELS

JONG SEUL LIM

ABSTRACT. This paper describes a single-server queue where the server is unavailable during some intervals of time, which is referred to as vacations. The major contribution of this work is to derive general formulas for the additional delay in the vacation models of the single vacations, head of line priority queues with non-preemptive service, and multiple vacations and idle time.

AMS Mathematics Subject Classification : 60K25
Key words and phrases : Delay, queue, priority, Poisson arrival.

1. Introduction

This paper describes a single-server queue where the server is unavailable during some intervals of time, which is referred to as vacations. The major contribution of this work is to derive general formulas for the additional delay in the basic model. The policy can be dependent on the past behavior of the system. The system admits the server to stay idle or to wait for a while before a vacation is taken after serving all customers. We focus on an queue with the complex vacation policies. The single vacations, head of line priority queues with non-preemptive service, and multiple vacations and idle time are considered. For these vacation policies, a general formula for the additional delay is derived. The formula can also be extended to cases with multiple types of vacations.

2. Vacation model with single vacations

Under this policy, the server takes a vacation after serving all customers. Upon return from a vacation, the server starts to serve the backlogged customers if any; otherwise,it simply remains idle and waits for the next arrival. To find P_0 by analyzing $Z_v(\hat{t})$, let a cycle be the time interval in $\hat{t}$ between two jumps of $Z_v(\hat{t})$. Each cycle consists of a vacation and a possible idle time. If no sampling

Received June 21, 2007.

point arrives during a vacation, the server has just finished and is returned from a vacation when $Z_v(\hat{t})$ decreases to zero. According to the policy, the server starts to stay idle in the system. This results in an idle time, which ends when the next sampling point which is a busy period arrives. Immediately after the sampling point, that is, the end of the busy period, another jump occurs and a new vacation starts in $Z_v(\hat{t})$. Let $v(\hat{t})$ be the probability density function (pdf) for a vacation length. The probability that no sampling point arriving in a vacation is $V^*(\lambda) = \int_0^\infty e^{-\lambda\hat{t}} v(\hat{t}) d\hat{t}$. As discussed earlier, the sampling points have an exponential interarrival time with mean $1/\lambda$. Hence, given that an idle time occurs, its average length is $1/\lambda$.

In contrast, if at least one sample point has arrived in a vacation, a jump occurs when $Z_v(\hat{t})$ becomes zero again. This is so because the server needs to take a vacation after all customers have been served. In this case, the cycle ends at the end of the vacation.

The PASTA(Poisson Arrivals See Time Averages) property implies that a Poisson probing stream provides an unbiased estimate of the desired time average, which states that, under very general conditions, the fraction of Poisson arrivals that observe an underlying process in a particular state is equal, asymptotically, to the fraction of time the process spends in that state. Combining both cases and using the PASTA property, the probability that an arbitrary sampling point finds $Z_v(\hat{t}) = 0$ is

$$P_0 = \frac{\textit{Avg Length of an Idle Time in a Cycle}}{\textit{Aug. Cycle Length}} = \frac{V^*(\lambda)/\lambda}{\overline{v} + V^*(\lambda)/\lambda} \tag{1}$$

After substituting (9) into (6) and some algebraic manipulation, we obtain the Laplace transform for the customer response time

$$T^*(s) = \frac{(1-\rho)}{\lambda\overline{v} + V^*(\lambda)} \left\{ \frac{\lambda[1 - V^*(s)] + sV^*(\lambda)}{s - \lambda + \lambda X^*(s)} \right\} X^* \tag{2}$$

which is identical to (22) in Levy and Yechiali (1975).

3. Head of line priority queues with non-preemptive service

Consider a head of line priority system with types of queues *H*, *P* and *L* customers where types *H* and *L* customers have the highest and lowest priority respectively, and type *P* customer has a priority which is not the highest and the lowest. Let $\lambda_H, \overline{x}_H, X_H^*(s)$ and $\rho_H (= \lambda\overline{x}_H)$, denote the arrival rate, average service time, the Laplace transform for service time and server utilization for type-*H* customers, respectively. The same notation applies to other customer types.

Let us find the Laplace transform for the waiting time for type-P customers. Since those type-H customers arriving during a customer service time are served prior to any type-P customers, the Laplace transforms for the effective service time of various types become

$$\begin{aligned} G_H^*(s) &= X_H^*(s+\lambda_H-\lambda_H G_H^*(s)),\\ G_P^*(s) &= X_P^*(s+\lambda_H-\lambda_H G_H^*(s)),\\ G_L^*(s) &= X_L^*(s+\lambda_H-\lambda_H G_H^*(s)). \end{aligned} \tag{3}$$

To find the waiting time for an arbitrary type-P customer, we first characterize the amount of work, U_{PH}, of types P and H customers in the system at the arrival instant. Now, let us combine queues H and P together to from a single queue PH with $\lambda_{PH} = \lambda_H + \lambda_P$ and the L.T. for service time, $X_{PH}^*(s) = [\lambda_H X_H^*(s) + \lambda_P X_P^*(s)]/\lambda_{PH}$.

As far as queue PH is concerned, service provided for a type-L customer can be treated as a "vacation." Note that the server takes a vacation to serve a type-L customer after queue PH becomes empty only if type-L customer(s) exist in the system. Clearly, at the time queue PH becomes empty, whether type-L customers exist or not at that point in time depends on the system behavior since the last time the system was empty. Let $U_{PH}^*(s)$ be the Laplace transform for the amount of work in queue PH found by an arbitrary arrival at the queue. We have

$$U_{PH}^*(s) = \frac{(1-\lambda_{PH}\overline{x}_{PH})s}{s-\lambda_{PH}-\lambda_{PH}X_{PH}^*(s)}\left[P_0 + (1-P_0)\frac{1-X_L^*(s)}{s\overline{x}_L}\right] \tag{4}$$

where $\overline{x}_{PH} = (\lambda_H\overline{x}_H + \lambda_P\overline{x}_P)/\lambda_{PH}$ and P_0 is the probability that a PH busy period starts when the server is idle.

To find P_0, we construct $Z_v(\hat{t})$ by contracting all PH busy periods into sampling. Let π_0 be the fraction of virtual time $\hat{t}$ that $Z_v(\hat{t}) = 0$. Since the service provided for all types H and P customers has already been removed from $Z_v(\hat{t})$ by the contraction operation, it is clear that $\pi_0 = (1-\rho_H-\rho_P-\rho_L)/(1-\rho_H-\rho_P)$. By the PASTA property, we have

$$P_0 = \pi_0 = \frac{1-\rho_H-\rho_P-\rho_L}{1-\rho_H-\rho_P}. \tag{5}$$

Substitution of (4) into (5), after algebraic manipulation, yields

$$U_{PH}^*(s) = \frac{(1-\rho_H-\rho_P-\rho_L)s + \lambda_L[1-X_L^*(s)]}{s-\lambda_H-\lambda_P+\lambda_H X_H^*(s)+\lambda_P X_P^*(s)} \tag{6}$$

Since queue PH actually has two independent streams of Poisson arrivals, $U^*_{PH}(s)$ also characterizes U_{PH} found by an arbitrary arriving type-P customer. As the subsequent arrivals of type-H customers are served prior to type-P customers, the waiting time for type-P customers can be obtained by the delay-cycle analysis with the initial delay given by U^*_{PH}. Hence, using (3) and (6), the Laplace transform for waiting time for type-P customers is

$$W^*_P(s) = U^*_{PH}(s + \lambda_H - \lambda_H G^*_H(s))$$
$$= \frac{(1-\rho_H-\rho_P-\rho_L)\left[s+\lambda_H-\lambda_H G^*_H(s)\right] + \lambda_L\left[1 - X^*_L(s+\lambda_H-\lambda_H G^*_H(s))\right]}{s-\lambda_P+\lambda_P X^*_P\left(s+\lambda_H-\lambda_H G^*_H(s)\right)}. \tag{7}$$

This is identical to (3.32) of Kleinrock (1976).

4. Multiple vacations and idle time

In this model, the server chooses to take a multiple vacation or remain idle in the system after all customers have been served. Once a decision is made to stay idle, the server cannot initiate a vacation until new customer(s) arrive and are exhaustively served. Assume that the vacation lengths are i.i.d. and the vacation decision and the subsequent vacation length(s) are mutually independent.

The process $Z_v(\hat{t})$ can be constructed for this policy as discussed above. Let β_M and $1-\beta_M$ be the long-time fraction of decision epochs at which the server initiates a multiple vacation and remains idle, respectively. A cycle is defined as the virtual time interval between two decision epochs. Using the above approach, by conditioning on whether the server chooses to take vacations or to stay idle, one can obtain the average cycle length as

$$\bar{c} = \beta_M \sum_{k=0}^{\infty} [V^*(\lambda)]^k \bar{v} + (1-\beta_M)\frac{1}{\lambda}. \tag{8}$$

Note that the last term of (5) is the average length of an idle time in a cycle. Let π_0 be the fraction of virtual time $\hat{t}$ at which the server is idle. Clearly, π_0 is equal to the ratio of that term to $\bar{c}$. Using the PASTA property again, we have $P_0 = \pi_0$ which is given by

$$P_0 = \frac{(1-\beta_M)/\lambda}{\bar{c}} = \frac{(1-\beta_M)[1-V^*(\lambda)]}{\lambda\bar{v}\beta_M + (1-\beta_M)[1-V^*(\lambda)]}. \tag{9}$$

Some algebraic manipulations using (9) yields

$$U^*_B(s) = \frac{(1-\rho)}{s-\lambda+\lambda X^*(s)}\left\{\frac{(1-\beta_M)[1-V^*(\lambda)]s + \lambda\beta_M[1-V^*(s)]}{\lambda\bar{v}\beta_M + (1-\beta_M)[1-V^*(\lambda)]}\right\}. \tag{10}$$

5. Conclusions

We have analyzed a queue with complex vacation. General formulas for the vacation policies in the vacation models of the single vacations, head of line priority queues with non-preemptive service, and multiple vacations and idle time have been obtained. The analysis approach in this paper can be applicable to other related queueing models, if they conform with the basic model considered in this paper. Further, these results can serve as a basis for the formulation and solution of certain optimization problems involved in the similar models.

REFERENCES

1. B.T. Doshi, *Conditional and Unconditional Distributions for M/G/1 Type Queue with Server Vacation,* to appear in Queueing Systems, 1990.
2. S.W. Fuhmann, and R.B. Cooper, *Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations,* Opns.Res., 33(1985), 1117-1129.
3. O. Kella, and U. Yechiali, *Priorities in M/G/1 Queues with Server Vacations,* Naval Res. Logistics, 35(1988), 23-34.
4. L. Kleinrock, *Queueing Wywtems Vol.II: Computer Application,* John Wiley & Sons, Inc., New York, 1976.
5. H. Levy and L. Kleinrock, *A Queue with Starter and A Queue with Vacations: Delay Analysis by Decomposition,* Opns.Res., 34(1986), 426-436.
6. Y. Levy and U. Yechiali, *Utilization of Idle Time in an M/G/1 Queueing System,* Managemant Sci., 22(1975), 202-211.
7. J.G. Shanthikumar, *Level Crossing Analysis of Priority Queues and a Conservation Identity for Vacation Models,* Naval Res. Logistics, 36(1989), 797-806.
8. R.W. Wolff, *Sample-Path Derivations of the Excess,Age, and Spread Distributions,* J.Appl.Prob. 25(1988), 432-436.
9. K. Altinkerner, *Average Waiting Time of Customers in An M/D/k Queue With Nonpreemptive Priorities,* Vol. 25, No 4, 317-328, 1999.

Jong Seul Lim received the BS degree from Seoul National University and the Ph.D degree in communications and operations engineerg from Ploytechnic University, New York. He has worked with Sunmoon University, Korea as a professor. His research interests focus on the queueing network analysis and mobile cellular communication systems.

Division of Information Communications Engineering, Sunmoon University, Chungnam, Korea

e-mail: `jlim@sunmoon.ac.kr`