I－3（S）
招待講演

1995年度日本オペレーションズ・リサーチ学会
秋季研究発表会

On a Distributional Version of Little's Law

Yoshitaka TAKAHASHI
NTT Telecommunication Networks Laboratories, 3-9-11 Midori, Musashino180, Tokyo

1. Introduction

The well known Little's Law ($L = \lambda W$) for queueing systems relates the (time-average) mean queue length to the (customer-average) mean waiting time in the system. Over thirty years since Little's law first appeared [5], its simplicity and importance have established it as a basic tool of queueing theory. Although there is considerable knowledge about Little's law and its extensions [10,11], its distributional version for many systems is still an open question [1,2,3,4]. For example, a distributional relationship between queue length and waiting time for a multi-class batch-arrival priority queue has been recently obtained in Takahashi & Miyazawa [8]. The primary purpose of this talk is to illustrate how such a relationship can be obtained via the point process approach. To simplify the presentation, we will only deal with a continuous-time system. For the discrete-time systems, see Takahashi & Miyazawa [9].

2. Preliminaries

We begin with the following assumptions. (A1) There exists a marked point process:

$$\omega = \{ (t_i, X_i, S_i(1), \cdots, S_i(X_i)) \}_{i=-\infty}^{+\infty}$$

which has a probability space $(\Omega, \mathcal{F}, P)$ and is strictly stationary with respect to the shift operator $\mathbb{T}$, where $\{t_i\}_{i=-\infty}^{+\infty}$ is a set of real numbers with no point accumulation such that $\cdots < t_{-1} < t_0 < 0 \le t_1 < t_2 < \cdots$. The X_i and $S_i(j)$ $(1 \le j \le X_i)$ take values in some measurable space $(K, \mathcal{K})$, while $\mathbb{T}$ is the operator on Ω such that $\mathbb{T}^s\omega = \{ (t_i + s, X_i, S_i(1), \cdots, S_i(X_i)) \}_{i=-\infty}^{+\infty}$ for real s. (A2) $X(t)(\omega)$ is a measurable function of (t,ω) from $(R\times\Omega, \mathcal{B}(R)\times\mathcal{F})$ to $(E, \mathcal{B}(E))$, i.e., $\{X(t)\}$ is a measurable process. (A3) For any s and t, $X(t)(\omega) = X(t-s)(\mathbb{T}^s\omega)$ $(\forall\omega\in\Omega)$.

Let N_b be the arrival point process of batches with intensity $\lambda_b \equiv EN_b[0,1) < +\infty$. Let P_{N_b} and E_{N_b} be the Palm distribution with respect to N_b, and its expectation, respectively. The following lemma is then verified as in Miyazawa[6]. For a multi-class extension, see Takahashi & Miyazawa [8].

Lemma 2.1 Consider a $GI^X/GI/1$ queue. Let T_1 the inter-arrival time between the 0-th and 1-st batches. For any bounded non-negative function u, we have

$$E(u(\omega)) = \lambda_b E_{N_b}[\int_0^{T_1} u(\mathbb{T}^s\omega)\, ds]. \tag{2.1}$$

Moreover, if there exists a process $\{Z(t)\}$ satisfying that $Z(s,\omega) = u(\mathbb{T}^s\omega)$ $(0 < s < T_1)$ (a.s. P_b) and that for each $s > 0$ $Z(s)$ and $\{T_1 > s\}$ are P_{N_b}-independent each other, then

$$E(u) = \lambda_b \int_0^{+\infty} P_{N_b}(T_1 \ge s) E_{N_b}[Z(s)] ds. \tag{2.2}$$

3. Batch arrival priority queue

We consider a $\overrightarrow{GI^X} / \overrightarrow{GI} /1$ priority queue with I classes. We assume that a customer with a smaller index has precedence over a customer with a greater index. For the priority queue, we use subindex p (signifying class p) on each of the corresponding notations for the single-class (non-priority) queue in the literature [6]. For example, $l_p(t)$ denotes the number of class p customers in the system at time t, and $W_{p,n}$ the waiting time of the first customer in the n-th batch of class p. Let $N_{p,b}$ be the arrival point process of class p batches. Let $P_{p,b}$ be the Palm distribution with respect to the point process $N_{p,b}$. Denote by $E_{p,b}$ (or E) the expectation with respect to $P_{p,b}$ (or P).

To denote functions (or transforms) for class p, we also use subindex p on each of the corresponding functions (or transforms) for the single-class queue. For example, $\mathcal{l}_p(z) \equiv E(z^{l_p(t)})$ denotes the pgf for queue-length distribution, and $W_p^*(s) \equiv E_{p,b}(e^{-sW_{p,n}})$ denotes the LST for the waiting time distribution of the first customer in a class p batch.

The completion time $C_{p,m}$ of the m-th customer of class p is defined as the interval from the moment at which the m-th class p customer enters service to the first moment at which there are no higher class $\{1,2,\cdots,p-1\}$ customers in the system. We denote by $C_p^*(s) \equiv E_{p,b}(e^{-sC_{p,n}})$ the LST for class p completion time distribution.

Suppose that a class p batch arrived at the queue at the origin of time axis $(1 \le p \le I)$. Consider an event $\{l_p(s) \ge j\}$ for $j \ge 0$ and s $(0 < s < t_{p,1})$ in the queue. We then note that

$$\{l_p(s) \ge j\} = \{\text{(at least) class p customers out of the ones who arrived at the queue before time 0 still remain in system}\}. \tag{3.1}$$

Denoting by i $(i \ge 0)$ be the index of the eldest (class p) customer out of the ones who still remain in the system at time s, we have

$$X_{p,-i} + X_{p,-i+1} + \cdots + X_{p,0} \ge j > X_{p,-i+1} + \cdots + X_{p,0}. \tag{3.2}$$

If we assume that $X_{p,-i+1} + \cdots + X_{p,0} = k$ and $X_{p,-i} = m$, we have the following correspondence under the NP (non-preemptive) priority rule.

{(at least) j - k class p customers out of $X_{p,-i} = m$ still remain in system at time s}

$$\stackrel{d}{=} \{W_{p,0} + C_{p,1} + \cdots + C_{p,m-(j-k)} + S_{p,m-(j-k)+1} > T_{p,1} + \cdots + T_{p,i} + s\} \quad (\text{a.s. } P_{p,b}). \tag{3.3}$$

Here, we used the convention for empty sum, e.g., $C_{p,1} + \cdots + C_{p,0} = 0$. We can also treat the wait-length process but we will omit here. Under the PR (preemptive-resume) rule, if we distinguish between the number of customers in the waiting room and that in *limbo* , a similar discussion can be developed. Applying a multi-calss version of Lemma 2.1 we have the following proposition.

Proposition 3.1 In a multi-class $\overrightarrow{GI^X}/\overrightarrow{GI}/1$ queue, we have for an individual class p

$$P(l_p \geq j) \equiv P(l_p(0) \geq j) = \lambda_{p,b} \int_0^{+\infty} P_{p,b}(T_p > s) P_{p,b}(l_p(s) \geq j) ds \quad (j \geq 0), \tag{3.4}$$

where l_p denote the class-p stationary queue length.

Proposition 3.1 finally yields the following theorem and corollary, which link the queue-length and waiting time distributions in a priority queue.

Theorem 3.2 Consider a $\overrightarrow{GI^X}/\overrightarrow{GI}/1$ priority queue with I classes. For an individual class p ($1 \leq p \leq I$), we have

$$\mathcal{T}_p(z) = 1 - \frac{1-z}{z} \lambda_{p,b} \sum_{m=1}^{\infty} \sum_{j=1}^{m} \int_0^{+\infty} P_{p,b}(W_{p,0} + C_{p,1} + \cdots + C_{p,m-(j-k)} + S_{p,m-(j-k)+1} > s)\, z^j P_{p,b}(X_p = m)$$
$$\times E_{p,b}(\tilde{X}_p(z)^{N_{p,b}(0,s)}) ds \qquad \text{under NP}, \tag{3.5}$$

$$\mathcal{T}_p(z) = 1 - \frac{1-z}{z} \lambda_{p,b} \sum_{m=1}^{\infty} \sum_{j=1}^{m} \int_0^{+\infty} P_{p,b}(W_{p,0} + C_{p,1} + \cdots + C_{p,m-(j-k)} + C_{p,m-(j-k)+1} > s)\, z^j P_{p,b}(X_p = m)$$
$$\times E_{p,b}(\tilde{X}_p(z)^{N_{p,b}(0,s)}) ds, \qquad \text{under PR}. \tag{3.6}$$

Corollary 3.3 Consider an $\overrightarrow{M^X}/\overrightarrow{GI}/1$ priority queue with I classes. For an individual class p ($1 \leq p \leq I$), we have

$$\mathcal{T}_p(z) = \frac{1-z}{1-\tilde{X}_p(z)} W_p^*(\sigma_p) S_p^*(\sigma_p) \frac{\tilde{X}_p(C_p^*(\sigma_p)) - \tilde{X}_p(z)}{C_p^*(\sigma_p) - z} \qquad \text{under NP}, \tag{3.7}$$

$$\mathcal{T}_p(z) = \frac{1-z}{1-\tilde{X}_p(z)} W_p^*(\sigma_p) C_p^*(\sigma_p) \frac{\tilde{X}_p(C_p^*(\sigma_p)) - \tilde{X}_p(z)}{C_p^*(\sigma_p) - z} \qquad \text{under PR}. \tag{3.8}$$

Here, $\sigma_p \equiv \sigma_p(z) \equiv \lambda_{p,b}(1 - \tilde{X}_p(z))$.

Remark 3.4 a) The functions $C_p^*(\sigma_p)$ and $W_p^*(\sigma_p)$ ($1 \leq p \leq I$) in (3.7) and (3.8) of Corollary 3.3 were previously obtained via the delay-cycle approach, see Takahashi & Shimogawa [7]. b) Note that (3.7) reduces to those of Keilson & Servi [4] for two-class (I = 2) Poisson (non-batch) input priority system. ∎

Acknowledgments I would like to thank Professor Masakiyo Miyazawa (SUT) for his helpful comments during the course of this work. I would like to thank Professor Yukio Takahashi (TIT), Professor Masao Mori (TIT), and Professor On Hashida (U of Tsukuba, Tokyo) for their patient teaching and constant encouragement. I am also indebted to Dr. Fumiaki Machihara (NTT) and Dr. Konosuke Kawashima (NTT) for their thoughtful discussions in my daily research.

References

[1] D. Berstsimas & D. Nakazato, The distributional Little's law and its applications, *Operations Res.*, **43** (1995) 298-310.

[2] S. L. Brumelle, A generalization of L = λW to moments of queue length and waiting times, *Operations Res.*, **20** (1972) 1127-1136.

[3] J. Keilson & L. D. Servi, A distributional form of Little's law, *Operations Res. Letters*, **7** (1988) 223-227.

[4] J. Keilson & L. D. Servi, The distributional form of Little's law and the Fuhrmann-Cooper decomposition, *Operations Res. Letters*, **9** (1990) 239-247.

[5] J. Little, A proof of the theorem L = λW, *Operations Res.*, **8** (1961) 383-387.

[6] M. Miyazawa, A formal approach to queueing processes in the steady state and their applications, *J. Appl. Prob.*, **16** (1979) 332-346.

[7] Yo. Takahashi & S. Shimogawa, Composite priority single-server queue with structured batch inputs, *Stochastic Models*, **7** (1991) 481-497.

[8] Yo. Takahashi & M. Miyazawa, Relationship between queue-length and waiting time distributions in a priority queue with batch arrivals, *JORSJ*, **37** (1994) 48-63.

[9] Yo. Takahashi & M. Miyazawa, Relationship between queue length and waiting time in a stationary discrete-time queue, *Stochastic Models*, **11** (1995) 249-271.

[10] S. Stidham, A last word on L = λW, *Operations Res.*, **22** (1974) 417-421.

[11] W. Whitt, A review of L = λW and extensions, *QUESTA*, **9** (1991) 235-268.