Derivation of Open System from Closed System Using Duality

Su Hyeon Namn
Dept. of MIS
Hannam University

Choong Nyoung Kim
Dept. of Business Admistration
Hallym University

Extended Abstract

Consider a closed queueing network (CQN) of queues of the "product-form" type with k stations and N packets. Station i has s_i parallel exponential servers ($s_i \geq 1$); each server at station i has a service rate μ_i ($0 < \mu_i < \infty$), $i=1,\cdots,k$. Packet routing follows a Markov chain which is finite and regular, with routing matrix [γ_{ij}], where γ_{ij} is the routing probability from station i to j. Let $\nu_i > 0$ $(i=1,\cdots,k)$ be the solutions to the following equations:

$$\sum_{j=1}^{k} \nu_i \gamma_{ji} = \nu_i \ (i=1,\cdots,k), \ \sum_{i=1}^{k} \nu_i = 1. \quad (1)$$

Let be the total number of packets at station i which includes the packets in service and in the buffer. We call X_i to be the queue length at station i. Then the equilibrium joint queue length distribution has the following product form:

$$P(X_1 = n_1, \cdots, X_k = n_k) =$$

$$G^{-1}(N) \cdot \prod_{i=1}^{k} \rho_i^{n_i} / \theta(n_i), \quad (2)$$

where $\rho_i = \nu_i / \mu_i$ $(i=1,\cdots,k)$ and $\theta(n_i) = n_i!$ (for $0 \leq n_i \leq s_i$) and $s_i! s_i^{n_i - s_i}$ (for $s_i \leq n_i \leq N$), and $G(N)$ to be the normalization constant:

$$G(N) = \sum_{\sum n_i = N} \prod_{i=1}^{k} \rho_i^{n_i} / \theta(n_i). \quad (3)$$

Then the throughput function of the network is:

$$TH = \sum_{i=1}^{k} TH_i = \sum_{i=1}^{k} \nu_i G(N-1)/G(N). \quad (4)$$

To evaluate TH in (4) exactly, such algorithms as convolution [3], mean value analysis [7], and generating function approach [4] have been proposed. However, the complexity of those algorithms grows directly as k, N, and s_i. Moreover, as the network becomes bigger, the computation of $G(N)$ exhibits numerical overflow problems. For these reasons, Whitt [9] introduced a method known as the Fixed Population Mean (FPM), which approximates the given closed network by the corresponding open network with specified expected equilibrium populations. In the FPM, he approximates each station in the CQN by an $M/M/s_i$ queues (with, of course, different parameters at each queue). Note that FPM treats each station independently. In this way he derives a procedure to obtain an approximate throughput.

In this paper we show an alternative approach to the FPM, which captures the dependency of a station on the remaining stations of the network more accurately. Our approach uses dual property of each station in the CQN. In particular, we use the role inversion [6]. In the following we propose the dual models of CQN along with approximate dual models. In addition, we will compare the quality of our approximate dual open system with that of FPM in the case where $s_i = 1$, for all i.

When $s_i = 1$ for all i, the CQN is composed of k stations, where station i may be viewed as an $A_i/M/1/N$. The arrival process A_i is non-renewal process and the distribution of service times is exponential with rate μ_i. The $A_i/M/1/N$ station has a dual station $M/A_i/N/N$ where we interchange the arrival and service processes. Thus, when j servers are idle in the $A_i/M/1/N$ station, j packets are in the $A_i/M/1/N$ system. Therefore, the CQN is a set of $A_i/M/1/N$, $i=1,\cdots,k$, stations, or a set of $M/A_i/N/N$, $i=1,\cdots,k$, stations where N servers are shared among k stations. This system we call a Dual Closed Queueing Network (DCQN). We approximate the $M/A_i/N/N$ station by the standard

$M/G/N/N$ station, where the general service times are independent and identically distributed. Since $M/G/N/N$ is just the Erlang loss system, we can make use of the previous results (for example see [5]).

Extending this approach to the case where $s_i \geq 1$ and using the similar approach to the above, station i in the CQN can be viewed as $A_i/M_i/s_i/1/N$, where the number of input sources is 1. Based on the queue description, the quasi-dual dual (see [6] for the definition) system becomes $M_i/A_i/1/s_i/N$. For this quasi-dual system to be stochastically equivalent to the corresponding station i in CQN, we need to introduce another object, cabs, which the arriving packets should take in order for them to get admitted in the network. Since the maximum number of packets in the network is limited to N, the dual system is given by $M_i/A_i/N/s_i/N$, $i=1,\cdots,k$. Note that $M_i/A_i/N/s_i/N$ is a generalized loss system with state dependent arrivals. For approximate dual system, we may invoke the result of [1], where $M/M/s$ finite source model is identical to $M/M/s$ machine interfernce model in most respects.

Finally we compare the quality of our dual system with that of FPM in terms of the network throughput when $s_i=1$. To do that, we take an upper bound of the Erlang loss system from [5]. The procedure by FPM reduces to

$$\sum_{i=1}^{k} \frac{TH_f \rho_i}{N-NTH_f \rho_i} = 1. \qquad (5)$$

On the other hand, the procedure by our dual system becomes

$$\sum_{i=1}^{k} \frac{TH_d \rho_i}{N-(N-1)TH_d \rho_i} = 1. \qquad (6)$$

TH_f and TH_d refer to the approximate throughput obtained by FPM and our dual model, respectively. Here we numerically compare the mean error of TH_d with that of TH_f from the exact throughput, TH, using a small network ($k=2$). Since the utilization (ρ_i) at a bottleneck station can always be scaled to be 1, we fix $\rho_1=1$. To examine the mean error we only need to vary ρ_2 from 0 to 1 and obtain

$$\int_0^1 [\ TH(\rho_1=1, \rho_2) - T(\rho_1=1, \rho_2)]\ d\rho_2,$$

where T is one of the approximate throughputs we will compare with the exact throughput. For numerical integration we use a symbolic manipulation language, Maple.

Mean Errors of TH_d (*T_d*), TH_f (*T_f*)

N	10		20		30	
Method	*T_d*	*T_f*	*T_d*	*T_f*	*T_d*	*T_f*
Error	0.006	0.093	0.003	0.050	0.002	0.034

As is indicated in the table, our dual system provides better results than the primal model of FPM.

Final comments go to the application of our proposed model: 1) Performance evaluation of telecommunication network protocols such as the sliding window [8], and 2) The analysis of the flexible manufacturing systems [2].

References

[1] Bunday, B. and R. Scraton, "The G/M/r Machine Interfernce Model", *European Journal of Operations Research*, Vol. 4 (1980), pp. 399-402.

[2] Buzacott, J. and J. Shanthikumar, *Stochastic Models of Manufacturing System*, Prentice Hall, New Jersey, 1993.

[3] Buzen, J., "Computational Algorithms for Closed Queueing Networks with Exponential Servers", *Communications of the ACM*, Vol. 16 (1973), pp. 527-531.

[4] Gordon, J., "The Evaluation of Normalizing Constants in Closed Queueing Networks", *Operations Research*, Vol. 38 (1990), pp. 863-869.

[5] Harel, A., "Sharp Bounds and Simple Approximations for the Erlang Delay and Loss Formulas", *Management Science*, Vol. 34 (1988), pp. 959-972.

[6] Hlynka, M. and T. Wang, "Comments on Duality of Queues with Finite Buffer Sizes", *Operations Research Letters*, Vol. 14 (1993), pp. 29-33.

[7] Reiser, M., "Mean Value Analysis and Convolution Method for Queue-Dependent Servers in Closed Queueing Networks", Performance Evaluation, Vol. 1 (1981), pp. 7-18.

[8] Schwartz, M., *Telecommunication Networks: Protocols, Modeling and Analysis*, Addison Wesley, Massachusetts, 1987.

[9] Whitt, W., 1984. "Open and Closed Models of Queues", *AT&T Bell Lab Technical Journal*, Vol. 63 (1984), pp. 1911-1979.