Performance Estimation of AS/RS using M/G/1 Queueing Model with Two Queues

M/G/1 대기모델을 이용한 자동창고 시스템의 성능 평가

Moon Hwan Lee, Si Yeong Lim, Sun Hur and Young Hae Lee
Department of Industrial Engineering, Hanyang University
이문환, 임시영, 허선, 이영해
한양대학교 산업공학과

Abstract

Many of the previous researchers have been studied for the performance estimation of an AS/RS with a static model or computer simulation. Especially, they assumes that the storage/retrieval (S/R) machine performs either only single command (SC) or dual command (DC) and their requests are known in advance. However, the S/R machine performs a SC or a DC, or both or becomes idle according to the operating policy and the status of system at an arbitrary point of time. In this paper, we propose a stochastic model for the performance estimation of a unit-load AS/RS by using a M/G/1 queueing model with a single-server and two queues. Expected numbers of waiting storage and retrieval commands, and the waiting time in queues for the storage and retrieval commands are found.

1. INTRODUCTION

Many researchers studied the travel time for an S/R machine under various storage assignment policies, and proposed analytical model by using statistical approaches [1–3]. Several researchers studied sequencing of retrievals in order to reduce the travel time of an S/R machine [4, 5]. However, all of these researches do not reflect the dynamic nature of an AS/RS, which is the realistic operating characteristic. Especially, they assumed that all storage and retrieval commands are known in advance. Under consideration of dynamic operating characteristic, their studies either under- or overestimate for the performance of S/R machine or can not provide feasible alternatives for the important design factors of an AS/RS such as the buffer size and utilization of S/R machine. The analytical stochastic analysis was presented by Bozer and White [6] for a mini-load AS/RS, and presented by Lee [7] for a unit-load AS/RS using queueing theory. Their analyses also do not reflect the dynamic aspect of the system either, since they assumed specific distribution types of service times.

In this paper, we present an analytical model for the performance estimation of a unit-load AS/RS that overcomes those drawbacks of the previous works. This model employs a queueing theory and stochastic analysis. It quickly estimate system performance for existing system or a large number of design alternatives such as expected numbers of waiting storage and retrieval commands.

2. MODEL AND DEFINITIONS

2.1 Description of model and assumptions

The system, considered in this paper, is depicted in Fig. 1. To analyze the model proposed in this research, we make the following assumptions.

(1) Customers (commands) arrive at the system according to the Poisson process with rates λ_1, λ_2 for storage and retrieval commands, respectively, and their arrival processes are independent.

(2) The services are independent and identically distributed. The service times, S, for SC and DC are generally distributed with mean E (S) = $1/\mu$.

(3) The service follows FCFS (First Come First Served) rule. If both customers exist in queue, the server performs a DC. If only one of either storage or retrieval commands exist, the server performs storage SC or retrieval SC. Otherwise it becomes idle at the I/O station.

(4) The buffer sizes for each storage and retrieval command are not limited.

We model the system as a two-dimensional Markov process with a single server and two queues: a storage commands queue and a retrieval commands queue. That is, a state (i, j) means that there are i storage commands waiting in storage queue and j retrieval commands waiting in retrieval queue. Thus, fore example, if the state is $(i+1, 0)$ or $(0, j+1)$, then the server performs storage SC or retrieval SC. If no command arrives during a service time, then the state becomes $(i, 0)$ or $(0, j)$, respectively. Fig. 2 shows the state-transition diagram based on these assumptions.

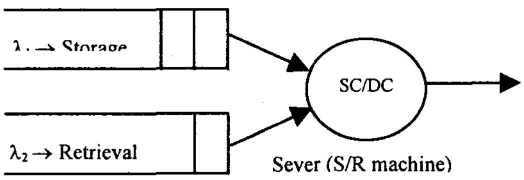

Fig. 1 Queueing model with a single server and two queues

2.2 Definitions

Let us define the following notations, probabilities and Laplace-stieltjes transforms (LST):

λ_1, λ_2	:	Arrival rates for storage and retrieval commands, respectively
μ	:	Service rate
$s(x)$	:	pdf of service time S
$S^*(\theta)$	:	LST of S
$N_S(t)$	:	# of commands in storage queue at time t
$N_R(t)$	:	# of commands in retrieval queue at time t
$S_+(t)$	:	Remaining service time for the commands in service at time t
$Q_0(t)$	=	$\Pr(N_S(t)=0,\ N_R(t)=0,\ Sever\ idle)$, $Q_0 = \lim_{t\to\infty} Q_0(t)$
$P_{i,j}(x,t)dx$	=	$\Pr(N_S(t)=i,\ N_R(t)=j,$ Sever busy, $x< S_+(t)\le x+dx)$, $i, j = 0, 1, 2, \cdots$
$P_{i,j}(x)dx$	=	$\lim_{t\to\infty} P_{i,j}(x, t)dx$
$P^*_{i,j}(\theta)$	=	$\int_0^\infty e^{-\theta x} P_{i,j}(x)dx$

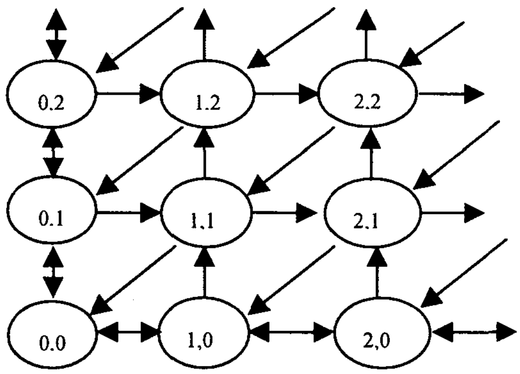

Fig.2 The state transition diagram of system

3. SYSTEM ANALYSIS

Using the above notations, we can derive the steady-state equations (1)~(5)(refer Appendix A). Let us define the following generating functions:

$$\overline{P^*}(z,w,\theta) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} P^*_{i,j}(\theta) z^i w^j$$

$$\overline{P}(z,w,0) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} P_{i,j}(0) z^i w^j$$

$$P^*(i,w,\theta) = \sum_{j=0}^{\infty} P^*_{i,j}(\theta) w^j$$

$$P(i,w,0) = \sum_{j=0}^{\infty} P_{i,j}(0) w^j,\ i \ge 0$$

$$P(z,j,0) = \sum_{i=0}^{\infty} P_{i,j}(0) z^i,\ j \ge 0$$

Taking LST to the equations (1) ~ (5) and applying the above generating functions, we get

$$(\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w - \theta)\overline{P^*}(z,w,\theta) =$$

$$(\tfrac{S^*(\theta)}{zw} - 1)\overline{P}(z,w,0) - \tfrac{S^*(\theta)}{zw}[(1-z)P(0,w,0) + (1-w)P(z,0,0) - (1-z)(1-w)P_{0,0}(0)] \quad (6)$$

From equation (6), we can derive the joint transform, $\overline{P^*}(z,w,\theta)$, of the number of customers in storage and retrieval queue and the remaining service time. From $\overline{P^*}(z,w,\theta)$ and letting θ=0, we obtain the probability generating function (pgf) of the joint distribution for the number of customers in storage and retrieval queues as follow:

$$\overline{P^*}(z,w,0) = \frac{S^*(\lambda_1(1-z)+\lambda_2(1-w))-1}{(\lambda_1(1-z)+\lambda_2(1-w))(zw-S^*((\lambda_1(1-z)+\lambda_2(1-w)))}$$
$$\cdot[(1-z)P(0,w,0) + (1-w)P(z,0,0) + (1-z)(1-w) P_{0,0}(0) \quad (7)$$

Then, we get the pgfs for the number of customers of storage queue and for the number of customers of retrieval queue from equation (7) and by letting w=1, θ=0 and z=1, θ=0, respectively.

$$\overline{P^*}(z,1,0) = \frac{S^*(\lambda_1(1-z))-1}{\lambda_1(z - S^*(\lambda_1(1-z)))} \cdot P(0,1,0) \quad (8)$$

$$\overline{P^*}(1,w,0) = \frac{S^*(\lambda_2(1-w))-1}{\lambda_2(w - S^*(\lambda_2(1-w)))} \cdot P(1,0,0) \quad (9)$$

By definition, $P(1,0,0) = \sum P_{i,0}(0)$, $P(0,1,0) = \sum P_{0,j}(0)$ and from the equation (8) and (9) we see

$$\overline{P^*}(z,1,0)\big|_{z=1} = 1 - Q_0,\ \overline{P^*}(1,w,0)\big|_{w=1} = 1 - Q_0 .$$

From the definition and these results, we obtain,

$$P(0,1,0) = \tfrac{1-\lambda_1 E(S)}{E(S)}(1-Q_0)\ ,\ P(1,0,0) = \tfrac{1-\lambda_2 E(S)}{E(S)}(1-Q_0)$$

which are constant values. Then, we get the distribution of the number of customers in both queues by substituting these values to the equations (8) ~ (9).

$$\overline{P^*}(z,1,0) = \frac{S^*(\lambda_1(1-z))-1}{\lambda_1(z - S^*(\lambda_1(1-z)))}\cdot[\frac{1-\lambda_1 E(S)}{E(S)}]\cdot(1-Q_0) \quad (10)$$

$$\overline{P^*}(1,w,0) = \frac{S^*(\lambda_2(1-w))-1}{\lambda_2(w - S^*(\lambda_2(1-w)))}\cdot[\frac{1-\lambda_2 E(S)}{E(S)}]\cdot(1-Q_0) \quad (11)$$

In order to get the mean number of waiting customers in storage and retrieval queues, we need Q_0, which can be obtained by the following theorem.

Theorem 3.1: Q_0 = *Pr (sever idle)*. Then, Q_0 satisfies the formula $Q_0 = (1-\rho_1)(1-\rho_2)$, where $\rho_i = \lambda_i/\mu$, i = 1, 2.

Proof: Put the virtual servers 1 and 2 before the storage and retrieval queues and assume each virtual servers independently serve storage commands and retrieval commands, respectively. The distributions of the service times of virtual servers are equal to those of the original server. Then, because these two queueing systems satisfy work-conservation law, the probabilities that each virtual servers are busy are $\rho_i = \lambda_i/\mu$, i=1, 2. Put, also, the switch box between the virtual servers 1 and 2. Then, let us suppose this switch is off only when the two virtual servers are all idle. Then, the following is obtained.

$$\Pr(switch\,on) = \Pr(server\,busy) = 1 - Q_0 .$$

Now, we have

Pr(switch on) = Pr (switch on | two virtual servers are all busy) Pr (two virtual servers are all busy)
+ Pr(switch on | only one of two virtual servers is busy)Pr(only one of two virtual servers is busy)
+ Pr(switch on | two virtual servers are all idle) Pr(two virtual servers are all idle)

$$= 1\cdot\rho_1\rho_2 + \{\rho_1(1-\rho_2) + \rho_2(1-\rho_1)\} + 0\cdot(1-\rho_1)(1-\rho_2)$$
$$= \rho_1 + \rho_2 - \rho_1\rho_2,$$

from which the theorem follows. {

Let $E(N_S)$ and $E(N_R)$ be the expected numbers of waiting customers in storage and retrieval queues, respectively. Then, from the equations (10), (11) and above theorem, $E(N_S)$ and $E(N_R)$ are calculated as follow:

$$E(N_S) = \frac{\lambda_1^2 E(S^2)}{2\lambda_1 E(S)(1 - \lambda_1 E(S))} \cdot (1 - Q_0) \quad (12)$$

$$E(N_R) = \frac{\lambda_2^2 E(S^2)}{2\lambda_2 E(S)(1 - \lambda_2 E(S))} \cdot (1 - Q_0) \quad (13)$$

We also can get the waiting time in system for each customer by the Little's formula, $L=\lambda W$.

4. CONCLUDING REMARKS

In this paper, we presented the probability distribution, the mean number of waiting customers in storage and retrieval queues and waiting time for each customer in system. We tested the proposed analytical model under the consideration of two cases that either each arrival rates of storage and retrieval commands are equal, $\lambda_1=\lambda_2$, or not, $\lambda_1\neq\lambda_2$ with three types of distribution for the service times: Exponential, Erlang and Discrete. Form the experiments, we found that the performance is accurate at a high traffic load and it is robust to the distribution of service times. The largest drawbacks in existing static analysis could not reflect the dynamic aspects of system and could lead a larger deviation from the corresponding simulation results.

However, the proposed model resolved these weakness and it could be used for quick estimation of the performance of an AS/RS and the design of an AS/RS.

References

1. Hausman, W H., Schwarz L B. and Graves S C. "Optimal storage assignment in automatic warehousing systems," *Management Science*, 1976, 22 (6), pp. 629-638

2. Bozer S L. and White J A. "Travel-time models for automated storage/retrieval systems." *IIE Transactions* 1984, 16 (4), pp. 329-338

3. Egbelu P J. and Wu C-T. "A comparison of dwell point rules in an automated storage/retrieval system," *Int'l J. Prod. Res.*, 1993, 31 (11), pp. 2515-2530

4. Han M H., McGinnis L F., Shieh J S. and White J A "Sequencing retrievals from automated S/R systems," *IIE Transactions*, 1987, 19 (1), pp. 56-66

5. Bozer Y A., Schorn E C. and Sharp G P "Geometric approaches to solve the Chebyshev traveling salesman problem," *IIE Transactions*, 1990, 22 (3), PP. 238-254

6. Bozer Y A. and White J A."Design and performance models for end-of-aisle order picking systems," *Management Science*, 1990, 36 (7), pp. 852-866

7. Lee H F. "Performance analysis for automated storage and retrieval systems," *IIE Transactions*, 1997, 29, pp. 15-28

Appendix A: Steady-state Equations (1) ~ (5)

$$0 = -(\lambda_1 + \lambda_2)Q_0 + P_{0,0}(0) \quad (1)$$

$$-\frac{d}{dx}P_{0,0}(x) = -(\lambda_1 + \lambda_2)P_{0,0}(x) + (\lambda_1 + \lambda_2)Q_0 s(x) + (P_{1,0}(0) + P_{0,1}(0) + P_{1,1}(0))\, s(x) \quad (2)$$

$$-\frac{d}{dx}P_{i,0}(x) = -(\lambda_1 + \lambda_2)P_{i,0}(x) + \lambda_1 P_{i-1,0}(x) + (P_{i+1,0}(0) + P_{i+1,1}(0))\, s(x),\ i \geq 1 \quad (3)$$

$$-\frac{d}{dx}P_{0,j}(x) = -(\lambda_1 + \lambda_2)P_{0,j}(x) + \lambda_2 P_{0,j-1}(x) + (P_{0,j+1}(0) + P_{1,j+1}(0))\, s(x),\ j \geq 1 \quad (4)$$

$$-\frac{d}{dx}P_{i,j}(x) = -(\lambda_1 + \lambda_2)P_{i,j}(x) + \lambda_1 P_{i-1,j}(x) + \lambda_2 P_{i,j-1}(x) + P_{i+1,j+1}(0)s(x),\ i, j \geq 1 \quad (5)$$