韓國經營科學會誌
第23卷 第3號
1998年 9月

Storage Assignment Policies in Automated Storage/Retrieval Systems*

Jeongseob Kim**

Abstract

Automated Storage and Retrieval Systems (AS/RSs) are an important facility for modern material management. The expected benefits of these capital-intensive facilities are gained when their control policies and their physical design parameters are determined simultaneously. In this paper we present several analytical models that capture the impact of the storage assignment policy and of the rack design on the expected storage and retrieval times. Sequential and interleaved service modes are considered for sequencing the storage and retrieval requests. We further investigate the impact of the rack structure on the relative performance of the following storage assignment policies: closest open location (random), full turnover-based policy, and class-based. Our analysis clearly indicates that significant savings in crane travel time are realized when implementing full turnover-based policy, rather than random. These savings become more and more pronounced as the profile of the storage racks approaches the square-in-time shape. Furthermore, it is shown that a class-based policy, with a small number of storage classes, will capture most of these savings and be easier to manage in practice.

1. Introduction

Automated Storage and Retrieval Systems (AS/RSs) are common in numerous manufacturing and distribution centers. These systems perform material storage and retrieval functions in a fully or semi-automated manner under a control of real-time computer systems. A typical system is composed of multiple parallel aisles of storage racks, a storage/retrieval crane for each aisle, and an input/output (I/O) pickup and deposit station. The crane has horizontal and vertical drives which op-

* This research was supported by the Taegu University Research Grant. 1997.
** Department of Business Administration. Taegu University

erate simultaneously in order to reduce the travel time. As a result the distance between any two points is measured by the Chebyshev (or the l_∞-norm) metric. In a dual-command operational mode each crane cycle begins with the crane at the I/O point: it picks up a load, travels to the designated storage location, deposits the load, travels empty (interleaves) to the other designated retrieval location, retrieves another load, and then travels back to the I/O point and deposits the load there. In a single command mode there is no interleaving and the crane returns to the I/O point following each storage or retrieval task. Several warehousing and manufacturing applications of AS/RSs are discussed by Seidmann (1988) and Sule (1988). Information storage and retrieval applications in two-dimensional mass storage systems are detailed by Wong (1980) and others.

There are many benefits of AS/RSs, such as reduced labor costs, high floor and cube space utilization, improved flow, and inventory controls. Unlike traditional warehouses, where the data record of location and inventory levels are collected manually, in AS/RSs such controls are maintained by the supervisory computer system. These controls account for significant savings in the inventory holding costs. In addition, using high-density high-rise racks reduce travel times and floor space. These AS/RS systems, however, require a high initial investment, and a thorough analysis to determine their economic viability. Realizing the benefits of these systems depends on specifying the appropriate configuration and on the development of effective management policies for operating these systems. Analytical models, in particular closed form expressions for determining the systems performance, are extremely useful for the rapid exploration of various operational policies and of potential economic savings.

A number of alternative operational policies for assignment of items to storage locations are studied in the literature [Ashayeri et al. (1985), Malmborg and Krishnakumar (1989), and White (1983)]. These studies also discuss how the storage policy use depends on the nature of the warehouse mission. For example, random storage is the preferred approach for unit-load (palletized) AS/RSs storing finished goods with a variable storage mix. On the other hand, full-turnover allocations are commonly used in mini-load systems where storage containers with frequently kitted components can be given preferential storage assignments. Using pairwise switching arguments Hausman et al. (1976) show that the expected one-way crane travel time is minimized under this policy. The studies by Hausman et al. (1976) and by Graves et al. (1977) present closed form expressions for both single and dual command cycle times under several storage allocation policies. A power function is used by these authors to describe the cumulative percentage demand for the various items in storage. Using this one dimensional probability distribution function limits the applicability of their models to square-in-time (SIT) shapes. Bozer and White (1984) deal with non-square-in-time (NSIT) shape warehouse which includes SIT as a special case. They use an order statistic method to derive the expected interleaving time assuming uniform (or random) storage allocation policy. Many storage facilities use either class-based [Rosenblatt and Eynan (1989)] or full-turnover based policy to reduce the expected travel times of the crane.

In this paper we extend and unify many of the earlier results cited above. We present exact closed form expression for computing the throughput rate of an AS/RS with the following storage allocation policies: closest open location (random), full-turnover, and class-based. Our analysis considers the more general NSIT case with both single command and dual command crane movements.

The significance of these new results is in providing a general framework for evaluating the relative benefits of the various storage policies as a function of the existing (or of the proposed) warehouse topology. For example, we found that random storage becomes less and less desirable as the aisle length increases, while the relative benefits of using class-based storage allocation policy are significantly less sensitive to changes in the aisle length. Several practical implications of these results for the management and design of modern AS/RSs are also discussed.

Section 2 presents the assumptions and parameters use in modeling the crane cycle time. Section 3 derives the closed form expressions for the expected crane cycle times under various operating policies. Several management and design issues in AS/RSs are dealt with in Section 4. Section 5 concludes the paper.

2. The Model

A storage rack is said to be NSIT if the time required to traverse the entire length of the rack (front to back) is not equal to the time required to traverse the entire height of the rack (bottom to top). Assume a rack with C columns and R rows and each pallet storage opening of width w and height h. Let $(x_j, y_j), x_j \in \{1, 2, \ldots, C\}, y_j \in \{1, 2, \ldots, R\}$, denote the coordinate in the rack of a storage opening j, $j=1,2,\ldots,N(=RC)$, where $(x_j, y_j) = (1, 1)$ represents the corner location adjacent to the I/O point. Let us denote the crane travel time from the I/O point to a storage opening j by z_{0j} and the travel time between two different storage openings i and j by z_{ij}. Then these times are expressed as

$$z_{oj} = \max\left\{\frac{w(x_j-1/2)}{V_x}, \frac{h(y_j-1/2)}{V_y}\right\} \quad (1)$$

$$z_{ij} = \max\left\{\frac{w\,|\,x_i-x_j\,|}{V_x}, \frac{h\,|\,y_i-y_j\,|}{V_y}\right\} \quad (2)$$

where, V_x and V_y denote the horizontal and the vertical speed of the crane. In (2) $z_{ij} = z_{ij}$ is assumed.

Without loss of generality we assume that the storage opening index j is assigned in such a way that $z_{oj} \leq z_{oj+1}$ holds and that the horizontal travel time is not less than the vertical time. If we assign the relative storage/retrieval frequency λ_j to each storage opening j, the relation $\lambda_j \geq \lambda_{j+1}$ represents the full-turnover storage assignment policy. Thus the expected cycle times of single and dual commands under full-turnover storage assignment policy in the (discrete) rack is given by

$$E_{disc}[SC] = 2\sum_{j=1}^{N} \lambda_j z_{oj} \tag{3}$$

$$E_{disc}[DC] = \sum_{j=1}^{N} \sum_{j=1, j\neq i}^{N} \lambda_i \frac{\lambda_j}{1-\lambda_i}(z_{oi} + z_{ij} + z_{oj}) \tag{4}$$

To simplify the notation used we normalize the horizontal and the vertical travel time into one unit of time and $b, b \leq 1$, unit of time, respectively. Doing so we transform the rack in physical length domain into a normalized time domain. Henceforth we discuss travel times on this normalized time domain.

We use a unit of travel time measurement to yield a storage space with unit width and height $b, b \leq 1$: the origin (0,0) is assumed to be the I/O point; the X axis and Y axis represent the horizontal and the vertical travel time, respectively. Since the pallet size is small compared with the rack, any point $(x, y) \in \{(u, v) | 0 < u \leq 1, 0 < v \leq b \leq 1\}$ is assumed to represent a location of a storage opening.

As shown in (1) and (2) the travel time from the I/O point to any storage opening depends on its horizontal and vertical locations. The coordinates for the locations of storage openings are not independent of each other since the crane travel time always dependent on the largest of the horizontal and the vertical travel times. Since this study aims at full-turnover storage allocation policies this immediately gives rise to a two-dimensional travel time density function. Following Hausman et al. (1976) we employed a parameterized fast-decaying power function to capture the item traffic intensity over the (1, b) rectangle.

We define in the above mentioned rectangle a random vector $X = (x, y)$. The probability density function of X is given by

$$f(x, y) = \begin{cases} f_1(y) \equiv ae^{\lambda x} & \text{if } 0 \leq y \leq x \leq 1 \\ f_2(y) \equiv ae^{\lambda x} & \text{if } 0 \leq x \leq y \leq b \end{cases}$$

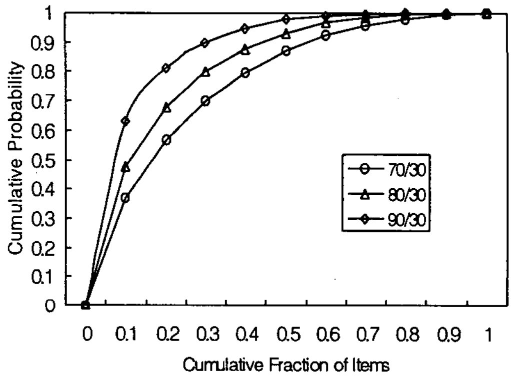
where

$$a = \begin{cases} [2(1-e^{-\lambda b})/\lambda - b(e^{-\lambda} + e^{-\lambda b})/\lambda]^{-1} & \text{if } \lambda > 0 \\ 1/b & \text{if } \lambda = 0 \end{cases} \tag{6}$$

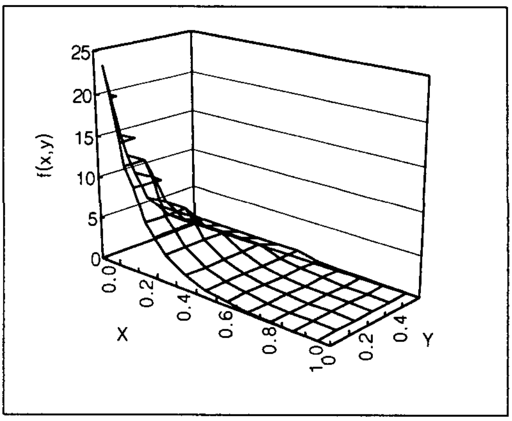
The probability mass of X has its highest value at $(0^+,0^+)$ and reduces as $X \to 1$ and $Y \to b$. This pdf is used in representing the traffic of items stored in the rack according to their relative demand, or turnover frequency. Under this storage policy the distance (in time) of each item from the I/O point is proportional to its relative demand. Hence, the fastest moving items are closest to the I/O point. The crane serving this area can simultaneously move both vertically and horizontally and at different speeds. The cumulative distribution function (cdf) is:

$$F(x,y)=\begin{cases} F_1(x,y) \equiv a[2/\lambda - 2e^{-\lambda y}/\lambda - ye^{-\lambda y} \\ \quad - ye^{-\lambda x}]/\lambda \text{ if } y \le x \le 1 \\ F_2(x,y) \equiv a[2/\lambda - 2e^{-\lambda x}/\lambda - xe^{-\lambda x} \\ \quad - xe^{-\lambda y}]/\lambda \text{ if } x \le y \le b \end{cases} \qquad (7)$$

When the items in storage are ranked according to their relative contribution to the total crane activities we get the well-known Pareto (or "ABC") curves. Figure 1 depicts the curves of three common Paretos when $b=0.5$. These curves go from (0%,0%) to (100%,100%) point through the β/γ point. This characterizes the percentage of the cumulative workload (β) attributable to the portion (γ) of the total item population. For instance, the 80/30 curve presents a case where 80% of the crane traffic is generated by 30% of the items. The pdf defined in (5) above with a given b can be fitted to this empirical data by adjusting the shape parameter λ. For example, when $b=0.5$, the value of λ corresponding to 80/30 Pareto is 6.507474. Figure 2 displays the pdf of the 80/30 ABC profile for $b=0.5$. Note that when $\lambda \to 0^+$ the above two-dimensional pdf



[Figure 1] The Cumulative Demand Distribution Function



[Figure 2] A Relative Traffic Intensity in an NSIT Rack with b=0.5, Full-turnover-based Storage Allocation Policy and an 80/30 Pareto reduces to the uniform distribution with density $1/b$. Further, when $b=1$ and $\lambda \to 0^+$, we get the special case of square-in-time (SIT) AS/RS with a random storage assignment policy.

In order to evaluate the crane travel time we define two new random variables, T and U, which represent the travel time from the I/O point to a point $X_1=(x_1, y_1)$ and the interleaving time between two random points $X_1=(x_1, y_1)$ and $X_2=(x_2, y_2)$ in the rectangle, respectively. This is gives us:

$$T=T(X_1)=\max\{X_1, Y_1\} \quad (8)$$

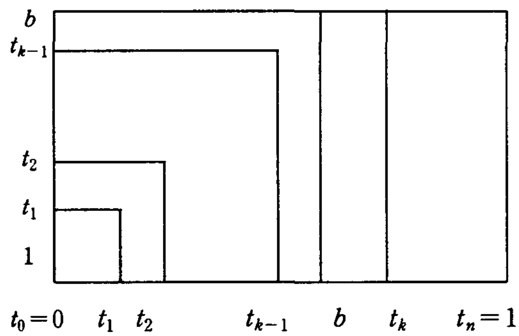
$$U=U(X_1, X_2)=\max\{|x_1-x_2|, |y_1-y_2|\} \quad (9)$$

We assume that two random points along the dual command cycle are independent and ignore the constant load/unload time of the crane [Hausman et al. (1976), Bozer and White (1984)]. Then we immediately obtain the following relations between the expected crane travel times of one round trip of the single command, $E[SC]$, and the dual command, $E[DC]$:

$$E[SC]=2E[T] \quad (10)$$

$$E[DC]=2E[T]+E[U]=E[SC]+E[U] \quad (11)$$

Another important storage assignment policy is the class-based one. This policy is commonly practiced in numerous warehouses when the actual turnover rate is not uniform. In this policy items are grouped based on their storage/retrieval request frequency and are then stored in corresponding segments of the warehouse as shown in Figure 3. The fastest moving items are classified as class 1 items and they are assigned to the segment nearest (in time) to the I/O point. It is assumed that the items within each class can be stored at any empty opening in the rack. This results in a uniform item storage distribution within each class.



[Figure 3] Multiclass Partition of the Rack

Class-based storage allocation policy is easier to manage in cases where the number of different units in storage is not constant. This is typical for a unit-load work in process (WIP) storage facility that interfaces between a production job shop and the regional distribution centers for finished goods. On the other hand, the full-turnover based storage allocation policy is more suitable for cases where the number of items in storage is constant. For example, when an AS/RS mini-load system is used to supply electronic components for a printed circuit board assembly line.

The distribution function $F(x, y)$ defined by (7) is used to represent the actual distribution of storage/retrieval frequency data. In Figure 3, t_i is the border point on the two axes which separates class i and class $i+1$ segments, for $i=1, 2, \ldots, n$, $t_0=0$, and $t_n=1$.

3. Expected Cycle Times

The following formulae specify the AS/RS cycle times under three different operating policies. These are derived in Appendix.[1)]

Formula 1. The expected cycle time of a single command under full-turnover storage allocation policy is

$$E\,[SC]=$$

$$\frac{2}{\lambda}\left(\frac{4e^{\lambda b}-(\lambda+\lambda^2)be^{\lambda(b-1)}-b^2\lambda^2-3b\lambda-4}{2e^{\lambda b}-b\lambda e^{\lambda(b-1)}-b\lambda-2}\right) \tag{12}$$

Notice that replacing b by 1 in (12) gives the special case of SIT [Kim and Seidmann (1990)],

$$E_{SIT}[SC]=\frac{4e^{\lambda}-2\lambda^2-4\lambda-4}{\lambda(e^{\lambda}-\lambda-1)}. \tag{13}$$

We also find that

$$\lim_{\lambda\to 0^+} E[SC]=\frac{1}{3}b^2+1, \tag{14}$$

which conforms with the earlier result given by Bozer and White (1984) for the expected cycle time of a single command of NSIT under random storage assignment policy. Furthermore the result of Hausman et al. (1976), for the expected cycle time of a single command SIT under random storage allocation policy, is derived by:

$$\lim_{b\to 1}\lim_{\lambda\to 0^+} E[SC]=\frac{4}{3}. \tag{15}$$

Formula 2. Let k be such that $t_{k-1} < b < t_k$. Then the expected cycle time of single command under n-class-based storage assignment policy is

$$E_{\text{n-class}}[SC]=\frac{4}{3}\sum_{i=1}^{k-1} p_i\left(\frac{t_i^3-t_{i-1}^3}{t_i^2-t_{i-1}^2}\right)$$
$$p_k\frac{b^3-4t_{k-1}^3+3bt_k^2}{3(bt_k-t_{k-1}^2)}$$
$$\sum_{i=k+1}^{n} p_i\cdot(t_{t-1}+t_i) \tag{16}$$

where $p_i = F(t_k, t_k)-F(t_{k-1}, t_{k-1})$.

Formula 2 generalizes the earlier expression given by Hausman et al. (1976) for two- and three-class-based policy. The closed form structure of (16) is helpful in determining the boundaries for the various storage classes and

1) All computations in this paper are performed using Mathematica™ and Symbolics™, which are very popular mathematical expert systems. The results from these two systems are cross-checked.

in determining the marginal effects of increasing the number of classes.

Formula 3. The expected cycle time of dual command under full-turnover storage allocation policy is given by:

(i) when $0 < b \le 1/2$:

$E[DC] =$

$$\begin{aligned}&2a\left[\frac{4}{\lambda^3} - \left(\frac{b^2}{\lambda} + \frac{3b}{\lambda^2} + \frac{4}{\lambda^3}\right)e^{-b\lambda} - \left(\frac{b}{\lambda} + \frac{b}{\lambda^2}\right)e^{-\lambda}\right] + \\ &a^2\left[-\frac{2e^{(b-2)\lambda}}{\lambda^5} + \left(\frac{2b^2 - b^3}{\lambda^2} + \frac{4b - 2b^2}{\lambda^3} - \frac{2b}{\lambda^4}\right)e^{-(1+b)\lambda}\right. \\ &-\left(\frac{4b^2}{\lambda^3} + \frac{6b}{\lambda^4} + \frac{7}{\lambda^5}\right)e^{-b\lambda} - \frac{64e^{-3b\lambda/2}}{3\lambda^5} \\ &+\left(\frac{b^3}{3\lambda^2} + \frac{2b^2}{\lambda^3} + \frac{8b}{\lambda^4} + \frac{21}{\lambda^5}\right)e^{-2b\lambda} - \frac{e^{-3b\lambda}}{3\lambda^5} \\ &\left.-\left(\frac{4b}{\lambda^3} - \frac{2b}{\lambda^4}\right)e^{-\lambda} + \left(\frac{b^3}{3\lambda^2} + \frac{2b}{\lambda^4} + \frac{2}{\lambda^5}\right)e^{-2\lambda} + \frac{23}{3\lambda^5}\right]\end{aligned} \tag{17}$$

(ii) when $1/2 \le b \le 1$:

$E[DC] =$

$$\begin{aligned}&2a\left[\frac{4}{\lambda^3} - \left(\frac{b^2}{\lambda} + \frac{3b}{\lambda^2} + \frac{4}{\lambda^3}\right)e^{-b\lambda} - \left(\frac{b}{\lambda} + \frac{b}{\lambda^2}\right)e^{-\lambda}\right] + \\ &a^2\left[\left(\frac{5b^3 - 6b^2 + 6b - 1}{3\lambda^2} + \frac{6b^2 - 4b + 2}{\lambda^3}\right.\right. \\ &\left.+\frac{18b - 10}{\lambda^4} + \frac{20}{\lambda^5}\right)e^{-(1+b)\lambda} - \frac{e^{(b-2)\lambda}}{\lambda^5} \\ &-\left(\frac{4b^2}{\lambda} + \frac{6b}{\lambda^4} + \frac{7}{\lambda^5}\right)e^{-b\lambda} - \frac{64e^{-3b\lambda/2}}{3\lambda^5} \\ &+\left(\frac{b^3}{3\lambda^2} + \frac{2b^2}{\lambda^3} + \frac{8b}{\lambda^4} + \frac{21}{\lambda^5}\right)e^{-2b\lambda} \\ &-\left(\frac{4b}{\lambda^3} - \frac{2b}{\lambda^4}\right)e^{-\lambda} - \frac{64e^{-3\lambda/2}}{3\lambda^5} \\ &\left.+\left(\frac{b^3}{3\lambda^2} + \frac{2b}{\lambda^4} + \frac{2}{\lambda^5}\right)e^{-2\lambda} + \frac{23}{3\lambda^5}\right]\end{aligned} \tag{18}$$

Replacing b by 1 in (18) gives the expression for the expected cycle time of dual command of SIT under full-turnover storage allocation policy as follows:

$E[DC] =$

$$\begin{aligned}&2a^2\left[\left(\frac{1}{\lambda^2} + \frac{3}{\lambda^3} + \frac{9}{\lambda^4} + \frac{43}{2\lambda^5}\right)e^{-\lambda 2}\right. \\ &-\frac{64}{3\lambda^5}e^{-3\lambda/2} - \left(\frac{4}{\lambda^3} + \frac{2}{\lambda^4} + \frac{4}{\lambda^5}\right)e^{-\lambda} \\ &\left.+\frac{23}{6\lambda^5}\right] + \frac{4e^{\lambda} - 2\lambda^2 - 4\lambda - 4}{\lambda(e^{\lambda} - \lambda - 1)}.\end{aligned} \tag{20}$$

Also notice that the expression given by Bozer and White (1984) for the cycle time of a dual command NSIT under random storage allocation policy is easily derived from our general expression:

$$\lim_{\lambda \to 0^+} E_{0 < b \le 1/2}[DC] = \lim_{\lambda \to 0^+} E_{1/2 < b \le 1}[DC]$$
$$= \frac{4}{3} + \frac{1}{2}b^2 - \frac{1}{30}b^3. \tag{21}$$

Furthermore, the expected cycle time of a dual command SIT under random storage allocation policy is given by

$$\lim_{b \to 1} \lim_{\lambda \to 0^+} E[DC] = 9/5. \tag{22}$$

This value conforms with the earlier result given by Graves et al. (1977) for this special case.

4. Warehouse Management and Design

In this section we consider some important issues in the management and design of AS/RSs. We begin by evaluating the relative merits of using a dual, rather than a single, command crane operating policy. To that end we first compare the expected cycle time under single and dual command modes of operation. Single command mode requires two round trips for one storage/retrieval: one round trip retrieving a pallet to the I/O point and the other to store it back on the rack. On the other hand, only one round trip with three legs is required in dual command mode. The relative saving of a dual command over a single command is given by, (see (10) and (11)),

$$S(\lambda,\ b)=\ \frac{2E[SC]-E[DC]}{2E[SC]}=\frac{1}{2}-\frac{1}{4}\cdot\frac{E[U]}{E[T]},\quad (0\le\lambda, 0<b\le 1). \tag{23}$$

where E[T] and E[U] are derived in (A-3), and (A-8, A-9). The following property gives the range of relative saving in the case of random storage allocation policy.

Property 1. In the case of the random storage allocation policy, (i.e., $\lambda=0.0$), the relative saving function $S(\lambda, b)$ is bounded by

$$39/120 \le S(\lambda, b) \le 40/120.$$

Proof: Inserting (14) and (21) into (23) we obtain

$$S(0,\ b)=\frac{1}{4}\left(1+\frac{b^3+5}{(5(b^2+3)}\right), \tag{24}$$

which is decreasing monotonically in b. Therefore, the two boundary values $S(0,1)$ = 13/40 and $S(0,0)=1/3$ complete the proof. □

Conjecture 1. For the most practical cases, the relative saving function $S(\lambda, b)$ is bounded on $0.1\le b\le 1$ by

$$0.26 \le S(\lambda,\ b) \le 0.34.$$

<Table 1> Expected cycle times for different b.

b	Random storage allocation policy			Full-turnover-based storage allocation policy								
				70/30			80/30			90/30		
	$E[SC]$	$E[DC]$	$S(\lambda,b)$	$E[SC]$	$E[DC]$	$S(\lambda,b)$	$E[SC]$	$E[DC]$	$S(\lambda,b)$	$E[SC]$	$E[DC]$	$S(\lambda,b)$
0.1	3.173	4.232	33.31	1.544	2.262	26.74	1.215	1.790	26.33	0.898	1.319	26.54
0.2	2.266	3.026	33.24	1.148	1.664	27.55	0.929	1.349	27.41	0.718	1.039	27.63
0.3	1.881	2.515	33.13	0.987	1.420	28.08	0.812	1.171	27.86	0.625	0.919	27.62
0.4	1.665	2.231	33.01	0.909	1.305	28.27	0.755	1.090	27.80	0.590	0.860	27.20
0.5	1.532	2.057	32.88	0.877	1.258	28.26	0.733	1.062	27.59	0.574	0.840	26.88
0.6	1.446	1.944	32.76	0.862	1.238	28.19	0.724	1.051	27.41	0.568	0.833	26.67
0.7	1.390	1.873	32.65	0.855	1.229	28.11	0.719	1.046	27.27	0.565	0.829	26.54
0.8	1.357	1.829	32.57	0.851	1.225	28.05	0.717	1.045	27.18	0.563	0.828	26.46
0.9	1.339	1.807	32.52	0.849	1.223	28.01	0.716	1.044	27.13	0.562	0.828	26.42
1.0	1.333	1.800	32.50	0.849	1.223	27.99	0.716	1.044	27.11	0.562	0.827	26.41

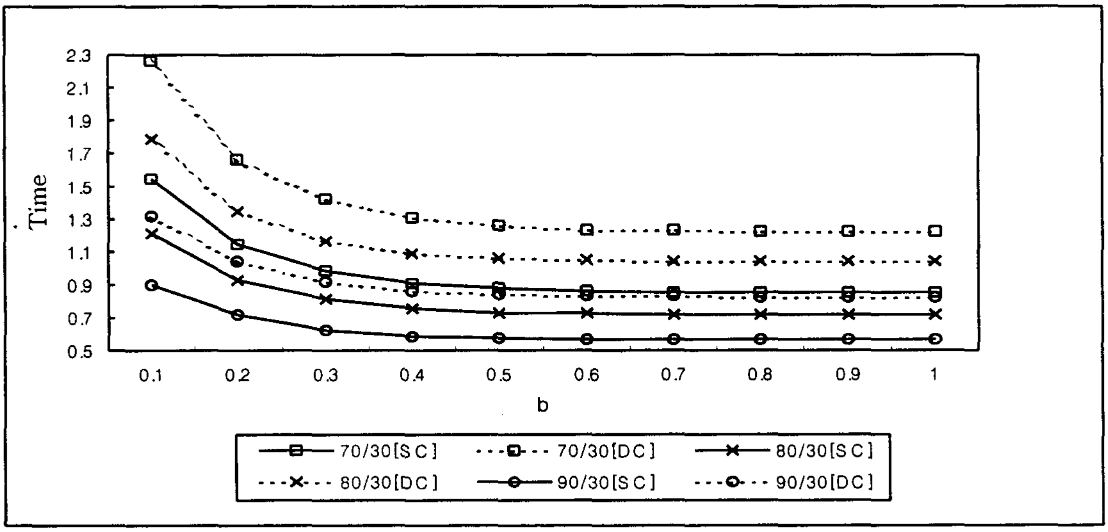
Remark: The horizontal travel time of SIT (b=1) is set to 1 unit of time. E[SC] and E[DC] are accordingly scaled up. Relative savings S(,b) are in percentages (%).

To verify this conjecture we found, by numerical search, that the relative saving function $S(\lambda, b)$ decreases, though not monotonically, converging to 26.04% as $b \to 1$ s and $\lambda \to \infty$.

Table 1 computes the values of this relative saving function for various values of b under random and full-turnover-based storage allocation policies. The normalized horizontal travel time under SIT ($b=1$) is set to 1 in Table 1 and all other cases are scaled up accordingly. It is clear that the relative savings of using a dual command over a single command are insensitive to the shape parameter b. For example, $S(\lambda, b)$ varies between 27.86% and 26.33% in the 80/30 case. Similarly, we observe in Table 1 that the relative saving from using dual rather than single command is practically constant in the case of random storage allocation policy. (See Property 1.)

Table 1 also presents the impact of changes in the shape parameter b on the expected cycle time under single and dual command modes. The figures in Table 1 show that the expected cycle times increase as the shape parameter b decreases. The relative increase in the expected cycle time is about the same for both command modes; it becomes less pronounced as the relative traffic distribution curve becomes steeper. This observation supports our conjecture that minimum cycle times occur under SIT. Figure 4 depicts the expected cycle times of some cases.

We next study the impact of the storage allocation policy on the expected crane cycle time. The results presented in Table 2 and Figure 5 imply that for smaller values of b,



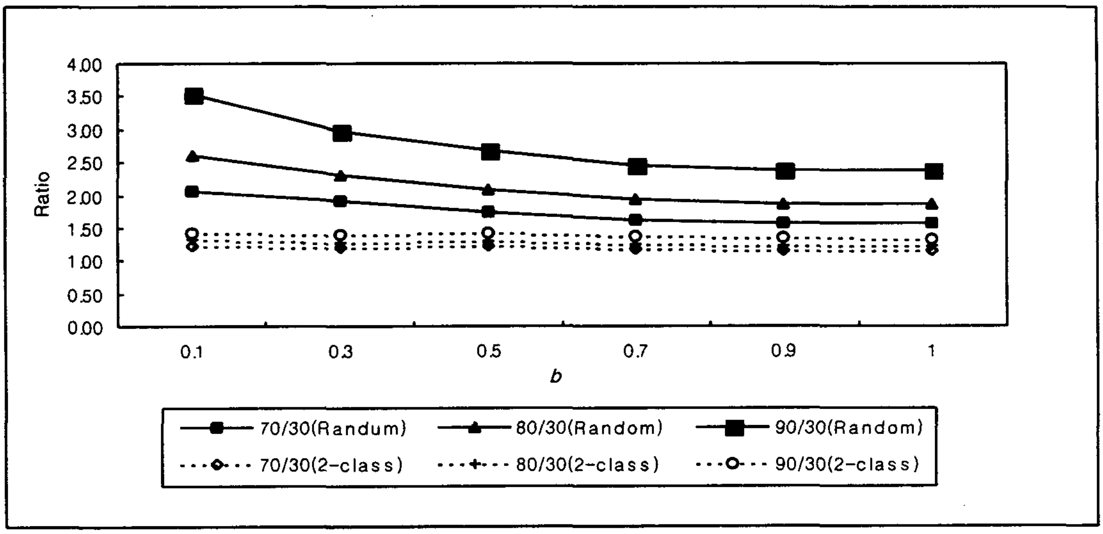
[Figure 4] Expected Cycle Times of Single and Dual Commands for Different Turnover Profiles e.g., 0.1, the cycle times for the random (or closest empty location first) policy can be two to three and a half times longer than those for the full-turnover-based policy. When *b* approaches 1, the relative performance of the random policy improves significantly but may still be one and a half to two and a half times as slow as under the full-turnover-based policy. The addition of constant crane load/ unload times (e.g., Bozer and White (1984)) still discloses a major performance degradation when using the random policy. Figure 5 also demonstrates the fact that the full-turnover-based policy becomes more desirable when the ABC turnover curve gets steeper. For example, in a single command case, moving from 70/30 to 90/30 curve increases the maximal value of the cycle time ratio from 2 to 3.5.

〈Table 2〉 Expected cycle times for single command under class-based policy for NSIT rack.

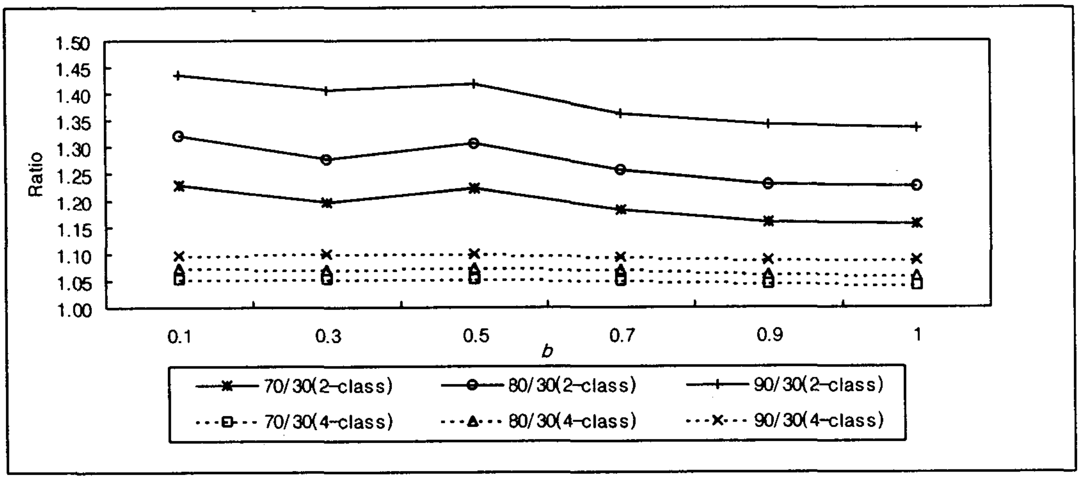
		Full	Random		2-Class		3-Class		4-Class	
ABC	*b*	$E_\infty[SC]$	$E_1[SC]$	*%Loss*	$E_2[SC]$	*%Loss*	$E_3[SC]$	*%Loss*	$E_4[SC]$	*%Loss*
70/30	0.1	1.544	3.173	105.55	1.897	22.87	1.694	9.73	1.627	5.39
	0.3	0.987	1.881	90.53	1.182	19.76	1.086	10.00	1.040	5.33
	0.5	0.877	1.532	74.77	1.072	22.23	0.957	9.20	0.923	5.30
	0.7	0.855	1.390	62.70	1.011	18.26	0.930	8.83	0.898	5.14
	0.9	0.849	1.339	57.59	0.984	15.88	0.912	7.41	0.886	4.30
	1.0	0.849	1.333	57.05	0.982	15.61	0.910	7.23	0.884	4.16
80/30	0.1	1.215	3.173	161.09	1.605	32.06	1.378	13.41	1.305	7.39
	0.3	0.812	1.881	131.64	1.037	27.79	0.919	13.24	0.870	7.23
	0.5	0.733	1.532	108.98	0.958	30.64	0.829	13.12	0.787	7.33
	0.7	0.719	1.390	93.28	0.903	25.54	0.807	12.12	0.770	7.07
	0.9	0.716	1.339	86.89	0.881	23.00	0.792	10.61	0.760	6.12
	1.0	0.716	1.333	86.18	0.879	22.70	0.791	10.42	0.759	5.97
90/30	0.1	0.898	3.173	253.51	1.289	43.64	1.055	17.55	0.985	9.71
	0.3	0.635	1.881	196.31	0.893	40.66	0.742	16.96	0.699	10.08
	0.5	0.575	1.532	166.65	0.815	41.90	0.682	18.78	0.632	10.00
	0.7	0.565	1.390	146.29	0.770	36.35	0.659	16.66	0.619	9.60
	0.9	0.562	1.339	138.11	0.753	33.89	0.648	15.30	0.612	8.86
	1.0	0.562	1.333	137.17	0.751	33.61	0.647	15.17	0.611	8.63

1. Cycles times are normalized into those of SIT (b=1).
2. The cycle times under class-based policies are computed after corresponding optimal boundaries are found by search method.
3. Figures under %Loss stand for the percentage *increase* of the expected crane cycle time under the relevant policy over full-turnover policy, $100(E_i[SC]/E_\infty[SC]-1)$.

Assuming that the number of aisles and the shape parameter (b) for each aisle are given the next issue discussed is the determination of the number of classes and the desired boundaries for a class-based storage system. Formula 2 can be incorporated with the simple line (or grid) search method in order to determine the optimal boundaries for any number of classes. Table 2 shows that two classes will yield about 60% to 89% of the



[Figure 5] a. Ratios of Single Command Cycle Times under Random and 2-Class-Based Policies to Full-Turnover-Based Policy for Different Turnover Profiles



[Figure 5] b. Ratios of Single Command Cycle Times under 2-and 4-Class-Based Policies to Full-Turnover-Based Policy for Different Turnover Profiles potential saving of the full-turnover- based policy and four classes will yield 90% or more. This table also indicates that increasing the number of classes results in diminishing returns on the expected relative savings. We found that these relative savings of class-based policy are highly sensitive to the steepness of the ABC curve. On the other hand, changes in the shape parameter b have a small effect on the potential savings for a given number of classes. For instance, in the 3-class, 90/30 case, the percentage difference varies between 15.17% and 18.78%. This is an important observation, meaning that changes in the number of classes, during the design phase or when expanding an existing facility, need not change the prevailing rack profile (b). Figure 5 plots the ratios of the expected cycle times of single command under random and class-based storage allocation policies to that of full-turnover-based policy.

Considering the AS/RS design problem, we assume that the demand distribution function, the overall storage area and the horizontal speed and the vertical speed of the crane are given. The designer has to determine the number of storage racks along with the physical dimensions and the shape parameter b for each rack. While an SIT rack leads to the minimal cycle time it may not be the most economical layout. Physical constraints on the warehouse height along with the nature of the construction cost function may mandate a rectangular rack with $b < 1$. Moreover, increasing the number of aisles increases the rack and crane costs but decreases the expected retrieval time per item.

5. Conclusions

In this paper, we have focused on storage assignment policies in AS/RSs and on their interaction effects with the crane operating modes and the shape of the warehouse. In particular, we investigated the following commonly used storage allocation policies: closest open location first (random), full-turnover-based and class-based. Using an NSIT rack profile, and a more general demand distribution function, we could extend and unify several of the earlier key studies. The practical significance of this analysis is in providing a closed form expression for simultaneously reviewing several managerial and design parameters.

Several general conclusions can be drawn from our study. First, a few alternative control measures are available in order to reduce the expected crane cycle times. Systems using single command mode can be improved by adopting a dual command mode. The savings in such a case may vary from 26% to 33% in most practical cases. We found that these savings are sensitive to the steepness of the ABC curve and insensitive to the shape parameter of the rack, *b*. Furthermore, the system throughput is increased by moving from a random to a full-turnover-based storage allocation policy. Our results show that the cycle times for the random policy are one and a half to three and a half times as long as for the full-turnover-based policy. The relative performance of the random policy tends to deteriorate as the rack profile moves away from SIT, and with the increase in the steepness of the ABC curve. It is clear that class based storage allocation is easier to implement than full-turnover, and with four classes it captures 90% or more of the potential savings of the latter policy.

The impact of changes in the demand distribution function has also been studied. The results seem to indicate that under the full-turnover policy the expected cycle times (for both single and dual command) diminish with the increase in the steepness of this curve. The increase in the steepness of this curve also reduces the relative savings of using a class based instead of the full turnover policy.

A stationary demand pattern has been considered by us in this study. We used this assumption since we were primarily interested in the long-run average behavior of the AS/RS. Other cases may also be important as the relative market demands for different products fluctuate over time. Similarly, additional research is needed in fine tuning our results to handle those cases where work schedules vary along the day with changing priorities for the storage and retrieval tasks, or where the order picking policy results in highly auto-correlated demands on the same crane routes.

Appendix

A. Derivation of Formula 1. [The expected cycle time of a single command under full-turnover storage allocation policy]

The distribution of one way travel time to a location in the rectangle from the origin is computed by

$$G_T(t) = \Pr\{X \le t, Y \le t\} = F(t, t)$$
$$= \begin{cases} F_2(t, t) = & a[2/\lambda - 2e^{-\lambda t}/\lambda] \\ & /\lambda \quad \text{if } 0 < t \le b \\ F_1(t, b) = & a[2/\lambda - 2e^{-\lambda b}/\lambda - be^{-\lambda b} - be^{\lambda t}] \\ & /\lambda \quad \text{if } b < t \le 1] \end{cases} \tag{A-1}$$

Taking derivative with respect to t, we get the pdf

$$g_T(t) = \frac{dG_t(t)}{dt} = \begin{cases} 2ate^{-\lambda t} & \text{if } 0 < t \le b \\ abe^{-\lambda t} & \text{if } b < t \le 1 \end{cases} \tag{A-2}$$

Thereby, denoting t as a realization of T, we get

$$
\begin{aligned}
&E[T] \\
&= \int_0^b 2at^2 e^{-\lambda t}dt + \int_b^1 abte^{-\lambda t}\,dt \\
&= a\left[\frac{4}{\lambda^3} - \left(\frac{b^2}{\lambda} + \frac{3b}{\lambda^2} + \frac{4}{\lambda^3}\right)e^{-b\lambda} - \left(\frac{b}{\lambda} + \frac{b}{\lambda^2}\right)e^{-\lambda}\right] \\
&= \frac{1}{\lambda}\left(\frac{4e^{\lambda b} - (\lambda + \lambda^2)be^{\lambda(b-1)} - b^2\lambda^2 - 3b\lambda - 4}{2e^{\lambda b} - b\lambda e^{\lambda(b-1)} - b\lambda - 2}\right).
\end{aligned} \tag{A-3}
$$

Hence we obtain Formula 1 from (10).□

B. Derivation of Formula 2. [The expected cycle time of a single command under class-based storage allocation policy]

The distribution function $F(x, y)$ defined in (7) is used to approximate the actual distribution of storage/retrieval frequency data. Let t_i be the border point on the x-axis (and the y-axis) which separates class i and class $i+1$ segments, for $i=1, \ldots, n$. $t_o=0$, and $t_n=1$. Let k be such that $t_{k-1} < b \le t_k$ and $h_i(x)$ be the pdf of a storage coordinates vector X given that it is in class i. Then the assumption on uniform distribution within each class gives us:

$$
h_i(x) = \begin{cases} \dfrac{1}{t_i^2 - t_{i-1}^2}, & i=1,\ldots,k-1 \\ \dfrac{1}{bt_i - t_{t-1}^2}, & i=k \\ \dfrac{1}{b(t_t - t_{t-1})}, & i=k+1,\ldots,n. \end{cases} \tag{A-4}
$$

For each class i, C_i the conditional expected cycle time of a single command is given by

$$
\begin{aligned}
&E[SC \mid X \in C_i] \\
&= 2\int_{X\in C_i} t(x)\,h_i(x)\,dx \\
&= \begin{cases} \dfrac{1}{t_i^2 - t_{i-1}^2}\left(\int_{t_{i-1}}^{t_i}\int_0^x xdydx + \int_{t_{i-1}}^{t_i}\int_0^y ydxdy\right), \quad i=1,\ldots,k-1 \\ \dfrac{2}{bt_i - t_{t-1}^2}\left(\int_{t_{i-1}}^{b}\int_0^y ydxdy + \int_{t_{i-1}}^{b}\int_y^{t_k} xdxdy \; \int_0^{t_{i-1}} + \int_{t_{k-1}}^{t_k} xdxdy\right), \quad i=k \\ \dfrac{2}{b(t_i - t_{i-1})}\int_0^b\int_{t_{k-1}}^{t_k} xdxdy, \quad i=k+1,\ldots,n \end{cases} \\
&= \begin{cases} \dfrac{4t_i^3 - t_{i-1}^3}{3t_i^2 - t_{i-1}^2}, & i=\ldots,k-1 \\ \dfrac{b^3 - 4t_{i-1}^3 + 3bt^2}{3(bt_i - t_{i-1}^2)}, & i=k \\ t_{i-1} + t_i, & i=k+1,\ldots,n. \end{cases}
\end{aligned} \tag{A-5}
$$

Hence the expected cycle time of a single command under n-class-based storage assignment policy is given as Formula 2 by the relationship

$$E[SC] = \sum_{i=1}^{n} E[SC] \mid X \in C_i \Pr\{X \in C_i\}. \;\square$$

C. Formula 3 [The expected cycle time of dual command under full-turnover storage allocation policy]

The first line of the right side of expressions (17) and (18) is $E[SC]$. Therefore, it suffices to show that the rest is the expected interleaving time, i.e., $E[U]$. Our procedure for deriving $E[U]$ is as follows:

Step 1. Find $G_{T,U}(t, u) = \Pr\{T \le t, U \le u\}$, the joint cdf of T, the one way travel time from the origin to a point in an NSIT rack, and U, the interleaving time of the dual command, under full-turnover-based storage allocation.

Step 2. By taking derivative of $G_{T,U}(t, u)$, find the corresponding joint pdf $g_{T,U}(t, u)$

Step 3. Find the marginal distribution $g_U(u)$ of the interleaving time U.

Step 4. Lastly, compute the desired expectation $E[U]$ by integrating $u g_U(u)$ over [0,1].

Since, once Step 1 is done, the other three steps are rather straightforward, it suffices to explain Step 1. However, we do not derive $g_{T,U}(t, u)$ completely here to save space. We now sketch the process of deriving $g_{T,U}(t, u)$

By probabilistic reflection we get the following:

$$
\begin{aligned}
&G_{T,U}(t,u)\\
&= \Pr\{T \le t, U \le u\}\\
&= \Pr\{X_1 \le t, Y_1 \le t, \ |X_1 - X_2| \le u, \ |Y_1 - Y_2| \le u\}\\
&= \Pr\{X_1 \le t, Y_1 \le t, X_1 - u \le X_2 \le X_1 + u, Y_1 - u, \le Y_2 \le Y_1 + u\}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{x \le t} \int_{y \le t} \Pr\{x - u \le X_2 \le x + u, y - u \le Y_2 \le y + u\} f(x,y)\,dy\,dx\\
&= \int_{x \le t} \int_{y \le t} [\, F(x+u, y+u) - F(x+u, y-u) - F(x-u, y+u) + F(x-u, y-u)] f(x,y)\,dy\,dx
\end{aligned}
\tag{A-6}
$$

That we have two different expressions of each of $f(x,y)$ and $F(x,y)$ and that the coordinates T and U must lie on a unit square require careful partitioning of the unit square to perform the algebra for (A-6). Let us take the first term, for example, to feel a flavor of partitioning involved.

$$
\int_{x \le t} \int_{y \le t} [F(x+u, Y+u)]\, f(x,y)\,dy\,dx
$$

$$
= \begin{cases}
\text{(i) for } t + u \le 1, \\
\int_0^t \left(\int_0^x F_1(x+u, y+u) f_1(x)\,dy + \int_x^t F_2(x+u, y+u) f_2(y)\,dy \right) dx; \\
\text{(ii) for } 1 \le + u \le 2, \\
\int_0^{1-u} \left(\int_0^x F_1(x+u, y+u) f_1(x)\,dy + \int_x^{1-u} F_2(x+u, y+u) f_2(y)\,dy + \int_{1-u}^t F_2(x+u, y+u) f_2(y)\,dy \right) dx \\
+ \int_{1-u}^t \left(\int_0^{1-u} F_1(1, y+u) f_1(x)\,dy + \int_{1-u}^x 1 f_2(y)\,dy + \int_x^t 1 f_2(y)\,dy \right) dx
\end{cases}
\tag{A-7}
$$

By similar probabilistic reflection and partitioning, we can complete Step 1 and thereby the other steps, too.

After considerable amount of algebra, we get the expected interleaving time E[U] as follows:

(i) when $0 < b < = 1/2$:

$$E[U] = a^2 \Big[\frac{-2e^{(b-2)\lambda}}{\lambda^5} + \Big(\frac{2b^2 - b^3}{\lambda^2} + \frac{4b - 2b^2}{\lambda^3} - \frac{2b}{\lambda^4} \Big) e^{-(1+b)\lambda} - \Big(\frac{4b^2}{\lambda^3} + \frac{6b}{\lambda^4} + \frac{7}{\lambda^5} \Big) e^{-b\lambda} - \frac{64e^{-3b\lambda/2}}{3\lambda^5} + \Big(\frac{b^3}{3\lambda^2} + \frac{2b^2}{\lambda^3} + \frac{8b}{\lambda^4} + \frac{21}{\lambda^5} \Big) e^{-2b\lambda} - \frac{e^{-3b\lambda}}{3\lambda^5} - \Big(\frac{4b}{\lambda^3} - \frac{2b}{\lambda^4} \Big) e^{-\lambda} + \Big(\frac{b^3}{3\lambda^2} + \frac{2b}{\lambda^4} + \frac{2}{\lambda^5} \Big) e^{-2\lambda} + \frac{23}{3\lambda^5} \Big] \quad \text{(A-8)}$$

(ii) when $1/2 \leq b \leq 1$:

$$E[U] = a^2 \Big[\Big(\frac{5b^3 - 6b^2 + 6b - 1}{3\lambda^2} + \frac{6b^2 - 4b + 2}{\lambda^3} + \frac{18b - 10}{\lambda^4} + \frac{20}{\lambda^5} \Big) e^{-(1+b)\lambda} - \frac{e^{(b-2)\lambda}}{\lambda^5} - \Big(\frac{4b^2}{\lambda} + \frac{6b}{\lambda^4} + \frac{7}{\lambda^5} \Big) e^{-b\lambda} - \frac{64e^{-3b\lambda/2}}{3\lambda^5} + \Big(\frac{b^3}{3\lambda^2} + \frac{2b^2}{\lambda^3} + \frac{8b}{\lambda^4} + \frac{21}{\lambda^5} \Big) e^{-2b\lambda} - \Big(\frac{4b}{\lambda^3} - \frac{2b}{\lambda^4} \Big) e^{-\lambda} - \frac{64e^{-3\lambda/2}}{3\lambda^5} + \Big(\frac{b^3}{3\lambda^2} + \frac{2b}{\lambda^4} + \frac{2}{\lambda^5} \Big) e^{-2\lambda} + \frac{23}{3\lambda^5} \Big] \quad \text{(A-9)}$$

Notice that these two are expected interleaving time part of equations (17) and (18), respectively.

REFERENCE

[1] Ashayeri, J. and Gelders, L. F., "Warehouse Design Optimization," *European Journal of Operational Research*, Vol.21 (1985), pp.285-294.

[2] Bozer, Y. A. and White, J. A., "Travel-Time Models for Automated Storage/Retrieval Systems," *IIE Transactions*, Vol. 16, No.4(1984), pp.329-338.

[3] Graves, S. C., Hausman, W. H., and Schwarz, L. B., "Storage-Retrieval Interleaving in Automatic Warehousing Systems," *Management Science*, Vol.23, No.9 (1977), pp.935-945.

[4] Hausman, W. H., Schwarz, L. B., and Graves, S.C., "Optimal Storage Assignment in Automatic Warehousing Systems," *Management Science*, Vol.22, No.6(1976), pp.629-638.

[5] Kim, J. and Seidmann, A., "A Framework for the Exact Evaluation of Expected Cycle Times in Automated Storage Systems with Full-Turnover Item Allocation and Random Service Requests," *Computers and Industrial Engineering*, Vol.18, No.4(1990), pp.601-612.

[6] Malmborg, C. J. and Krishnakumar, B., "Optimal Storage Assignment Policies for Multiaddress Warehousing Systems," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.19, No.1(1989), pp.197-204.

[7] Rosenblatt, M. J., "Issues in Design and

Control of AS/RS," *in Advanced Information Technology for Industrial Material Flow Systems*, S. Y. Nof and C. C. Moodie (Ed.), Springer Verlag, Berlin 1989.

[8] Rosenblatt, M. J. and Eynan, A., "Deriving the Optimal Boundaries for Class-Based Automatic Storage/Retrieval Systems," *Management Science*, Vol.35, No.12(1989), pp.1519-1524.

[9] Sule, D. R., *Manufacturing Facilities*, Kent Publishing, Boston, MA. 1988.

[10] Seidmann, A., "Intelligent Control Schemes for Automated Storage and Retrieval Systems (AS/RSs)," *International Journal of Production Research*, Vol.6, No.5(1988), pp.931-952.

[11] White, J. A., "Trends in Warehousing," *Proceedings of the Fifth International Conference on Warehousing, Atlanta, GA.* 1983.

[12] Wong, C. K., "Minimizing the Expected Head Movement in One-Dimensional and Two-Dimensional Mass Storage Systems," *Computing Surveys*, Vol.12, No.2(1980), pp.167-178.