FERDINAND F. LEIMKUHLER

Systems Analysis in University Libraries

A comprehensive engineering approach to the analysis and functional design of library systems is described in terms of the fundamental space-time relationships which characterize university libraries. Long-run trends in acquisitions and circulation are related to the relative obsolescence of stored materials, and the uncertainty of short-run demand patterns is related to the need for excess service capability. The spatial dispersion of library resources among specialized information centers and central depositories is considered with respect to availability, retrieval, duplication, and efficient storage.

T HIS PAPER attempts to outline a fairly comprehensive, engineering approach to the functional design of a university library system. It is comprehensive in the sense that it considers both long-run and short-run patterns in library operations in an integrated manner. It is an engineering approach in that it attempts to isolate the fundamental space-time relationships that characterize a library and to treat them in an analytic manner.

Although a library is essentially a social institution steeped in human values of all kinds, it can be viewed as a complex communication system charged with the task of transferring information through space and time, and as such, it is particularly amenable to engineering analysis. The mathematical models which emerge from such an analysis may seem far removed from the librarian's view of the library, but they are almost

Dr. Leimkuhler is Associate Professor of Industrial Engineering in Purdue University. This paper was read before the 1965 annual meeting of the American Society for Engineering Education in Chicago.

certain to provide a much better basis for the design and operation of library systems than is now available.

The models developed below are the result of several years of discussions and joint research efforts by librarians and industrial engineers at Purdue University. There is no attempt here to make forced applications of industrial techniques, although the general methodology of operations research is readily applicable in the library environment. The appropriate goal is neither the advent of total library automation nor the complete mathematical description of a library system. Rather, it is a search for reliable methods of measurement and analysis which are compatible with the established objectives of library administration.

In the remainder of the paper, models are developed to help to explain how acquisition, circulation, storage, loan period, and duplication policies and patterns interact over the long run. Circulation, as a measure of library activity, becomes the one factor related to all of

the others. Total circulation is shown to depend on the size of the collection, its rate of growth, the usage of new acquisitions, and their obsolescence rate. The circulation of individual items is later seen to be closely bound up with the demand pattern, the loan period, and the availability of duplicate or substitute materials.

Availability can be increased by purchasing duplicate copies and maintaining branch libraries but with a consequent increase in the eventual accumulation of obsolete materials. Tighter control of acquisitions and loan policies is seen to reduce obsolescence to a limited extent, but eventually every library will have a considerable amount of inactive materials. Optimal selection rules for

withdrawing these items from active storage, and optimal shelving rules for storing the selected materials are two promising methods for relieving the pressure on today's bulging libraries.

The models presented have been deliberately simplified in order to emphasize their contribution to an overview of the library as a system. Particular applications would require considerably more detail and complexity in the models. In their present state, however, they do serve to indicate some general trends, isolate the important factors associated with the trends, and indicate areas for further study.

The development is divided into three parts: *(a)* growth models, *(b)* storage models, and *(c)* availability models. The
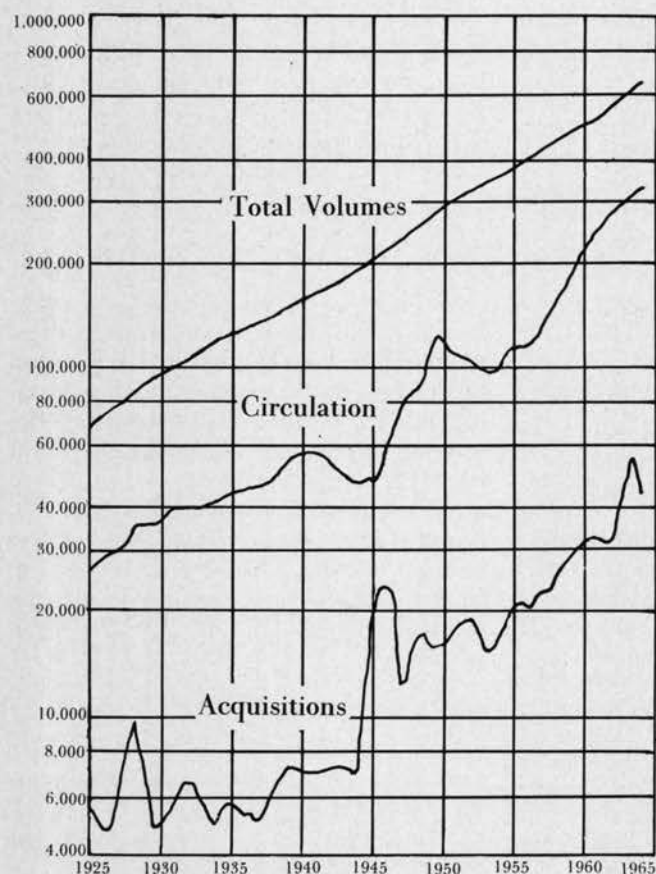

Fig. 1—Purdue libraries

first involves a time-series analysis of acquisitions, holdings, and circulation. The second part considers several aspects of book storage in a manner that is analogous to modern inventory theory. The third part treats the library as a stochastic service system and applies some elementary queueing models to determine the availability and circulation of single and duplicate copies at central or branch libraries.

ACQUISITIONS, CIRCULATION, AND OBSOLESCENCE

The twentieth-century explosion in knowledge is reflected in the growth of university libraries. A study of Danton's[1] figures for the holdings of the ten largest American universities indicates a fairly steady rate of growth of 4.5 per cent per annum since the year 1850. This is the same rate Johnson[2] cites for the growth in scientific journals since the year 1700.

The history of the Purdue University libraries over the past forty years shows an amazingly steady increase of about 6 per cent per annum in both acquisitions and holdings. The evidence suggests an elementary growth model of the form:

$$N_t = (1 + a) N_{t-1} = (1 + a)^t N_0 \quad (1)$$

where N_t defines the size of the collection at the end of year t as a function of the annual growth rate, a, and the size t years ago, N_0. It follows that the acquisitions, A_t in year t are proportional to the size of the collection, since:

$$A_t = N_t - N_{t-1} = a N_{t-1} = \frac{a}{1+a} N_t . \quad (2)$$

Circulation records are a measure of the service performed by a library, but such data has only recently been given the scrutiny it deserves. A forty-year

record of circulation (home use) at the Purdue libraries suggests a long-run trend that parallels that of acquisitions and holdings, with some fluctuation both above and below the trend line. At all times, however, the total number of books checked out is considerably less than the total holdings, with the implication that a large number of books in any one year are inactive and that the number of such volumes is increasing steadily (Fig. 1).

Studies by Fussler and Simon[3] at Chicago and by Jain[4] at Purdue have shown that the average circulation of library materials decreases consistently with the length of time since their publication or acquisition. A preliminary analysis of their data suggests that this decrease occurs at a fairly steady rate. Jain's data suggests a 6 per cent annual decline since acquisition and a 4.5 per cent rate since publication, but these estimates are based on a relatively small sample.

If the annual decrease in activity occurs at some constant "obsolesence" rate, β, then the circulation in year t of all books acquired in year t-j can be defined as follows:

$$C_t (A_{t-j}) = C_0 (1 - \beta)^j A_{t-j} \quad (3)$$

where C_0 is their average circulation rate in the first year of acquisition. By combining equations (1), (2), and (3), the annual circulation of the entire collection can now be defined by:

$$C_t (N_t) = \frac{aC_0}{a + \beta} N_t + \frac{\beta C_0}{a + \beta} (1 - \beta)^t N_0 . \quad (4)$$

Here, the first term defines the long-run trend, since the second term disappears for large t. However, any change in a, β, or C_0 would cause an asymptotic shift in circulation to a new trend line.

[1] J. P. Danton, *Book Selection and Collections: A Comparison of German and American Universities* (New York; Columbia, 1963).

[2] E. A. Johnson, "The Crisis in Science and Technology and Its Effect on Military Development," *Operations Research*, VI (No. 1, 1958), 11-34.

[3] H. H. Fussler and J. L. Simon, *Patterns in the Use of Books in Large Research Libraries* (Chicago: The University of Chicago Library, 1961).

[4] A. K. Jain, *A Sampled Data Study of Book Usage in the Purdue University Libraries* (Lafayette, Ind.: Purdue University, 1965).

CIRCULATION, STORAGE, AND SHELVING

Many libraries are contemplating or practicing some form of depository storage of less active materials. Methods of selecting items for storage have been proposed by Fussler and Simon, Jain, and Trueswell,[5] all of which seek to minimize the anticipated circulation of the stored items as a collection, *i.e.* to select *least used* material for storage.

One such plan would be simply to relocate all items acquired more than d years ago, as is sometimes done with periodicals. The relative advantages of such a plan can be demonstrated with the use of equations (1) and (4). The proportion, n_d, of the total collection stored under this plan is found from equation (1) to be:

$$n_d = \frac{N_{t-d}}{N_t} = \frac{1}{(1+a)^d}. \quad (5)$$

From equation (4), the proportion of the total circulation which is due to the stored items is defined by:

$$c_d = \frac{C_t(N_t-d)}{C_t(N_t)} = \left(\frac{1-\beta}{1+a}\right)^d \quad (6)$$

where a is the growth rate and β is the obsolesence rate as before.

Values of n_d and c_d are shown in Table

[5] R. W. Trueswell, "A Quantitative Measure of User Circulation Requirements and Its Possible Effect on Stack Thinning and Multiple Copy Determination," *American Documentation*, XVI (January 1965), 20-25.

TABLE 1
RELATIVE USE OF STORED ITEMS

Age, d, in Years, When Stored	Proportion of Collection in Storage*	Proportion of Circulation from Storage
5	0.75	0.55
10	.56	.30
12	.50	.24
15	.42	.17
20	.31	.09
25	.23	.05
30	.17	.03
35	.13	.02
40	.10	.01

* A 6 per cent acquisition and obsolesence rate assumed.

TABLE 2
EFFICIENCY OF STORAGE BY SIZE

NUMBER OF SIZE CLASSES	POTENTIAL CAPACITY INCREASE	
	Storage by Height, Per Cent	Storage by Width, Per Cent
1	0	27
2	38	94
3	47	111
4	51	116
5	53	121
10	58	128

1 for a hypothetical library with $a = \beta = 0.06$. Such a library can expect to satisfy 95 per cent of its circulation with its acquisitions of the past twenty-five years, or about 77 per cent of the total collection. Acquisitions of the past twelve years alone account for half of the collection and 76 per cent of the circulation.

Better utilization of shelf space can often be made in the storage of inactive materials. Leimkuhler and Cox[6] developed an exact method for evaluating the spatial efficiency of shelving books by size, and Cox[7] has extended the model to include various aspects of shelf construction. Raffel[8] has recently published a very efficient programing model for evaluating shelf storage by both height and width. These models are capable of specifying the best size classifications (shelf-heights) to use for any given collection.

Table 2 gives the results for shelving a representative sample of the Purdue collection at optimal shelf heights. It can be seen that the additional gain in capacity decreases quite rapidly as the number of size classes increases. The Purdue studies indicate that no more than three or four classes would ever be needed in a compact storage area. Further studies of shelving systems are needed especially with regard to their cost and compatibility in the total library context.

[6] F. F. Leimkuhler and J. G. Cox, "Compact Book Storage in Libraries," *Operations Research*, XII (May-June 1964), 419-27.
[7] J. G. Cox, *Optimum Storage of Library Materials* (Lafayette: Purdue University Libraries, 1964).
[8] L. J. Raffel, "Compact Book Storage Models" (Master's thesis, Purdue University, 1965).

It is worth remembering that Dewey[9] recommended shelving by size. In general, subdivisions of large collections in various ways, by size, use, etc., permits more efficient operations on a suboptimal level. The bigger problem is to integrate the parts into a unified and effective system.

AVAILABILITY, CIRCULATION, AND DUPLICATION

Studies at Massachusetts Institute of Technology by Morse[10] and others have focused on the random nature of the demand for library services. This is a crucial element in the design of almost all kinds of service facilities and usually implies the need for considerable "over design" in the system or the provision of excess service capability. A well developed engineering treatment of this topic has evolved over the past fifty years within the telephone industry and recently has found wide application in such diverse places as supermarkets and superhighways. The theory is bound to play an increasingly important role in the future design of library systems.

As an elementary application, consider the circulation pattern of a single volume, which is requested randomly or independently in time but at a steady rate, R. (This is the pattern observed by Morse.) The average time interval between checkouts for the book can be divided into two parts: The mean loan interval, T_L, and the mean interval between checkin and checkout, T_R, which in the "random" case is equal to $1/R$. Thus, the mean circulation rate C_1 of the volume can be written:

$$C_1 = \frac{1}{T_R + T_L} = \frac{R}{1 + \frac{T_L}{T_R}} \qquad (7)$$

which indicates that the circulation rate is always less than the request rate.

[9] Melvil Dewey, "Library Shelving, Definitions and General Principles," *Library Notes*, II (September 1887), 95-122.
[10] P. M. Morse, "Probabilistic Models for Library Operations," Minutes of 63rd Meeting, Association of Research Libraries (Chicago: January, 1964), p. 9-13.

The proportion of satisfied customers, C_1/R, is a useful measure of the availability of the book; and the fraction, T_L/T_r, is a measure of the service load on the book, commonly called "traffic intensity." Where $T_L = T_r$, the book is only available for 50 per cent of the requests. As the request rate diminishes, availability increases but circulation decreases. If the request rate increases, circulation increases also, but availability falls off rapidly, so that the circulation increase is accompanied by a considerable rise in the number of disappointed customers.

Two common library responses to the demand for popular books are: (a) the shortened loan period, and (b) the acquisition of extra copies. Equation (7) can serve to evaluate the first of these strategies, since a reduction in T_L reduces the traffic intensity and increases both availability and circulation, assuming no change in demand. A 50 per cent reduction in the loan period produces the same effect on availability that a 50 per cent decrease in the request rate would produce. However, circulation is increased in the former case and reduced in the latter case.

The analysis of the effects of duplication on circulation and availability requires an extension of equation (7). The circulation rate, C_2, for two copies can be defined in terms of the rate, C_1, with one copy as follows:

$$C_2 = \frac{2C_1}{1 + \left(\frac{C_1}{R}\right)^2} . \qquad (8)$$

Since the availability with one copy, C_1/R, is always less than unity, it can be seen immediately that two copies can never succeed in doubling circulation. For example, if the single copy was available only half of the time, duplication will increase total circulation for the two volumes only 60 per cent above that of the single volume. Availability is increased in the same proportion, *i.e.* from 0.5 to 0.8.

What is even more significant for the

long run is the fact that duplication always reduces the *average* circulation per volume. For the previous example, the circulation per volume was *reduced* 20 per cent in order to increase availability. As the demand for a particular book diminishes, duplicate copies add less and less to the availability but contribute significantly to the apparent obsolescence of the collection.

The above models also shed some light on duplication among branch libraries, since this plan is analogous to the division of requests into two separate streams each serviced by its own copy. If loans between branches are not made, then it can be shown that the increase in circulation is always less than if the two copies are held at the same location.

The best that can be done under such circumstances is to divide the demand equally between the branches, which yields a total circulation of:

$$C_{2B} = \frac{2C_1}{1 + \dfrac{C_1}{R}}. \qquad (9)$$

For the above example, where $C_1/R = 0.5$, the total circulation is increased only 33 per cent by the branch policy of duplication, and circulation per volume is decreased 33 per cent. The availability at each branch would be 0.67 as contrasted to 0.5 for the single volume and 0.8 for two copies at a central point. With unequal division of demand between the branches the total circulation is increased even less and circulation per volume reduced even more.

These models serve to indicate the effects of various policy decisions. It is unlikely that they are precise estimators of actual performance because of the assumptions required in their development, as, for example, that the demand rate for a book is independent of its availability or the loan policy. Much more complex models would be required to account for such reader reactions as discouragement and impatience and the substitution of similar books. Such models can be or have been developed to varying degrees of perfection within the voluminous literature of queueing theory. The difficulties of adapting these theories to library practice should not be underestimated, but there is no question of the feasibility and eventual benefits of such an undertaking. ∎∎

ACRL MEMBERSHIP DEC. 29, 1965

Total	9,911
Subject Specialists	2,049
Junior College	874
University	3,837
College	3,099
Rare Books	1,121
Institutional Memberships	1,793

Please note that many members do not select membership in sections.

Choice Address Change

THE EDITORIAL, advertising, and production offices of *CHOICE: Books for College Libraries* are once again being moved to larger quarters—on December 31, 1965. The new address is 100 Riverview Center, Middletown, Connecticut 06457. The telephone number remains the same: (203) 347-6933.