Queueing Systems, 4 (1989) 319–338

MODELING CIRCULATION SYSTEMS IN BUILDINGS USING STATE DEPENDENT QUEUEING MODELS *

Steven J. YUHASKI, Jr and J. MACGREGOR SMITH

Department of Industrial Engineering and Operations Research, University of Massachusetts, Amherst MA 01003, U.S.A.

Received 1 February 1988; revised 12 December 1988

Abstract

Circulation systems within buildings are analyzed using $M/G/C/C$ queueing models. Congestion aspects of the traffic flow are represented by introducing state dependent service rates as a function of the number of occupants in each region of the circulation system. Analytical models for unidirectional and multi-source/single sink flows are presented. Finally, use of the queueing models to analytically determine the optimal size and capacity of the links of the circulation systems is incorporated into a series of software programs available from the authors.

Keywords: State dependent, queueing networks, and facility planning.

1. Introduction

In this paper, we develop state dependent queueing models for capturing the congestion effects of movement through circulation systems of buildings. The circulation system includes the corridors, ramps, elevators, stairways and other physical paths of movement within buildings. We focus on the detailed development of single-level corridor models which can be expanded to represent multiple corridors and multi-level circulation systems.

Whether an architect or engineer is concerned with optimal accessibility requirements of a new or remodeled facility in order to accomodate expected customer demands or is concerned with the optimal egress of occupants due to emergency situations, pedestrian flow in the circulation systems of buildings is one of the most prevelant problems in facility design. Given these customer flows within a facility, a *deterministic* decision must be made with regards to the size or finite capacity of the circulation system of the facility i.e. its nodes and links while at the same time recognizing the stochastic nature of the customer flows. Certainly, this is a very complex stochastic programming problem.

What we intend to do in this paper is offer some congestion models to capture this stochastic customer flow and at the same time illustrate how these stochastic

* This material is based upon work supported by the National Science Foundation under grant #MSM-8417942 and #MSM-8715152.

queueing models can be used to optimally size and determine the circulation links of the system. Albeit, our focus will be on one aspect of the circulation system, but given its central importance, it is readily applicable to many other parts of the circulation system.

DEFINITION OF PROBLEM

Congestion within buildings occurs mainly in the circulation system, of the building, viz. the corridors, ramps, hallways, elevators, and stairwells. This congestion is due largely to the increased customer traffic seeking to occupy the limited space available within the circulation elements of the building.

There are two crucial aspects of these movement systems:

1) The service rate of the movements system decays with increasing traffic.
2) The amount of available space within the movement system is finite.

Naturally these characteristics lead us to the question: how should one model the congestion within the circulation system of a building? What models have been used in the past; what theory is available; and, how should we "best" capture the congestion effects?

FACILITY REPRESENTATION

Every floor plan of a facility can be represented as a planar graph $G'(V', E')$. Thus, our queueing network is essentially the Dual Graph of G', Smith [13–15]. Also, facilities have two spatial entities which require a distinction within our queueing network model: the activity network and the circulation network. Thus, V is partitioned into two sets $V = \{A, S\}$, see fig. 1:

A set of activity nodes $A = \{A_1, A_2, \ldots, A_N\}$ of size N which characterize the rooms, departments or activity areas. The notion of "activity area" is common parlance in facility design.

A set of circulation nodes $S = \{S_1, S_2, \ldots, S_M\}$ of size M which characterize the hallways, corridors, stairwells, elevators, and other movement pathways.

Each facility can be best described as a hierarchical queueing network of nodes and arcs. For our purposes, the hierarchical queueing network has two levels. At the upper level, we have termed the resource activities, $V_i = (1, 2, \ldots, I)$ where $I = |N| + |M|$. At the lower level, we have the subnetwork of activities within each resource activity, $v_{it} = (1, 2, \ldots, T)$.

The incidence function that regulates the flow of customers among the subnodes is a subnetwork transition matrix and has the following form:

$$P_{(i,j,l)} = \begin{pmatrix} (p_{11})_{i,j,l} & (p_{12})_{i,j,l} & \cdots & (p_{1t})_{i,j,l} \\ (p_{21})_{i,j,l} & (p_{22})_{i,j,l} & \cdots & (p_{2t})_{i,j,l} \\ \vdots & \vdots & \ddots & \vdots \\ (p_{t1})_{i,j,l} & (p_{t2})_{i,j,l} & \cdots & (p_{tt})_{i,j,l} \end{pmatrix}.$$

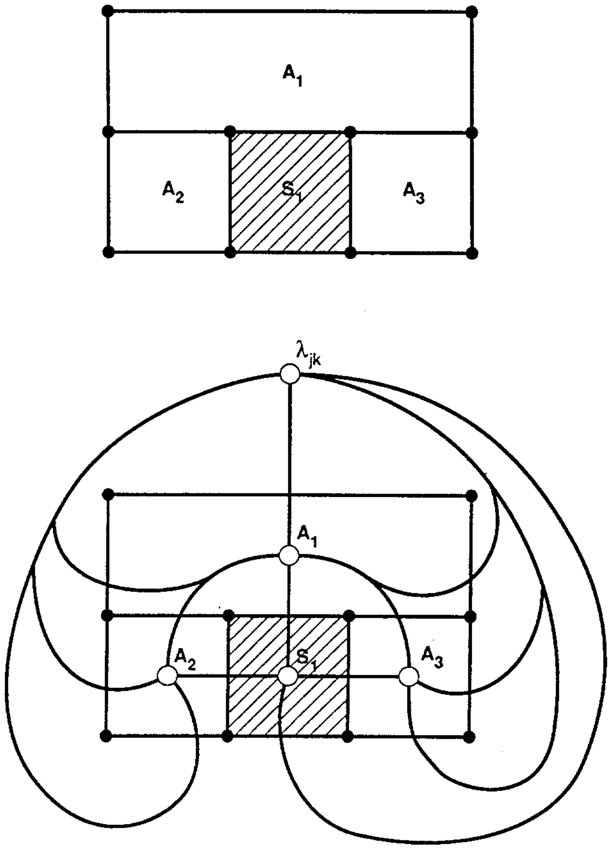

Fig. 1. Planar graph $G'(V', E')$ and dual graph $G(V, E)$.

Transitions within resource activities are normally accounted for by the amount of time spent at each subactivity v_{it} by a customer following his route through $G(V, E)$. Arcs between resource activities (A_i, A_j) represent travel time over the circulation network S of the facility.

One of the unique features of facility modelling is that transitions within the network are not virtually instantaneous Lee [9] as they are in computer and communication networks Kleinrock [8] and Sauer [12]. Thus, the need for the set S. Vertices from the set S represent additional nodes within a facility designed to handle the intermediate flow of customers from origin source nodes A_i to A_k without interrupting service. Often the cardinality and configuration of the nodes in S is determined by the intended building use, building codes, and zoning ordinances for certain building types due to emergency egress conditions or other public access requirements.

We like to refer to these nodes as Steiner nodes, after the renowned Steiner network problem Smith [13], since there is a similarity between the nodes from the set S within queueing network modelling of facilities and their presence in minimizing overall network costs in the Steiner network problem. It is the customer traffic flow within S that is of primary interest to us in this paper.

Previously, we have modelled movement within the circulation system using an $M/G/\infty$ queue to capture customer traffic flow Smith [13–16] in each circulation node S_j. Use of $M/G/\infty$ queues is quite prevelant in the modelling of computer systems Kleinrock [8].

In addition, we utilized a congestion factor to slow-down traffic through the circulation system based on the *traffic intensity* as measured by ρ to this $M/G/\infty$ queue. Use of this congestion factor is tantamount to showing that the service rate in the queue is state dependent on the number of customers passing through the circulation system of a building.

Modelling of state dependence originates with the work of Conway and Maxwell [2], who studied the situation where the service rate of a server increases with increasing traffic. They introduced the notion of "pressure coefficients" to capture the nonlinear effects on the service rate as increased traffic entered the system. Hillier et al. [7] later extended this work to account for multiple-servers.

In their model, if the service rate increases, then

$$\mu_n = n^{\gamma}\mu$$

where $\gamma \geqslant 0$.

This model is not defined for $\gamma < 0$ because the infinite series used to calculate the probabilities in the Chapman-Kolmogorov equations do not converge Gross and Harris [5]. However, $\gamma < 0$ is the type of parameter expression which corresponds to a deterioration in service rate, which is exactly what we need in our circulation congestion model. This need to capture a service rate decay in S led to the development of the finite, state-dependent congestion model which now follows.

ASSUMPTIONS

The common assumptions we have made over the years in modelling facilities with queueing networks are as follows.

COMPLEX MIXING OF ARRIVAL AND DEPARTURE PROCESSES

There are J customer types (classes) which seek to utilize a facility drawn from an infinite population each with K generating sources. The average arrival rate per unit time of type j customer from source k is λ_{jk} $(j = 1, 2, \dots J;\ k = 1, 2, \dots K)$. It is further assumed that this arrival process from each source is a renewal process *.

* Certainly, in some facilities such as airports, there are batch arrivals, yet for the most part, customers arrive singly, independently, and randomly. Furthermore, customers tend to follow deterministic rather than random routes through a facility due to their origin and destination plans.

DETERMINISTIC ROUTING CHAIN

Type jk customers will be routed through the facility using a deterministic routing vector, sometimes referred to as a "customer chain". This routing vector has elements r_{jkl} ($l = 1, 2, \ldots L_{jk}$) where the lth element of the routing vector marks the destination of the resource activity to which the customer is directed after visiting the previous activity on its route. Thus, customers of types jk ($j = 1, 2, \ldots J$; $k = 1, 2, \ldots, K$) enter a system of queues l ($l = 1, 2, \ldots, L_{jk}$) in independent Poisson streams at rate λ_{jk} and pass through a sequence of queues:

$$\left[(j, k, 1), (j, k, 2), \ldots, (j, k, L_{jk})\right]$$

before leaving the facility. Therefore at stage l ($l = 1, 2, \ldots, L_{jk}$) along its route a customer of type jk will be at queue $Q(j, k, l)$.

$G(V, E)$ IS A DUAL GRAPH

$G(V, E)$ can be embedded in the plane R^2 with a corresponding distance metric (Euclidean or Rectilinear), since the flow of customers and goods through the nodes requires a distance movement. Thus, the location of each $v_i \in V$ is in R^2, or with Cartesian coordinates (x_i, y_i). In the queueing network, the location of the nodes should arguably be the centroid of the spatial entity it models. Although one may model a facility without the network embedding, one needs to characterize the transition time between activity nodes in the network and this requires a distance movement.

In the material that follows, we present the congestion model in §2 for capturing the state dependent customer flow; then in §3 we develop the queueing model; in §4 we discuss some computational considerations; and, finally, in §5 we perform some sensitivity analysis to illustrate how the state dependent queueing model can be utilized in the planning and design of circulation systems within buildings.

2. Congestion model

It is reasonable to believe that as the number of people traveling along the corridor increases, the average speed of the pedestrians will tend to decrease. As more people occupy the limited floor-space of the corridor, there will be more of a tendency for the slower pedestrians to impede the progress of the faster pedestrians since a greater degree of congestion will restrict the lateral movement needed by the faster pedestrians for passing and avoiding the slower walkers.

A quantitative understanding of congestion can best be attained by considering unidirectional traffic along a corridor of length, L, and constant width, W. This elementary shape will henceforth be referred to as a single corridor. Obviously

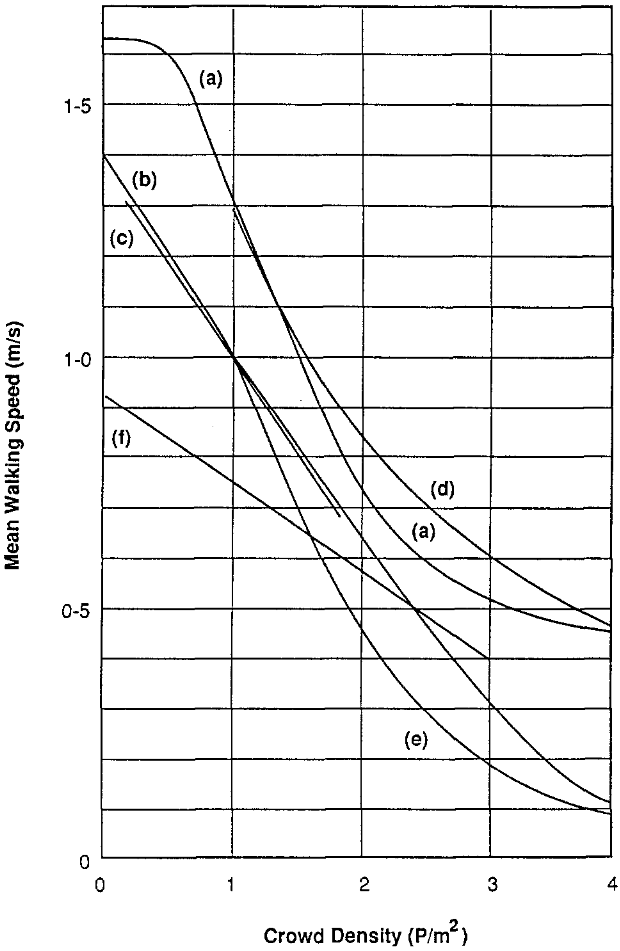

Fig. 2. Variation of mean walking speed with crowd density (a) [HANK 58], (b) [OFLA 72], (c) [OLDE 68], (d) [TOGA 55], (e) [TOGA 55], (f) [FOOT 73].

there is a limit on the number of pedestrians that can move through the corridor at any given moment; we will refer to such an upper bound as the capacity, C, of the corridor. When there are n pedestrians occupying the single-corridor, we will say that they travel at an average walking-speed, V_n, for $n = 1, \ldots, C$.

A useful way to describe congestion is in terms of the number of pedestrians per floor area in our single corridor; we will refer to this as the population density, ρ, of the corridor. Empirical relationships between walking-speeds and population-density play an essential role in the development of mathematical relations that determine the overall walking-speed of the pedestrians as a function of the population-density. These empirical relationships are usually presented in graphical form Fruin [4] and Tregenza [18], see fig. 2 which is re-created from Tregenza.

Figure 2 provides a number of implications in terms of what the most definitive mathematical model may be. The shapes of the curves suggest either a linear or an exponential relationship (depending on the curve) for the walking-speed of the pedestrians as a function of the population-density. Due to practical considerations; however, we will examine the walking-speed, V_n, as a function of either the corridor capacity, C, or the area of the floor space, and the number of pedestrians, n, occupying the corridor.

The concept of corridor capacity is a central concept with respect to our congestion modelling. According to Tregenza [18], pedestrian flow comes to a halt when the population-density approaches five pedestrians per square meter (i.e., 5 peds/m^2). Therefore, the corridor capacity, C, is equal to the highest integer that is less than or equal to five times the area of the floor-space in square meters (m^2). Thus, the capacity is expressed as,

$$C = [5LW], \tag{1}$$

when L and W are given in meters.

When using a linear relationship to approximate the overall walking-speed of the pedestrians through the single corridor, we need to take note of the fact that the speed of a lone occupant in the corridor, V_n, is typically 1.5 m/sec, according to [18]. We also need to be aware of the idea that as n approaches C (i.e., $n \to C$) the pedestrian flow comes to a halt; however, there may still be some forward movement at $n = C$. For this reason, we can say that since a population of $n = C + 1$ is an impossibility, we can set $V_n = 0$ for all $n \geqslant C + 1$. Thus, we are deriving a linear relation that satisfies $V_1 = 1.5$ m/sec and $V_{C+1} = 0$, which is shown by eq. (2).

$$V_n = \frac{1.5}{C}(C + 1 - n). \tag{2}$$

The appearance of the curves in Tregenza [18] suggests strongly that the exponential relationship may be a more accurate approximation of the average walking-speed of the pedestrians. The form of the exponential relation for our single corridor model is, for $n = 1, 2, \ldots, C$:

$$V_n = A \exp\left[-\left(\frac{n-1}{\beta}\right)^{\gamma}\right]. \tag{3}$$

The amplitude $A = 1.5$ m/sec. Parameters β and γ will be referred to as the scale and shape parameters respectively. By carefully approximating the positions of three representative points among the six curves in Tregenza [18], we have the following coordinates:

$$V_n = 1.50 \text{ m/sec at } \rho = 1/LW \text{ peds/m}^2 \Leftrightarrow n = 1$$

$$V_n = 0.64 \text{ m/sec at } \rho = 2 \text{ peds/m}^2 \Leftrightarrow n = a = 2LW$$

$$V_n = 0.25 \text{ m/sec at } \rho = 4 \text{ peds/m}^2 \Leftrightarrow n = b = 4LW.$$

Fitting the points $(1, A)$, (a, V_a), and (b, V_b) gives us the algebraic relationships shown below:

$$\gamma = \ln\left[\frac{\ln(V_a/A)}{\ln(V_b/A)}\right] \Big/ \ln\left(\frac{a-1}{b-1}\right) \tag{4}$$

$$\beta = \frac{a-1}{\left[\ln\left(\frac{A}{V_a}\right)\right]^{1/\gamma}} = \frac{b-1}{\left[\ln\left(\frac{A}{V_b}\right)\right]^{1/\gamma}}. \tag{5}$$

In eqs. (4) and (5), $A = 1.5$ m/sec. Note that γ must be determined prior to determining β; this is done by using eq. (4) prior to using eq. (5).

Note that weighted nonlinear regression could be used instead of judiciously selecting three points from the curves of Tregenza [18]. Each of the curves could be assigned a relative weight between zero and unity (such that all six weights sum to unity), and a predetermined number of points could be taken uniformly along the population-density abscissa over the range of 0 to 4 peds/m. At this stage of development in the research, however, the difference in results between the three-point approximation and the weighted nonlinear regression would be expected to have a conceptual impact that is insignificant.

Other mathematical relations could be used to relate the walking-speed of the pedestrians to the pedestrian-density or to the number of occupants on a given floor-space. Piecewise linear approximations of the curves would be a more accurate approximation than a curve-fit that is based on only one continuous line, such as that of eq. (2). A series of piecewise linear and exponential curve fits would be even more accurate; however, the level of accuracy of the data from which the walking-speed approximations would be determined, along with our stage of development in this line of investigation, does not justify such an effort on purely a conceptual basis. For purposes of illustration and experimentation, we will restrict ourselves to strictly linear or exponential approximations of the curves.

3. Analytical model for the single corridor

To gain a clear understanding of the concepts involved in the stochastic modeling of congestion in corridors, we address the methodology involved in modeling the single corridor as a queueing system.

As discussed in the previous section, the single corridor of length, L, and constant width, W, has a capacity $C = [5LW]$, when L and W are given in meters. We assume an exponential service rate, μ_n, when the corridor is occupied by n pedestrians. In other words, the occupants enter the corridor with the behavior of a Poisson stream of rate, λ, and the time that the occupant spends in the corridor is exponentially distributed with rate, μ_n. Note that the service rate,

μ_n, is state-dependent, i.e., a function of the number of occupants. Since the corridor behaves as a server to its occupants, we can model it as a queue having C servers and a capacity of C. Thus, in Kendall notation, our queueing model is described by $\mathrm{M/M}/C/C$.

It is usually of interest to note that the queueing model $\mathrm{M/M}/C/C$ is stochastically equivalent to the model $\mathrm{M/G}/C/C$ provided the mean service rates of the two models are equal; details can be obtained by consulting Gross and Harris [5] among other sources.

The general solutions to the Chapman-Kolmogorov steady-state difference equations for the state probabilities $p_1, p_2, \ldots, p_C$, are shown in eqs. (6) and (7).

$$p_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} p_0 \tag{6}$$

such that,

$$\frac{1}{p_0} = 1 + \sum_{n=1}^{C} \left[\frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} \right]. \tag{7}$$

In the context of our investigation, the arrival rates are not influenced by n; thus, we define λ, such that, $\lambda = \lambda_0 = \lambda_1 = \cdots = \lambda_C$, which gives,

$$p_n = \frac{\lambda^n}{\prod_{i=1}^{n} \mu_i} p_0, \text{ for } n = 1, \cdots, C \tag{8}$$

and

$$\frac{1}{p_0} = 1 + \sum_{n=1}^{C} \left\{ \frac{\lambda^n}{\left[\prod_{i=1}^{n} \mu_i \right]} \right\} \tag{9}$$

where μ_i, for $i = 1, \ldots, C$, is a function of i, the state of the system (i.e., the number of occupants). Since congestion directly affects the service rates of our queueing model, we use the linear congestion model and the exponential congestion model to describe μ_n. Both of the congestion models are discussed within the next two sections.

The linear congestion model is based on the idea that the service rate of the servers in the $\mathrm{M/G}/C/C$ queueing model is a linear function of the number of occupants in the system. Equation (2) gives the walking-speed of n occupants in the single corridor. Note that the service rate, r_n, of each of n individuals in the corridor, is the average of the inverse of the time it takes these individuals to traverse the length of the corridor; therefore,

$$r_n = V_n / L. \tag{10}$$

Using eq. (2), this gives us,

$$r_n = \frac{1.5}{CL}(C + 1 - n). \tag{11}$$

The service rate of the queueing system (overall) is equivalent to the number of servers in operation (i.e., occupied) multiplied by the rate of each server. Since all n servers, in a state-dependent $M/G/C/C$ queueing model, operate at the same rate, r_n, we have,

$$\mu_n = nr_n = \frac{1.5}{CL}(C + 1 - n)n. \tag{12}$$

We derive the expressions for the state probabilities by substituting the expression for μ_n, equation (12), into equations (8) and (9) to obtain equations (13) and (14).

$$p_n = \frac{\lambda^n}{\left(\frac{A}{LC}\right)^n \prod_{i=1}^{n}(C - i + 1)i} p_0 \tag{13}$$

$$\frac{1}{p_0} = 1 + \sum_{n=1}^{C} \frac{k^n}{\prod_{i=1}^{n}(C - i + 1)i} \tag{14}$$

where $A = 1.5$ m/sec, and $k = \lambda LC/A$. Note that L is expressed in meters and λ is expressed in $\sec^{-1}$.

In developing the exponential congestion model, we assume that, r_n, the service rate of each of the n occupied servers, is related to the number of occupants by an exponential function. The form of the exponential function is based on the equation for the walking-speed, as depicted by relation (3). Combining eqs. (3) and (10) gives,

$$r_n = \frac{A}{L} \exp\left[-\left(\frac{n-1}{\beta}\right)^{\gamma}\right] \tag{15}$$

where $A = 1.5$ m/sec. Therefore, we can express the overall service rate of our $M/G/C/C$ queueing model as,

$$\mu_n = nr_n = n\frac{A}{L} \exp\left[-\left(\frac{n-1}{\beta}\right)^{\gamma}\right] \tag{16}$$

We obtain equations for the state probabilities by substituting our expression for μ_n, equation (16), into eqs. (8) and (9) to obtain,

$$p_n = \frac{\lambda^n}{\prod_{i=1}^{n} i\left(\frac{A}{L}\right) \exp\left\{\left[-\left(\frac{i-1}{\beta}\right)^{\gamma}\right]\right\}} p_0 \tag{17}$$

where

$$\frac{1}{p_0}=1+\sum_{n=1}^{C}\left\{\frac{\lambda^n}{\prod_{i=1}^{n} i\left(\frac{A}{L}\right)\exp\left[-\left(\frac{i-1}{\beta}\right)^{\gamma}\right]}\right\}. \tag{18}$$

Note that $A = 1.5$ m/sec, L is expressed in meters, and λ is expressed in $\sec^{-1}$.

SINGLE CORRIDOR WITH MULTIPLE SOURCES

Consider a single corridor of length, L, and width, W, having multiple customer sources (arrival streams), see assumption #1, $\lambda_1, \lambda_2, \ldots, \lambda_K$ whose traveling distances to the exit of the corridor are $L_1, L_2, \ldots, L_K$, respectively. Such a situation can be modeled as another single corridor of length, L', and arrival rate, λ', such that:

$$\lambda' = \sum_{i=1}^{K} \lambda_i \tag{19}$$

$$L' = \frac{\sum_{i=1}^{K} \lambda_i L_i}{\sum_{i=1}^{K} \lambda_i} \tag{20}$$

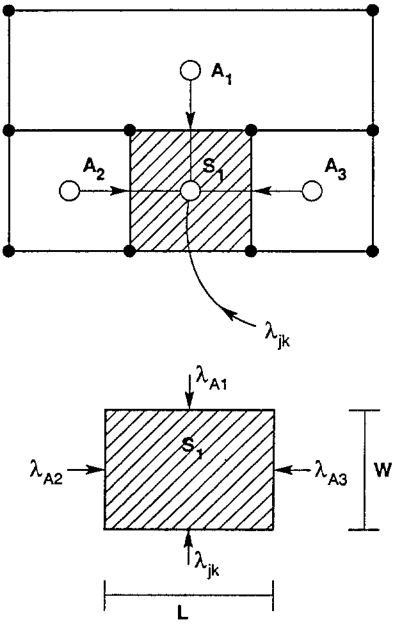

Fig. 3.

Thus, L' of eq. (20) is the weighted average according to $\lambda_1, \lambda_2, \ldots, \lambda_K$. Therefore, L' is the average distance traveled by all the arrivals through the corridor. Figure 3 represents the transformation process for our simple example floor plan embodied in fig. 1, where the activity areas incident to the circulation node act now as inputs λ_i.

We wish to design the dimensions of this type of circulation node so that the traffic flowing through it is adequately accomodated. We shall do so for this class of examples in the §5 on sensitivity analysis but before we do so, we need to discuss some computational considerations for solving our state dependent model.

4. Computational considerations

All computing systems have a limited range of values over which they can process. Some computing systems contain diagnostic error message that report underflows (i.e., very small values) as well as overflows (i.e., very large values). Other systems cope with underflows by automatically setting them equal to zero. Regardless of the type of computing system used in applying the concepts in the preceeding sections of this paper, the magnitude in the quantities involved can result in overflows when the capacity, C, of our single corridor is high enough.

Consider the basic relations of our state-dependent $M/G/C/C$ model, which are depicted by eqs. (8) and (9). If we define terms, T_n, such that $T_0 = 1$, and

$$T_n = \frac{\lambda^n}{\prod_{i=1}^{n} \mu_i}, \text{ for } n = 1, \ldots, C, \tag{21}$$

then, we can see that,

$$p_n = \frac{T_n}{\sum_{i=0}^{C} T_i} \text{ for } n = 0, \ldots, C. \tag{22}$$

For capacities of C on the order of 10000, the maximum T_n, $T^{(\max)}$, can exceed 10^{100}; some computing systems cannot cope with such a situation.

To deal effectively with the overflows, it is necessary to compute T_n and p_n in terms of their logarithms, using eqs. (21) and (22), as follows:

$$\ln(T_n) = n \ln(\lambda) - \sum_{i=1}^{n} \ln(\mu_n), \text{ for } n = 1, \ldots, C \tag{23}$$

$$\ln(p_n) = \ln(T_n) - \ln\left(\sum_{i=0}^{n} T_i\right), \text{ for } n = 0, \ldots, C. \tag{24}$$

The problem of overflows could be dealt with in the calculation of p_n if it where not for the logarithm of the summation term in (24). This problem can be dealt with by dividing (i.e., scaling down) each T_i by a very large (positive) quantity D as shown,

$$\ln(p_n) = \ln(T_n/D) - \ln\left(\sum_{i=0}^{n} T_i/D\right), \text{ for } n = 0, \dots, C. \tag{25}$$

The value of D is chosen such that the value of $(T_n^{(\max)}/D)$ is slightly less than the highest value, $V^{(\max)}$, the computing system will allow. Thus,

$$\ln(D) = \ln\left[T_n^{(\max)}\right] - \ln\left[V^{(\max)}\right]. \tag{26}$$

Note that $T^{(\max)}$ can be detected by determining its logarithm, using eq. (23).

The (T_i/D) terms that are smaller than the lowest value the computing system can cope with are set equal to zero. If the computing system doesn't do this automatically (i.e., reports an underflow instead), then such small values can be handled by determining their logarithms, using eq. (23). The logarithm is examined to determine whether it is less than a negative quantity that is slightly greater than the logarithm of the lowest quantity that the computing system can cope with. If the logarithm of the (T_i/D) term is less than the negative quantity, then the (T_i/D) term is set equal to zero.

The errors in setting (T_i/D) terms that would otherwise cause underflow problems equal to zero are (vastly) insignificant. The difference in absolute values of the upper and lower bounds of quantities that any computing system can handle is considerably larger than the mantissa of the double precision of such systems spans. This is the reason that the value of the summation term in eq. (25) is unaffected when potential underflows are set equal to zero.

5. Optimization and sensitivity analysis

An important factor in the design of most floor plans is whether severe crowding will bring pedestrian traffic to a standstill. Therefore, the effects of the traffic flow rate, λ, on the balking probability, p_C, becomes a significant issue. The dimensional variables, L, W, and therefore the capacity, C, are also factors to be considered. Suppose that an upper limit is imposed on p_C and we are given fixed values for some of the other variables. Under these conditions, it is important to be able to study the response, i.e., the sensitivity, of p_C with respect to the behavior of the remaining (unfixed) variables.

In other contexts, the effective flow rate, λ_{eff}; that is, the rate at which the arrivals are routed through the corridor (floor space), is of paramount concern. For this reason, the study of λ_{eff}, also known as the throughput, as a function of λ, is of practical importance.

In the sub-sections that follow, the issues above are examined. In the next three sections, the focus is placed on the effect that λ, L, and W each have on the value of p_C and how an upper bound on p_C limits the range of values of λ, L, and W. In the last section on sensitivity analysis, the throughput as a function of the arrival rate is examined. Computer programs that are useful in studying the interrelationships of λ, λ_{eff}, L, W, C and p_C discussed in the text are available from the authors.

λ TRAFFIC ANALYSIS

Consider a public corridor with dimensions L and W as depicted in fig. 3 which is our floor plan representation of a single Steiner node with 4 inputs into an equivalent single corridor model. If we know that there is an upper bound on p_C, which will be referred to as p_{CU}, then we will want to know how high the arrival rate, λ, can be without allowing p_C to exceed p_{CU}. We will refer to such an arrival rate as λ_U.

A program named LEAM, available from the authors, is useful as a subroutine to calculate p_C given λ, L, and W until a value of λ is found that gives p_C close enough to p_{CU} to satisfy an arbitrary stopping condition. Another program called LESLAM, also available from the authors, uses bisection as a unidimensional search technique over many values of λ. The interval containing λ_U, which yields p_{CU}, is halved until it is small enough to satisfy the stopping criterion. The approximation to λ is given as the lower bound on the interval that contains the actual λ, since p_C increases monotonically with λ for fixed values on L and W.

Another method for determining λ, given L, W, and p_{CU} involves the use of graphs such as those shown in figs. 4 and 5. The capacity, C, is determined from L and W, by using eqn. (1). The curves of constant capacity are used to project the value of p_C to the axis corresponding to $\rho = \lambda/\mu_1$. Since we are given the value of p_C and can determine μ_1, because $\mu_1 = 1.5/L$ (see eqs. (10) and (12)), we can calculate $\lambda = \rho\mu_1$. It is important to note that the capacity curves in figs. 4 and 5 are the result of values of L and W such that the product $5LW$ is equal to an integer value.

A few interesting observations about the curves in figs. 4 and 5 can be made. First of all, the curves that correspond to higher capacities have lower values of p_C (i.e., the curves never intersect); this should be intuitively obvious. Also, the bend on the curves corresponding to the linear congestion model are much sharper and occur at higher probabilities than those corresponding to the exponential congestion model. Finally, the curves corresponding to the exponential congestion model approach $p_C = 1$ asymptotically much more slowly then those corresponding to the linear congestion model.

p_{CU} ANALYSIS

Suppose that a corridor is to be placed in a location in which it will be exposed to a traffic flow of rate λ; that either one, but only one, of the dimensions (L or W) is specified; and that an upper bound, p_{CU}, is specified for p_C. Under these

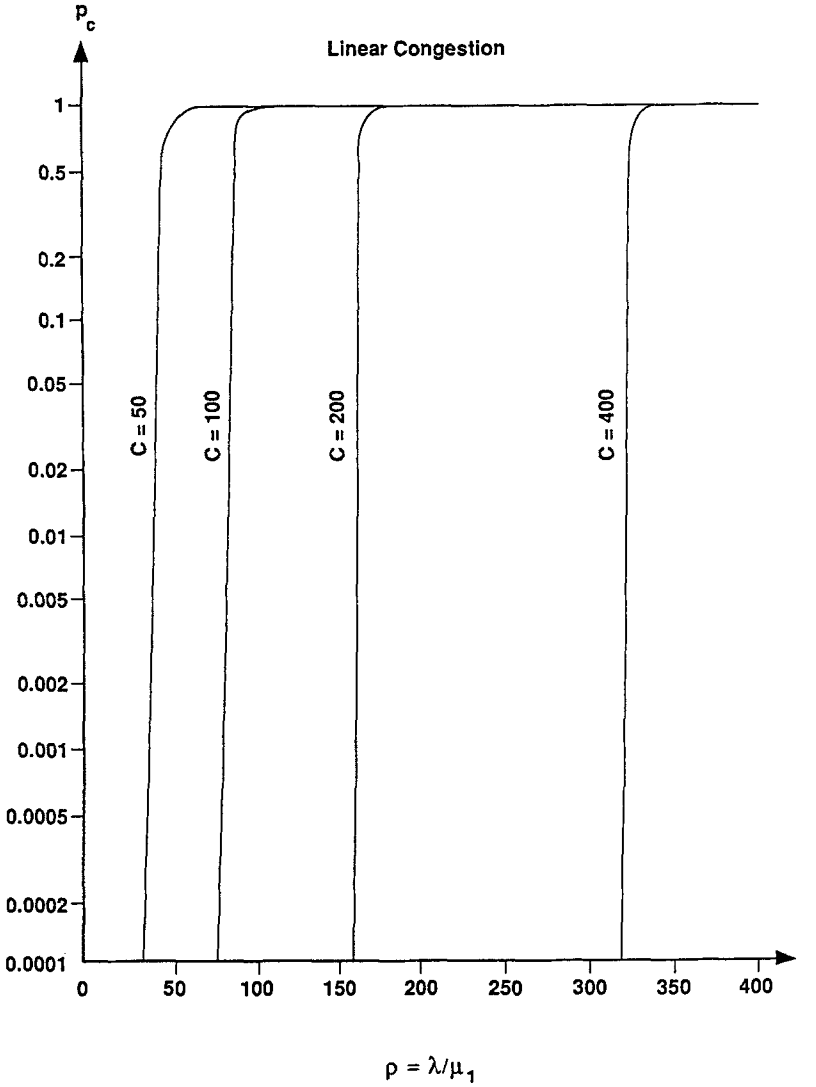

Fig. 4.

circumstances, it would be enlightening to be able to study the behavior of the unfixed dimension (L or W) in terms of its effect on p_C.

Consider a situation in which λ, p_{CU}, and width W, are specified. We would need to investigate the behavior of $p_C \leqslant p_{CU}$ for variations in the corridor length, L. An increase in L has two opposing influences on the traffic flow in terms of p_C. By virtue of increasing the capacity of the corridor, the sensitivity of p_C to changes in congestion decreases since the area of the corridor increases; however, this benefit is adversely affected by the fact that increasing the length gives the

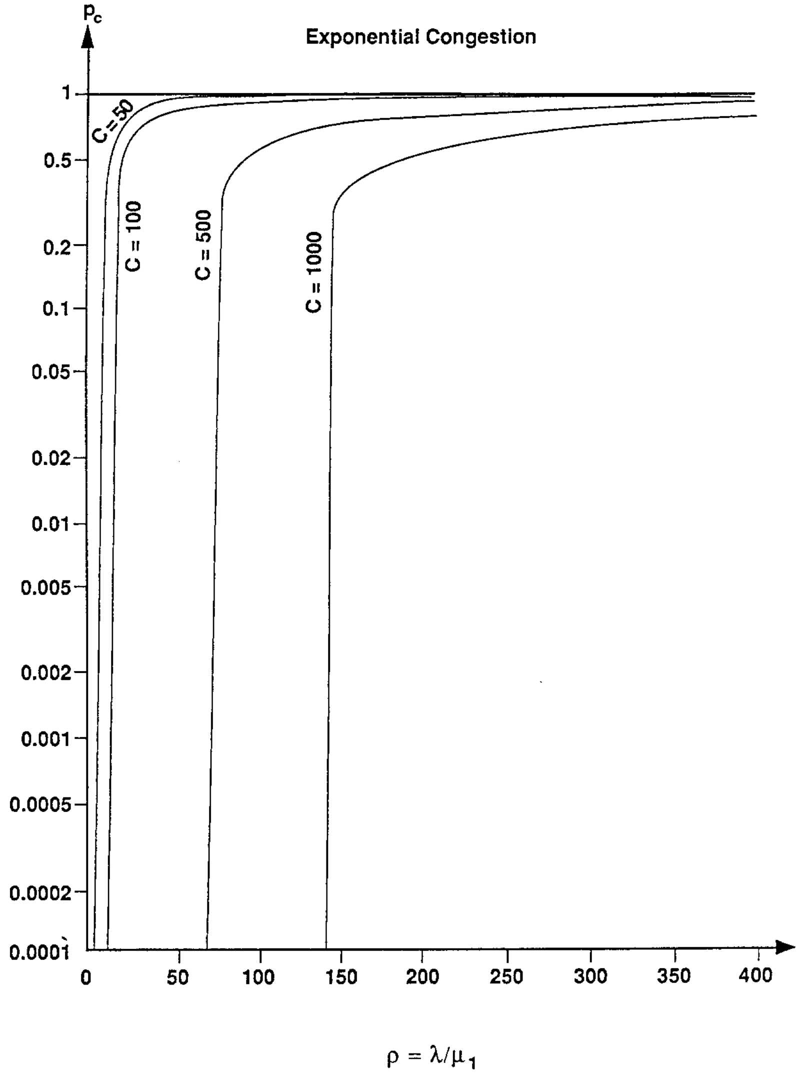

Fig. 5.

occupants a longer distance to travel within the corridor. Because of the behavior of these opposing influences, there is a minimum value of p_C, which we will refer to as $p_C^{(\mathrm{min})}$, with its corresponding length, L^*, and capacity, C^*. For all values of $L \geqslant L^*$ the effect of the occupants having to travel a longer distance to leave has a stronger overall effect on the congestion than the integral increase in capacity that will accompany an increase in L. Therefore, increasing L, when $L \geqslant L^*$, will increase p_C.

There is a more interesting phenomenon that occurs within the interval of values of L, $L_{(\mathrm{min})}^{(C)}$ to $L_{(\mathrm{max})}^{(C)}$, for a given C. For very small (i.e., infinitesimal)

$\epsilon > 0$, define

$$L_{(\min)}^{(C)} = \frac{C}{5W} \tag{27}$$

$$L_{(\max)}^{(C)} = \frac{C+1}{5W} - \epsilon. \tag{28}$$

Thus $L_{(\max)}^{(C)}$ is the largest value for L in a corridor of width, W, and capacity, C. Within such an interval, as L increases, p_C increases in the linear congestion model. The reverse is true, however, for the exponential congestion model (p_C decreases). As the length increases infinitesimally from $L_{(\max)}^{(C)}$ to $L_{(\min)}^{(C+1)}$, there is a step increase (decrease) in p_C in the exponential (linear) congestion model.

Thus to determine the maximum length of the corridor, greater then L^*, that will give $p_C \leqslant p_{CU}$, we need to specify which kind of length, $L_{(\min)}^{(C)}$ or $L_{(\max)}^{(C)}$, we are interested in determining. Thus, given λ, W, p_{CU}, and the kind of length that interests us, we solve for C first by experimental calculations of p_{CU} for different values of C and its corresponding value of L. The C that yields the largest value of p_C such that $p_C \leqslant p_{CU}$, is the capacity we are trying to determine. The value of L that corresponds to p_C is the length of interest to us.

A program that determines the maximum length, for the kind of length that we are interested in, is also available from the authors; it is named LESL. A golden-section search over the (integer) values of C, is performed to locate C^* (and therefore L^*). Next, bisection is performed over all values of C, starting with $C^* \leqslant C \leqslant 10000$, until the value of C for max $p_C \leqslant p_{CU}$ is determined. The value of the kind of length that we are interested in, either $L_{(\min)}^{(C)}$ or $L_{(\max)}^{(C)}$, is then determined by using either eq. (27) or eq. (28), with $\epsilon = 0$ in LESL.

Suppose λ, p_{CU}, and L are specified. We would like to determine the minimum width of the corridor that would give $p_C \leqslant p_{CU}$. Note that the only effect of decreasing the width would be to increase the congestion of the traffic flow; therefore, it is obvious that p_C increases monotonically as the width is decreased.

Also available from the authors is a program, named LESW, that performs a bisectional search for $C_{\min}$. Using eq. (1), the corresponding width is then determined.

λ_{eff} ANALYSIS

Consider a corridor of dimensions L and W. We want to determine how much flow the corridor can handle on the average (i.e., over a long period of time). The effective flow rate, λ_{eff}, known as the throughput, is related to λ and p_C as shown by eq. (29).

$$\lambda_{\text{eff}} = \lambda(1 - p_C). \tag{29}$$

We know that as λ increases, p_C also increases; thus, there must be a value, λ^*, that maximizes λ_{eff}. An experimental approach can be used to find (a close

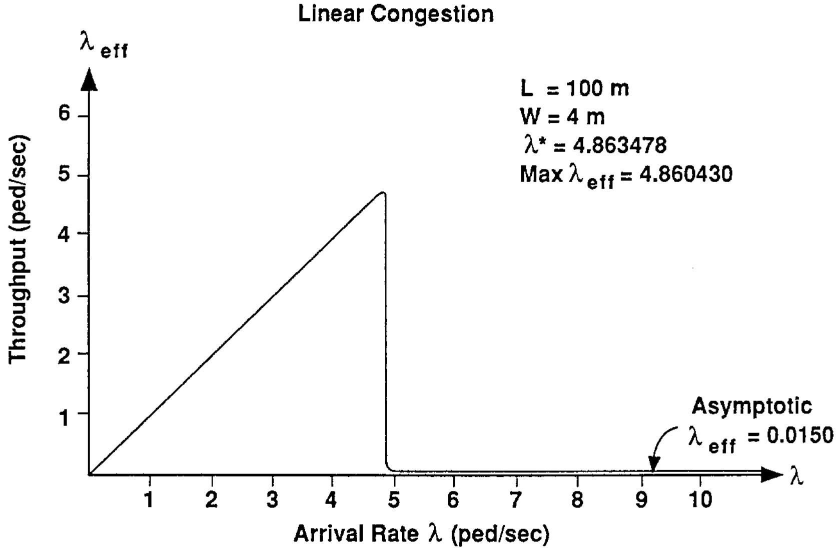

Fig. 6.

approximation to) λ^* by calculating the value of p_C for each trial value of λ, however, available form the authors is a program, named LESMEL, that performs a golden-section search for λ^*. After the size of the interval containing λ^* is reduced below the size that is specified by the stopping criterion, the λ that corresponds to the highest value of λ_{eff} is used as a close approximation to λ^*.

Some observations on the behavior of λ_{eff}, with respect to λ, may be of interest to some investigators; see figs. 6 and 7. One such observation is that the

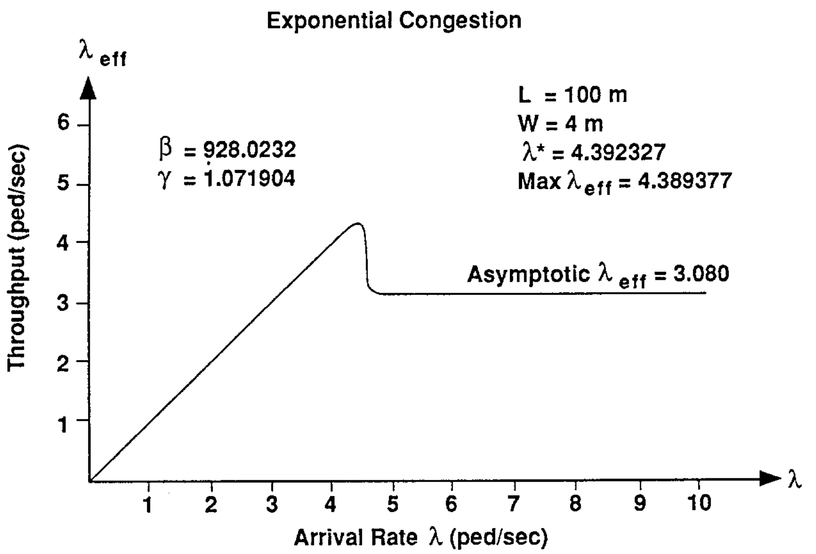

Fig. 7.

value of λ for linear congestion will drop in value much more rapidly than that of exponential congestion once the value of λ exceeds λ^*. Also, it is obvious that the asymptotic value of λ_{eff} (for very large λ) is a far more sizeable fraction of the maximum λ_{eff} for the case of exponential congestion than it is for the case of linear congestion, which is typically close to zero. This behavior, as illustrated for the case of $L = 100$ meters and $W = 4$ meters in figs. 6 and 7, is typical of all corridors having constant dimensions.

Because the contrast in the behavior of λ_{eff} between linear and exponential congestion is so significant for values of $\lambda > \lambda^*$, it may be practical to use a weighted average of the pedestrian velocities of the linear and exponential congestion models when modeling the behavior of a real situation involving pedestrians in a corridor of constant dimensions whenever $\lambda > \lambda^*$.

6. Summary and conclusions

In this paper, state dependent queueing models for capturing the congestion effects of pedestrian traffic flow within circulation systems of buildings has been proposed and developed. Linear and exponential state dependent service models were developed to account for the decreasing service rate within horizontal, single-flow circulation systems as the density of customers within the circulation system increases. Not only do the models afford an analyst the means to analyze existing circulation systems for throughputs, sojourn times, blocking probabilities, mean delays and number of customers within the system, they afford a means to design a circulation system to achieve a certain performance level by varying the arrival rate, λ, and critical design variables L, W, and C. Finally, along with their mathematical development, FORTRAN-77 computer implementations of the models described in the text are available from the authors upon request.

Future extensions of this work will investigate multi-directional traffic flows and the related design problems of *cross and t-intersections* as well as their integration into larger queueing network models of circulation systems and facilities.

References

[1] A.O. Allen, *Probability, Statistics, and Queueing Theory with Computer Applications* (Academic Press, New York, 1978).

[2] R.W. Conway and W.L. Maxwell, A queueing model with state dependent service rates, Journal of Industrial Engineering 12 (1962) 132–136.

[3] N.I.S. Foot, Pedestrian traffic flows, DMG, DRS Journal: Design Research and Methods 7, part 2 (1973) 162–167.

[4] J.J. Fruin, *Pedestrian Planning and Design* (Metropolitan Association of Urban Designers and Environmental Planners, Inc. New York, N.Y., 1971).

[5] D. Gross and C.M. Harris, *Fundamentals of Queueing Theory* (John Wiley and Sons, New York, 1985).
[6] J.D. Hankin and R.A. Wright, Passenger flow in subways, Operational Research Quarterly 9, part 2 (1958) 81–88.
[7] F.S. Hillier, R.W. Conway and W.L. Maxwell, A multiple server queueing model with state dependent service rates, Journal of Industrial Engineering 15 (1964) 153–157.
[8] L. Kleinrock, *Queueing Systems, Vol. I: Theory* (John Wiley and Sons, New York, 1975).
[9] Alec Lee, *Applied Queueing Theory* (St. Marin's Press, New York, 1966).
[10] C.A. O'Flaherty and M.H. Parkinson, Movement on a city centre footway, Traffic Engineering and Control 13 (1972) 434–438.
[11] S.J. Older, Movement of pedestrains on footways in shopping streets, Traffic Engineering and Control 10 (1968) 160–163.
[12] C.H. Sauer and K.M. Chandy, *Computer Systems and Performance Modeling* (Prentice-Hall, Englewood Cliffs, N.J., 1981).
[13] J. MacGregor Smith and W.B. Rouse, Application of queueing network models to optimization of resource allocation within libraries, JASIS 30 (5) (1979) 250–263.
[14] J. MacGregor Smith, Queueing networks and facility planning, Building and Environment 17 (1) (1982) 33–45.
[15] J.M. Smith and B. Bouanaka, Queueing network decomposition in facilities planning, Computers and Operations Research 12 (1) (1985) 1–16.
[16] J. MacGregor Smith, R.J. Graves and L. Kerbache, QNET: an open queueing network model for material handling system analysis, Material Flow 3 (1986) 225–242.
[17] K. Togawa, Study on fire escapes based on the observation of multitude currents, Report #14, Building Research Institute, Japan, 1955.
[18] P. Tregenza, *The Design of Interior Circulation* (Van Norstand Reinhold Company, New York, N.Y., 1976).