Title
Design for local Area Freight Networks

Author
Hall, Randolph W.

Publication Date
1991-05-01

Design for Local Area Freight Networks

Randolph W. Hall

May 1991
Working Paper No. 63

The University of California Transportation Center

University of California
Berkeley, CA 94720

The University of California Transportation Center

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California State Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, and Irvine; the Institute of Urban and Regional Development at Berkeley; the Graduate School of Architecture and Urban Planning at Los Angeles; and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in Center activities. Researchers at other universities within the region also have opportunities to collaborate on selected studies. Currently faculty at California State University, Long Beach, and at Arizona State University, Tempe, are active participants.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. For a list of publications in print, write to the address below.

University of California Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel: 415/643-7378
FAX: 415/643-5456

Authors of papers reporting on UCTC-sponsored research are solely responsible for their content. This research was supported by the U.S. Department of Transportation and the California State Department of Transportation, neither of which assumes liability for its content or use.

Design for Local Area Freight Networks

Randolph W. Hall

Department of Industrial Engineering & Operations Research
University of California at Berkeley

Working Paper No. 63

May 1991

The University of California Transportation Center
University of California at Berkeley

TABLE OF CONTENTS

DESIGN FOR LOCAL AREA FREIGHT NETWORKS

Randolph W. Hall
Department of Industrial Engineering & Operations Research
University of California, Berkeley CA 94720

ABSTRACT

Local area freight networks (LANs) are used to collect and distribute freight within metropolitan regions. Focusing on common carriers, this paper classifies LAN topologies, then shows how the optimal topology depends on demand characteristics. Continuous space approximations are used to analyze topologies as well as to analyze the relationship between cost and number of terminals. Key findings are summarized in Table 1.

INTRODUCTION

The spatial properties of communication and transportation networks have much in common. The former transmit data between terminals, computers and devices (nodes) via wire, optical cables and microwave (arcs). The latter transport goods and people between terminals, factories, residences and retailers (nodes) via roads, railroads, waterways and airways (arcs). Both network types are designed to optimize tradeoffs between operating costs for transmission/transportation and investment costs for constructing the network. In one extreme, if network flows are very small, the design problem resembles the minimum spanning tree problem, where the sole objective is to minimize the cost of building the network. In another extreme, if network flows are very large, the sole objective is to minimize transmission/transportation costs, so direct connections are constructed between all node pairs (see Prager, 1959, for example).

The similarity between freight networks and communication networks is especially striking. On a nationwide scale, goods/data are transported over wide-area-networks (WAN) that connect metropolitan regions. Typically, the WAN provides direct routes between gateway terminals situated within metropolitan regions. On a regional scale, goods/data are transported over local-area-networks (LAN), which connect local terminals to the gateway and to each other. Unlike the WAN, however, traffic flows often do not justify direct connections between local terminals. Instead, interregional flows may be concentrated through the gateway to save on network investment costs, at the expense of transmission/transportation costs. On a smaller scale, goods are transported to local terminals via multiple stop pickup and delivery routes. From the design perspective, these routes resemble the ring topology used in computer LANs.

The objective of this paper is to identify principles for the design of LANs for transporting freight, with emphasis on common carriers. The issues raised in this analysis are similar to those raised in the design of LANs for transmitting data. However, some aspects of the cost structure are unique to transportation and, therefore, the results are not necessarily transferable.

The methodology follows from work by Daganzo (1987,1990), and Hall (1987). These papers used continuous space approximations to identify near-optimal designs for wide-area-networks. The primary limitation of these works is that vehicle size is assumed to be identical on all network links, and vehicles are assumed to be filled to capacity. While these assumptions are appropriate for WANs, they are not for LANs. Within a LAN, different vehicle types are invariably used for P&D and interterminal operations, due to time restrictions on the length of P&D routes.

Daganzo and Newell (1986) studied networks with different vehicle types. However, their research is limited to many-to-one and one-to-many traffic patterns. The questions of how to efficiently consolidate shipments from small vehicles to large vehicles and how to create a hierarchy of consolidation terminals were examined, but the issue of how to sort and distribute many-to-many shipments was not. In the following sections, the design of a LAN for a many-to-many traffic pattern, with different vehicle sizes, is analyzed.

The remainder of the paper is divided into five parts. The first part describes the design problem and introduces four network topologies. The second part presents a cost model for transporting freight across an arc in a LAN. The third develops systemwide cost models. The fourth section optimizes the cost models and interprets the results. Finally, a fifth section covers adjustments to the model.

A key assumption of this paper is that the network design, as defined by the placement of terminals and the routing between terminals, is held constant over a reasonably long period of time (a month or more). During this period, traffic flows are likely to vary, which are accommodated by changing the number of vehicles dispatched per day over the routes. Therefore, arc cost will reflect the average daily cost, which accounts for these variations.

NETWORK TOPOLOGIES

The local area network (LAN) serves a metropolitan **region** (hereafter, referred to as just region) from a set of local pickup and delivery (P&D) terminals and a gateway terminal. The **P&D terminals** serve as bases for the trucks that retrieve shipments from shippers and deliver shipments to receivers. Each P&D terminal retrieves all pickups that originate in its unique **territory**, and delivers all pickups that are destined for its territory. The **gateway terminal** serves as the consolidation point for all shipments leaving or entering the region. The gateway terminal may also serve as an intermodal terminal, for transshipment between truck and air or between truck and rail. An **internal shipment** will refer to a shipment that has both origin and destination within the region, and an **external shipment** will refer to a shipment that travels between regions.

Each external shipment must travel through both its P&D terminal and the gateway terminal before leaving, or after entering, the region. In addition, the shipment will likely be processed at other terminals over a wide-area network that connects different regions. The design of the wide-area network is outside the scope of this paper (see Daganzo, 1987; or Hall, 1987).

The network topology for transporting internal shipments can be any of several types. A shipment can travel from its pickup terminal to the gateway terminal to its delivery terminal (Figure 1a, a hub-and-spoke, or star, design); a shipment can travel direct from its pickup terminal to its destination terminal (Figure 1b, a complete network); or, a shipment can travel via multiple-stop collecting/peddling routes (variations of star and complete, Figures 1c and 1d). Finally, if a shipment's origin and destination fall in the same territory, and pickups are sorted at the P&D terminal, then the shipment can forego interterminal transportation.

The above description is not exhaustive. It precludes several options, such as direct transportation from origin to destination, or direct delivery from P&D terminals to destinations located outside their territories. From an organizational standpoint, such alternatives are difficult to control and not commonly practiced by common carriers, whose shipments tend to be small.

The relevant costs for operating the LAN include terminal costs and transportation costs. Inventory costs are excluded because shipping frequency is governed by daily work cycles and not by network design, and because inventory costs are not directly borne by the common carrier. Shipping less frequently than once per day would result in lost business; shipping more frequently than once per day would provide minimal benefit to customers. Hence, vehicles are dispatched once a day across each link in the network.

Terminal costs can be divided into those that are fixed with respect to traffic volume and those that increase as volume increases. The latter primarily consists of sorting and handling costs, and the former, real estate and administrative costs. Transportation costs can be divided into interterminal and P&D costs. These can be further divided into line-haul and

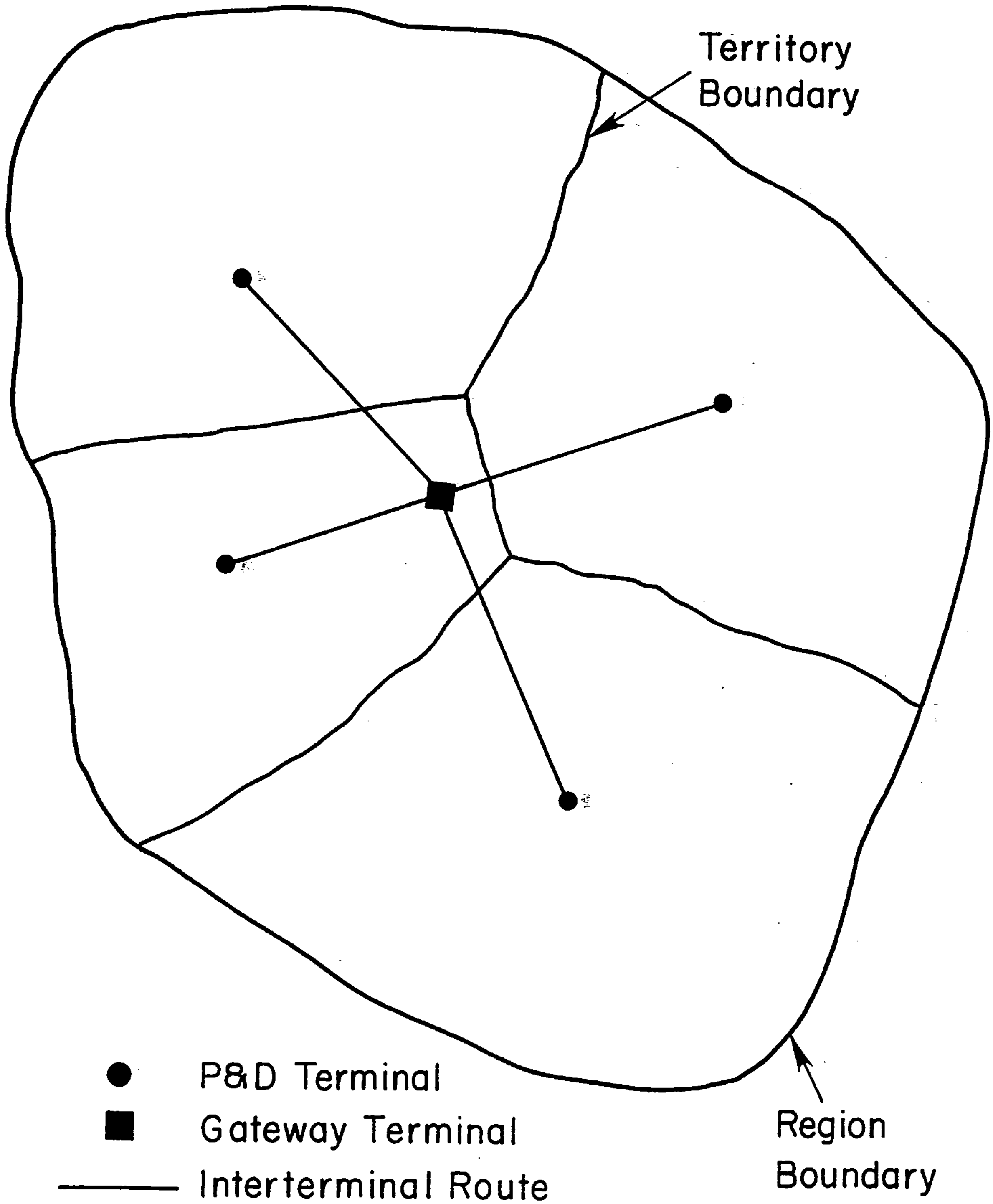

Figure 1a. Hub-and-spoke (star) network topology.

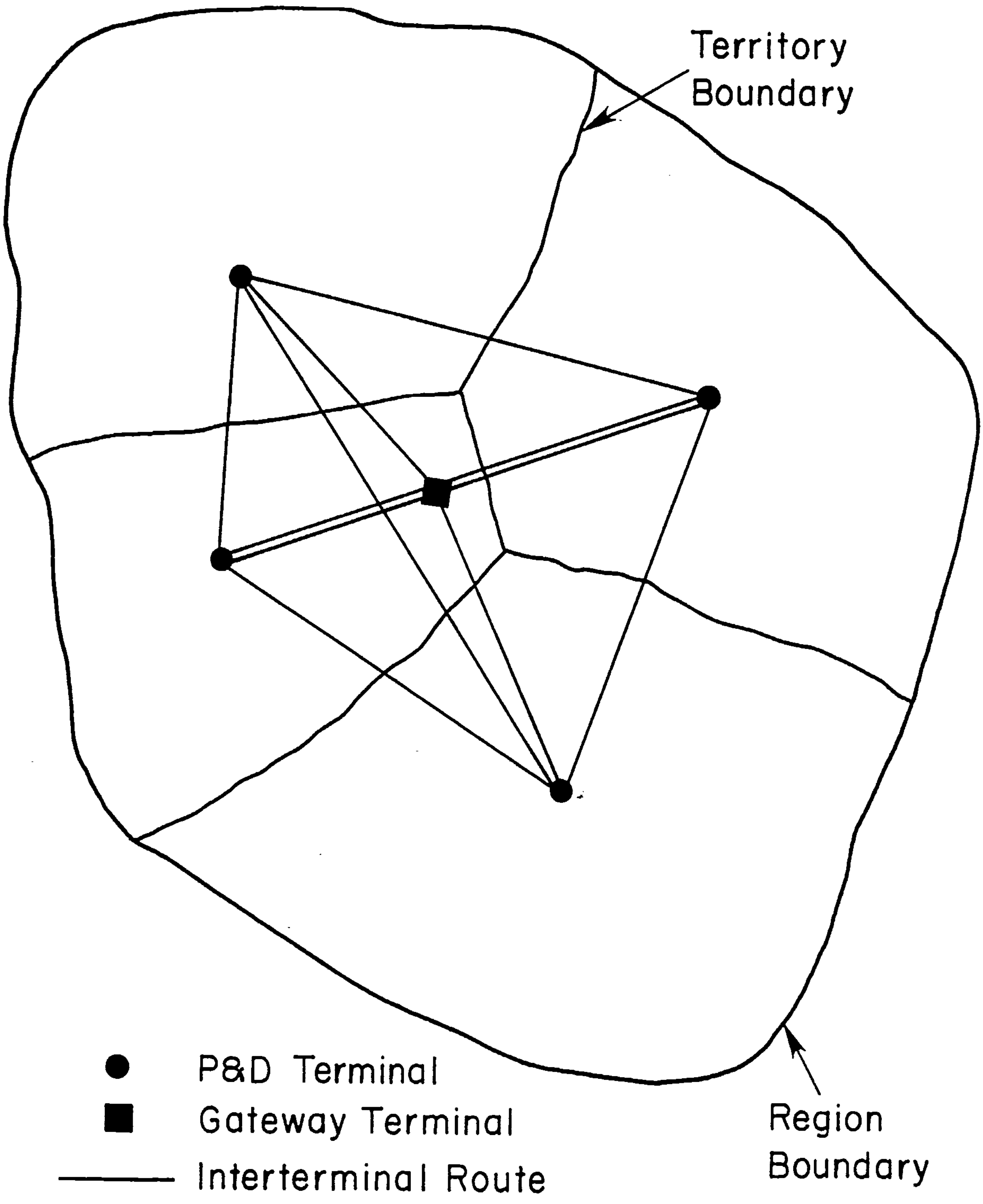

Figure 1b. Complete network topology

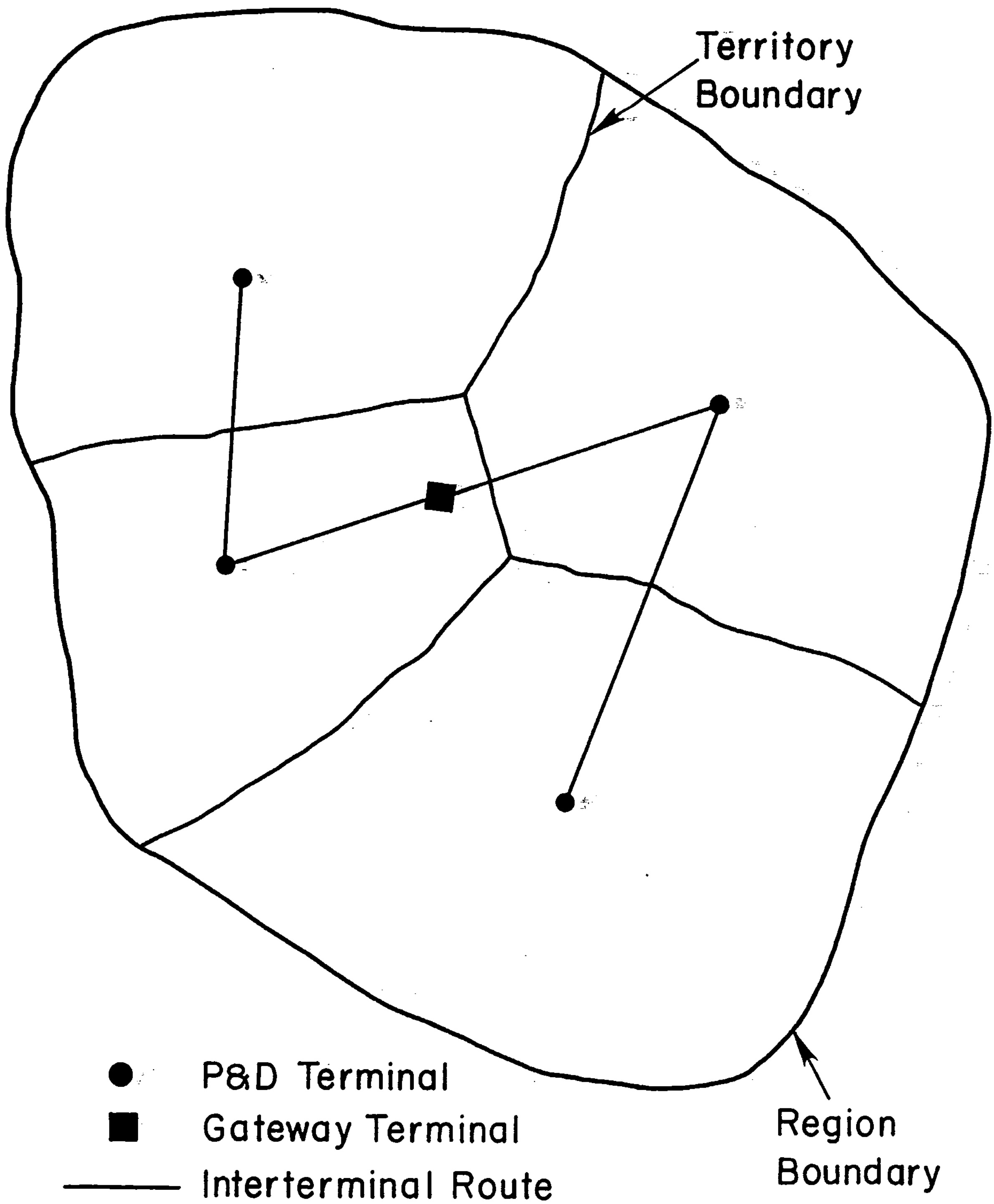

Figure 1c. Hub-and-spoke (star) topology with multiple stop routes.

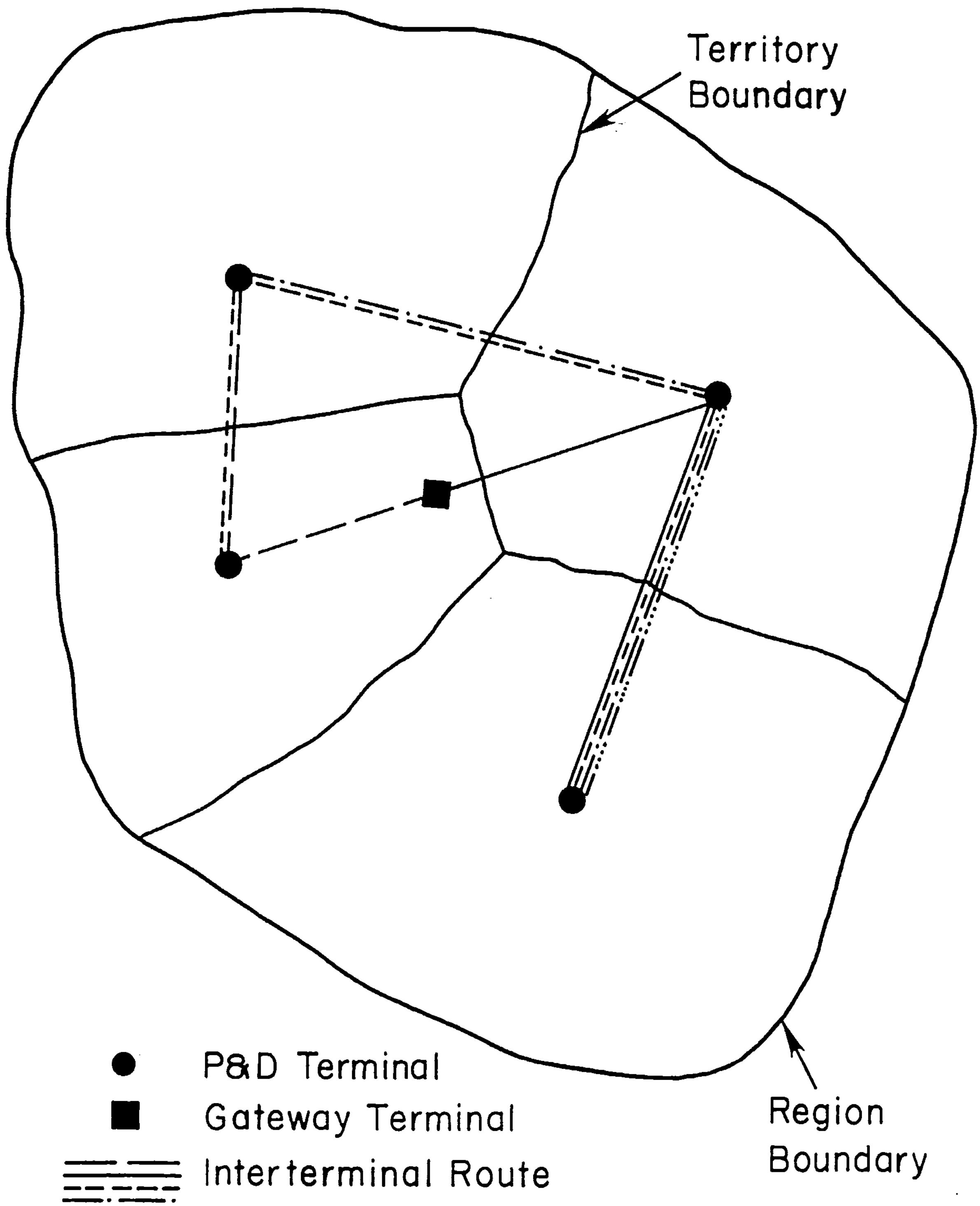

Figure 1d. Complete topology with multiple stop routes
(pickups not mixed with deliveries)

local costs, the former being a function of direct distance and load size, the latter being a function of inter-stop spacing and stop time.

The optimal LAN design will be defined by the number of P&D terminals and the topology. In the following sections, cost models are developed, then optimized, to determine the best design as a function of demand attributes.

TRANSPORTATION COST ON A ROUTE

Assume that shipment size and vehicle size can be measured with a single attribute, such as volume. Further, assume that whenever a route is served, that enough vehicles are available to allow all waiting shipments to be transported simultaneously at minimum feasible cost.

The cost of transporting the freight depends on the number, and type, of vehicles required to accommodate the freight. For example, if the volume is very small, then a van might be sufficient; if it is somewhat larger then a fixed truck might be called for; and if it is larger still then a tractor-trailer rig might be needed. If the volume is extremely large, then multiple trucks would be dispatched. Overall, the relationship between cost and volume, C(v), would resemble Figure 2, which shows discrete steps in cost as new vehicle combinations are mandated, as well as a general decline in cost/volume as volume increases (due to larger, more cost efficient, vehicle sizes).

If the route structure is fixed over some time period, such as a week, month or year, then volume, and cost, are sure to vary from day to day. The expected cost incurred per day could be calculated as a function of mean daily volume:

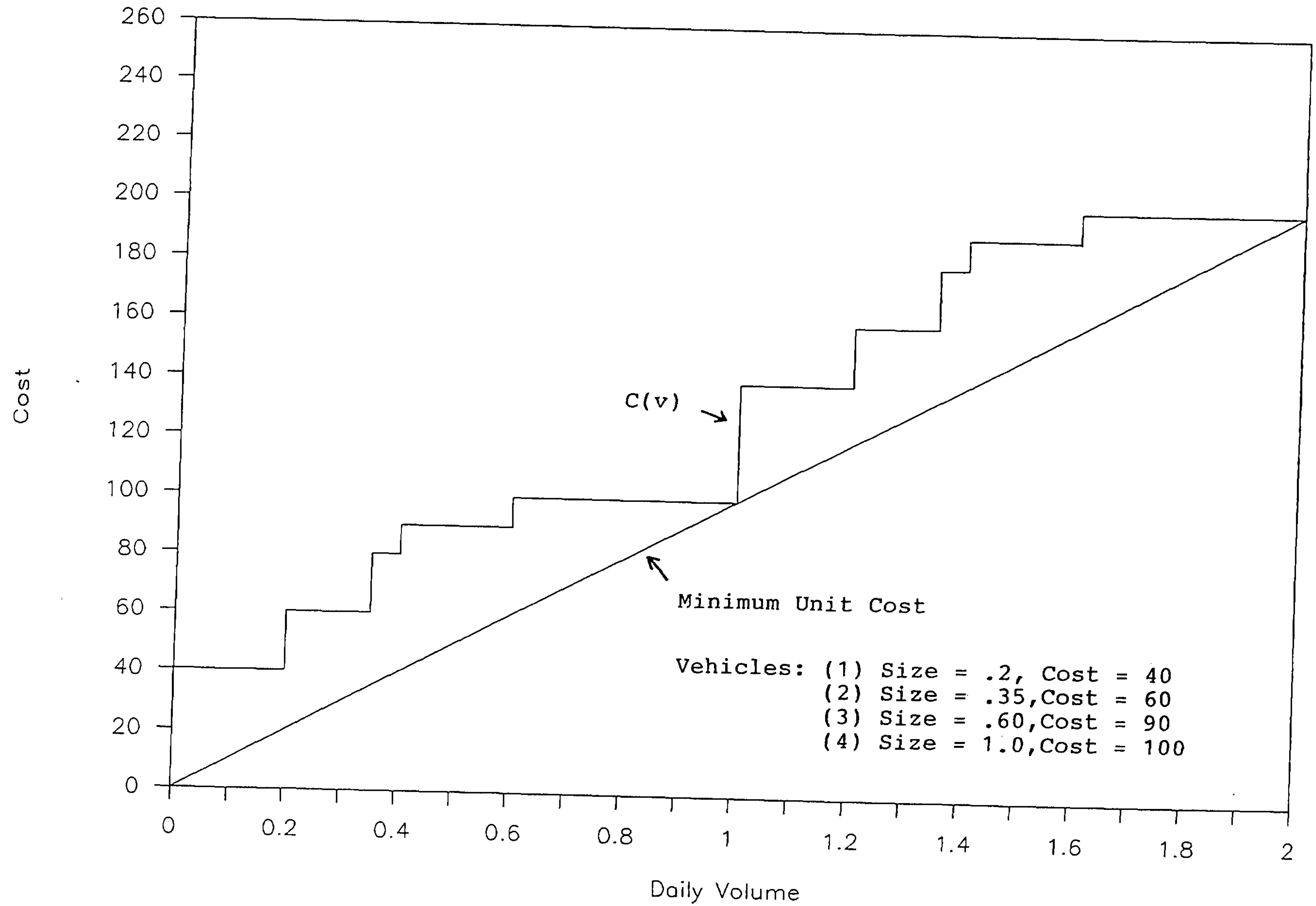

Figure 2. Cost versus daily volume, with four vehicle sizes and deterministic volume.

$$\tilde{C}(\bar{v}) = \int_0^{\infty} C(v)f(v,\bar{v})dv , \tag{1}$$

where $f(v,\bar{v})$ is the probability density function for the volume per day, when the mean daily volume is $\bar{v}$. Random variations cause the steps in $C(v)$ to be smoothed out, as in Figure 3 (which applies to a normal volume distribution). In the deterministic case, a small increase in volume either causes no cost change or a very large cost change. When randomness is accounted for, the marginal cost falls in a much narrower range. The range is even smaller for the incremental cost (the cost of changing the flow by some finite, non-differential, quantity), as shown in Figure 4.

Referring to the figures, $\tilde{C}(\bar{v})$ behaves in a rather simple way when the coefficient of variation in traffic volume is large, and a single vehicle type is available:

$$\tilde{C}(\bar{v}) \approx \begin{cases} a, & \bar{v} \leq s/2 \\ a/2 + a(\bar{v}/s), & \bar{v} > s/2 \end{cases} , \tag{2}$$

where: a = cost of dispatching vehicle

s = vehicle size .

Eq. 2 will be adopted as the route cost model in later portions of this paper.

When several vehicle types are available, the behavior is more difficult to characterize for values of $\bar{v}$ below the capacity of the largest vehicle. However, for large values of $\bar{v}$, cost can be approximated by the following, which assumes that $v - [v/s]^-$ is uniformly distributed over $[0,s]$:

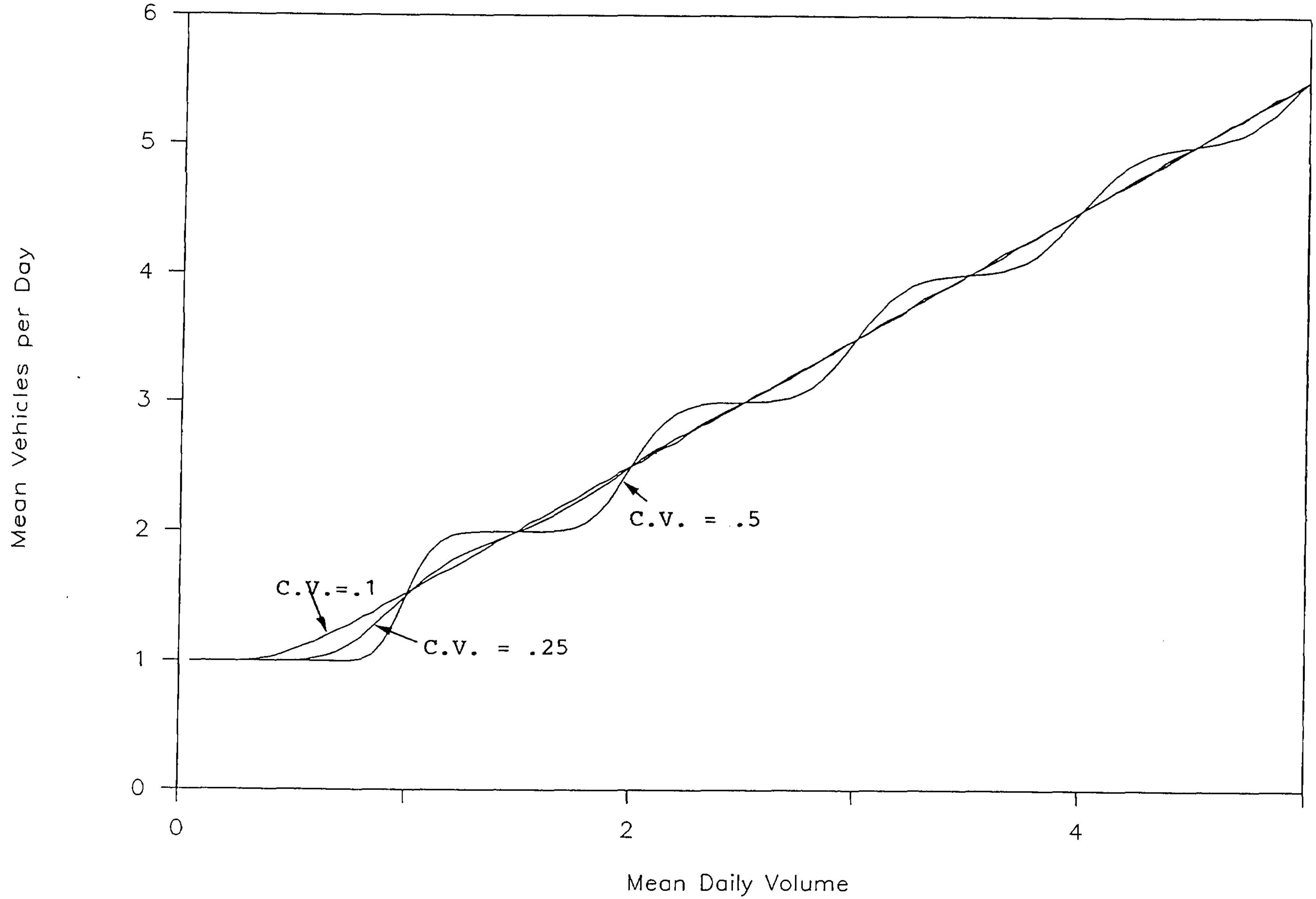

Figure 3a. Mean Vehicles per Day versus Mean Daily Volume, Normally Distributed Daily Volume with Constant Coefficient of Variation.

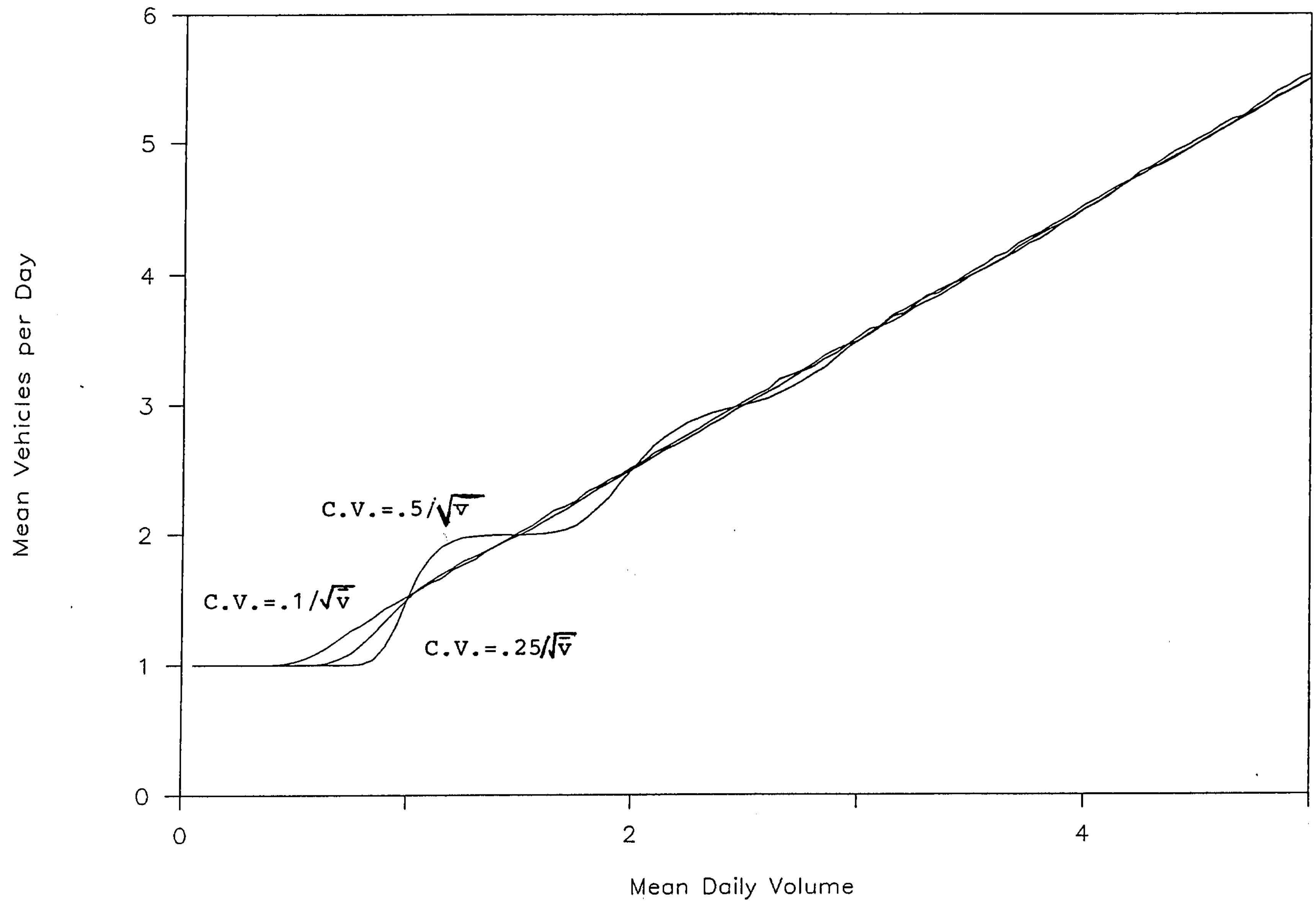

Figure 3b. Mean Vehicles per Day versus Mean Daily Volume: Normally Distributed Daily Volume with Coefficient of Variation Inversely Proportional to Square-root of Mean Daily Volume

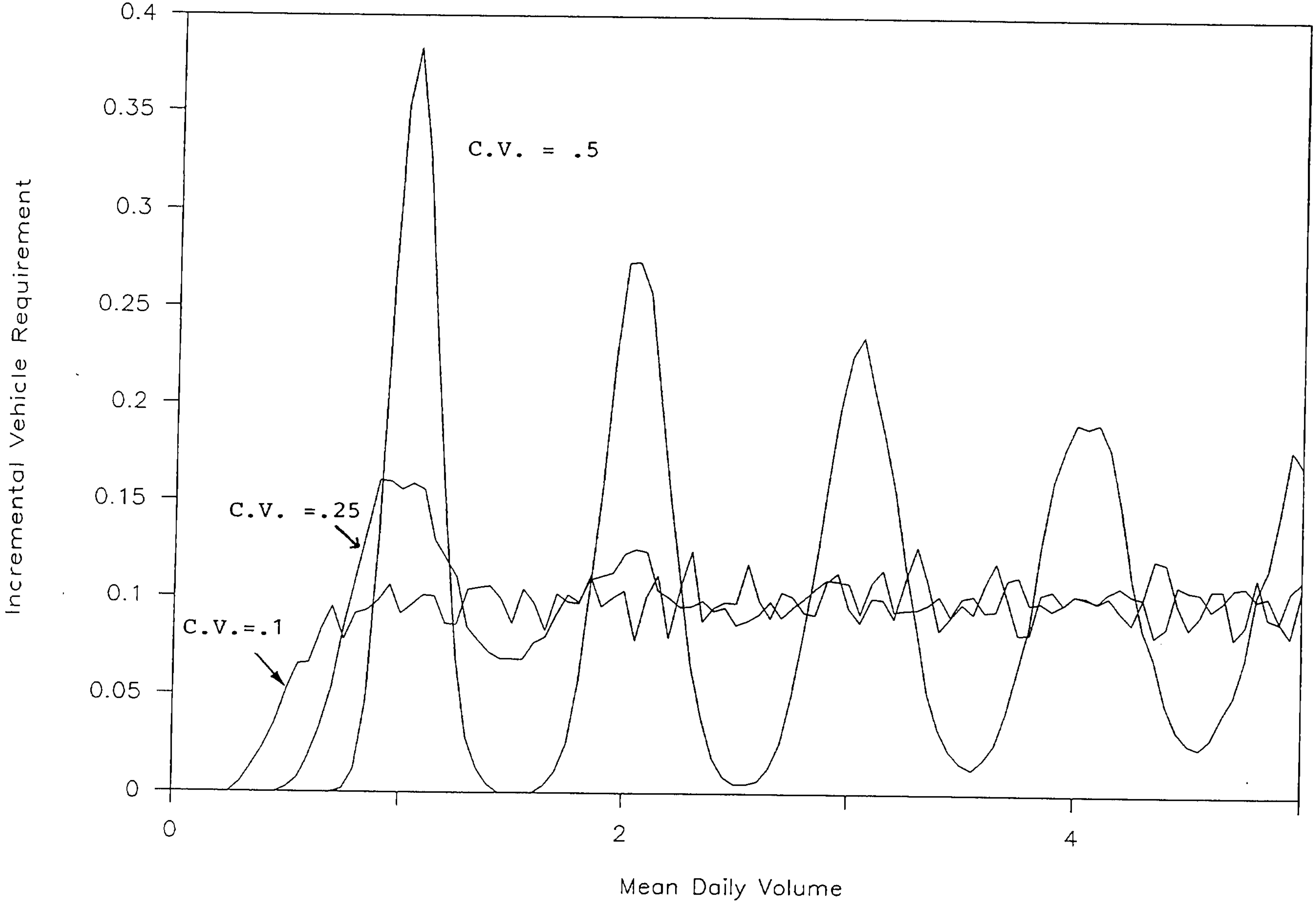

Figure 4a. Incremental Vehicle Requirement when Mean Daily Volume is Increased by .1 Vehicle Loads: C.V. constant.

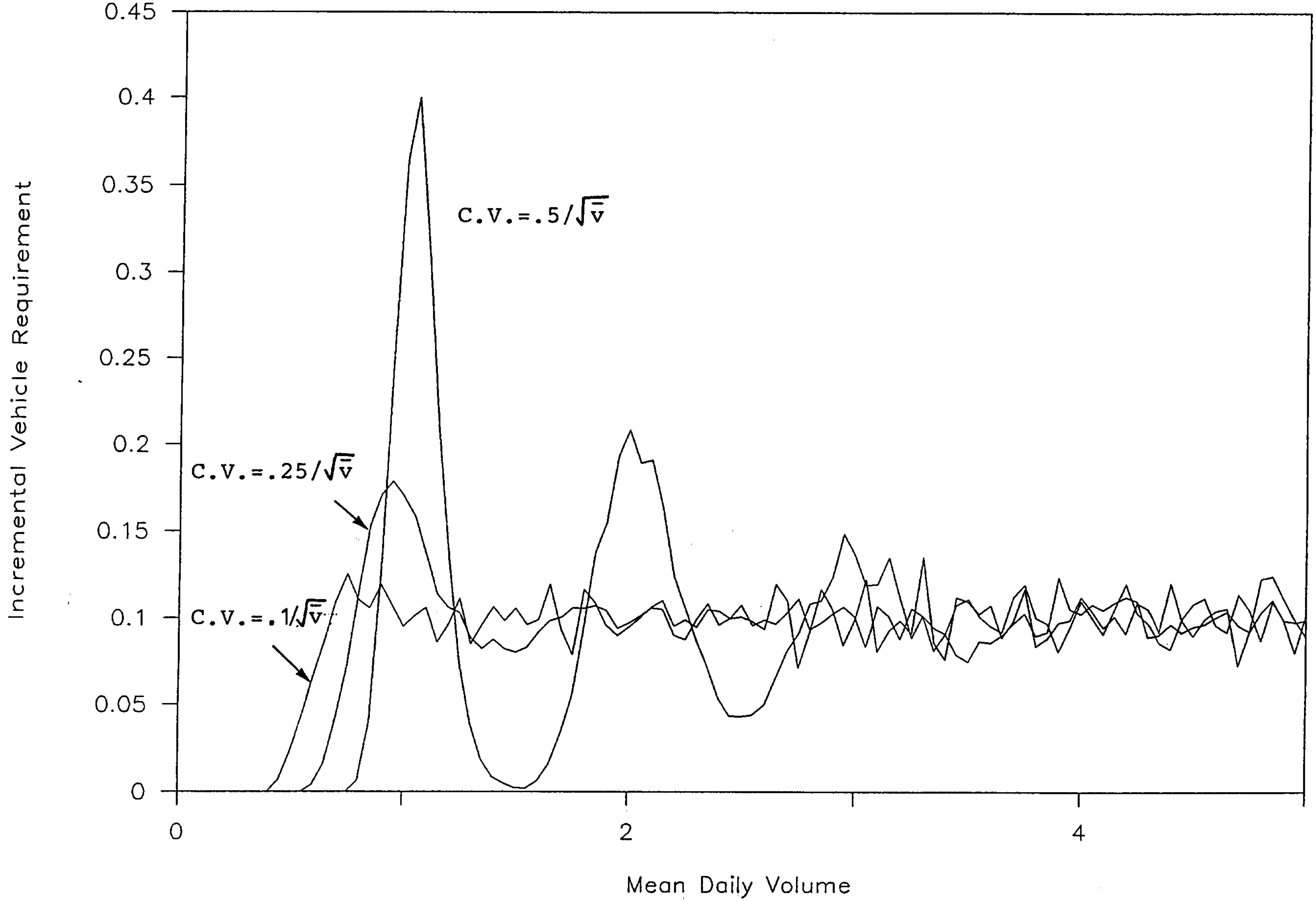

Figure 4b. Incremental Vehicle Requirement when Mean Daily Volume is Increased by .1 Vehicle Loads: C.V. Inversely Proportional to Square-root of Mean Daily Volume.

$$\tilde{C}(\bar{v}) \approx \left[(1/s)\int_0^s [C(v)-a(v/s)]dv\right] + a(\bar{v}/s) , \quad \bar{v} > s , \qquad (3)$$

where s and a are interpreted as the size and cost for the largest, and most cost efficient, vehicle available. In either case, for $\bar{v} > s$, total cost is approximated by the sum of a term that is linear with respect to $\bar{v}$, and a term that is fixed with respect to $\bar{v}$ (which represents the excess cost from failing to utilize the most cost efficient vehicle to full capacity).

It may be that the number of vehicles sent across a route on a given day is defined by the maximum of two volumes, each representing a direction of travel. If each volume is normally distributed, then the maximum is approximately normal with mean given by (Clark, 1961):

$$\mu = \bar{v}_1 + (\bar{v}_2 - \bar{v}_1)\Phi(\alpha) + a\phi(\alpha) \qquad (4)$$

where: $\bar{v}_i$ = mean volume direction i

$a = (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}^2)^{.5}$

σ_i^2 = variance of volume, direction i

σ_{12}^2 = covariance in volume

$\alpha = (\bar{v}_2 - \bar{v}_1)/a$,

and where $\Phi(\alpha)$ and $\phi(\alpha)$ are the normal distribution and normal density functions, respectively.

In the special case where the volumes are independent and identically distributed, $\mu = \bar{v}_1 + .56\sigma_1$, and the relationship between cost and volume behaves like Figure 3, excepting that μ replaces $\bar{v}_1$ and σ is modified to approximate the standard deviation of the maximum (see Clark, 1961). If the volumes are not identically distributed, then the relationship between cost and $\bar{v}_1$ would only resemble Figure 3 if $\bar{v}_1 >> \bar{v}_2$ (i.e., if direction one is dominant). If $\bar{v}_1 << \bar{v}_2$, then cost would be nearly constant with respect to $\bar{v}_1$, but as $\bar{v}_1$ approaches $\bar{v}_2$, cost would increase at an increasing rate.

SYSTEM COST MODELS

In this section, cost models are formulated for each topology, then optimized with respect to number of P&D terminals. The optimal topology is later found as a function of demand attributes by comparing the optimized costs.

On a systemwide basis, increasing the number of P&D terminals acts to increased fixed terminal costs. It also tends to increase interterminal transportation costs, due to the increased number of routes required (this cost is offset somewhat by reduced travel distance). However, these penalties must be balanced against decreased P&D line-haul costs, the principal benefit of adding P&D terminals. This basic tradeoff holds true for all topologies, though the exact costs vary from case to case.

To highlight the key design issues, the cost analysis will make the following assumptions. It is not difficult to relax the assumptions to analyze alternative scenarios.

- Traffic flows are balanced between all terminal pairs, so that empty backhauls are not required on interterminal routes.
- The amount of freight entering the region equals the amount leaving.
- A single vehicle type is available for interterminal transportation and a single type is available for P&D transportation (not necessarily the same).
- Pickups are never mixed with deliveries in the vehicle.
- Local delivery vehicles must return to their P&D terminal before making pickups, but interterminal vehicles pickup freight in the backhaul direction.

- The gateway terminal does not also serve as a P&D terminal (as is reasonable if the gateway is intermodal).
- Demand is uniformly distributed over the region.

The following symbols are used in the models:

T = number of P&D terminals in metropolitan region
A = area of metropolitan region
J = number of systemwide interterminal trips completed per day
M = capacity of interterminal vehicle, as ratio to P&D vehicle capacity.
N = number of pickup tours required per day
(= number of delivery tours required per day)
P = proportion of freight picked up that is internal
(= proportion of freight delivered that is internal)
$\bar{S}$ = average number of stops per trip
X = number of external terminals connected to gateway
α_1 = cost/distance per vehicle for P&D linehaul
α_2 = cost/distance per vehicle for interterminal travel
β_1 = fixed cost for initiating a route
β_2 = cost for each delivery or pickup stop on route
γ_1 = fixed cost of owning and operating a terminal, per day
γ_2 = variable cost of handling unit of freight at a terminal per P&D load quantity
γ_3 = terminal sorting cost parameter.

The following sections present models for four cost components: (1) P&D transportation cost (C_1), (2) Interterminal linehaul cost (C_2), (3) Interterminal stop costs (C_3), and (4) Terminal costs (C_4).

Pickup & Delivery Transportation Cost

The P&D transportation cost is the same for all topologies, for any fixed number of P&D terminals. The cost is the sum of the local and linehaul costs. The former does not change appreciably as the number of terminals (T) changes, and will be excluded from the analysis. The line-haul cost, on the other hand, does depend on T, and equals twice the average linehaul distance (allowing for an empty movement), multiplied by the cost per unit distance, multiplied by the number of trips/day (Daganzo, 1984):

$$C_1 = \text{linehaul P\&D cost/day} \approx 2\alpha_1 (k_1 \sqrt{A/T}) \cdot N, \qquad (5)$$

where: k_1 is a constant defined by the travel metric ($k_1 \approx .38$ for Euclidean with centrally located P&D and uniform spatial demand distribution.)

If the gateway is not centrally located, more complicated expressions can be employed (see Eilon, et al, 1971).

If routes are constrained not to exceed the length of a driver's workshift, then any decrease in T will increase the linehaul P&D time, and reduce the time available in a workshift for local P&D. The result is that N must increase as T decreases, in the following way:

$$N \approx \frac{\tilde{N}t}{\tau - 2(k_1/\nu)\sqrt{A/T}} = \frac{\text{Local Stop Time}}{\text{Local time Available/Route/Day}}, \qquad (6)$$

where: $\tilde{N}$ = number of stops made per day in region
τ = length of workshift
t = local time per stop
ν = linehaul velocity.

Ordinarily, N is insensitive to changes in T and can be treated as a fixed value, as will be done later in the paper.

Interterminal Linehaul Cost

The interterminal cost varies among topologies. Each route must be covered at least once per day, and perhaps more than once if the volume fills multiple vehicles. Let $\bar{L}$ represent average one-way route length. Then the interterminal linehaul cost is approximated by substituting $\bar{L}\alpha_2$ for a in Eq. 2, and multiplying the result by the number of routes.

Star Topology If single-stop routes are employed, the number of routes is twice the number of terminals. $\bar{L}$ is nearly independent of T, except in the special case where there is just one P&D terminal (in which case $\bar{L}$ can be as low as zero if the gateway coincides with the P&D). For a centrally located gateway, uniform demand and the Euclidean metric, $\bar{L}$ is approximately $.38\sqrt{A}$ (allowing for a loaded backhaul). The volume per route, measured in interterminal vehicle loads, is N/MT, because M pickup tours feed each P&D to gateway route and M delivery tours serve each gateway to P&D route. Substituting the appropriate parameters in Eq. 2 and multiplying by the number of routes (2T):

$$C_2^S \approx \begin{cases} 2[(N/M) + T/2]\alpha_2\sqrt{A}(.38), & T/2 \leq (N/M) \\ 2T\alpha_2\sqrt{A}(.38), & T/2 > (N/M) \end{cases} . \qquad (7)$$

If vehicles make multiple stops, with an average of $\bar{S}$ pickups on trips to the gateway and an average of $\bar{S}$ deliveries on the return, then the number of trips is reduced by a factor of $\bar{S}$, and:

$$C_2^S \approx \begin{cases} 2[(N/M) + (T/2\bar{S})]\alpha_2\sqrt{A}(.38), & (T/2\bar{S}) \leq (N/M) \\ 2(T/\bar{S})\alpha_2\cdot\sqrt{A}(.38), & (T/2\bar{S}) > (N/M) \end{cases} . \qquad (8)$$

Complete Topology Interterminal routes can be classified into P&D to gateway routes, and P&D to P&D routes. If single-stop routes are employed, there are 2T routes of the former type and T^2 routes of the latter type (though T of these routes begin and end at the same terminal; hence, they have zero length and zero cost). Average route length for the former is the same as the average for the star topology, $.38\sqrt{A}$. Average route length for the latter varies as the number of terminals increases, due to more terminals being located near to the fringe of the region. $\bar{L}$ is approximated by $(k_2 - k_3/T)\sqrt{A}$ ($k_2 \approx .51$ and $k_3 \approx .51$ for the Euclidean metric and uniform demand; Hall, 1984). Substituting the appropriate parameters in Eq. 2:

$$C_2^c = C_{2pd}^c + C_{2g}^c, \quad \text{where} \tag{9a}$$

$$C_{2pd}^c \approx \begin{cases} [(N/M)P + T^2/2]\cdot\alpha_2\cdot\sqrt{A}(.51-.51/T), & T^2/2 \geq (N/M)P \\ T^2\cdot\alpha_2\cdot\sqrt{A}\cdot(.51-.51/T), & T^2/2 < (N/M)P \end{cases} \tag{9b}$$

$$C_{2g}^c \approx \begin{cases} 2[(N/M)(1-P) + T/2]\alpha_2\sqrt{A}(.38), & T/2 \leq (N/M)(1-P) \\ 2T\alpha_2\cdot\sqrt{A}(.38), & T/2 > (N/M)(1-P) \end{cases} . \tag{9c}$$

If multiple stop routes are employed, then the T^2 terms would be replaced by $(T^2/\bar{S}_{pd})$ in Eq. 9b and the T terms would be replaced by $T/\bar{S}_g$ in Eq. 9c, where $\bar{S}_{pd}$ is the average number of stops per P&D/P&D trip, and $\bar{S}_g$ is the average number of stops per P&D/gateway trip.

Interterminal Stop Costs

In addition to the linehaul distance cost, a fixed charge is incurred per route and per delivery stop. For multiple stop routes, an interterminal distance cost is also incurred. With J trips per day, these costs amount to:

$$C_3 \approx [\beta_1 + \beta_2\bar{S} + \alpha_2 f(\bar{S},T,A)]\cdot J, \tag{10}$$

where: $f(\bar{S},T,A)$ = added local distance, with $\bar{S}$ stops per trip over a region of size A with T terminals.

If single-stop routes are employed between terminals, then $\bar{S}$ would equal 1 and $f(\bar{S},T,A)$ would equal zero. For larger values of $\bar{S}$, without mixed pickups and deliveries, each stop added to a tour increments route length by an amount comparable to the terminal spacing, $\sqrt{A/T}$. Therefore, $f(\bar{S},T,A)$ is roughly $(\bar{S}-1)\sqrt{A/T}$.

The number of trips is approximated from the model of Eq. 2. For the star topology, the number of trips is:

$$J^S \approx \begin{cases} 2[(N/M) + (T/2S)], & (T/2S) \leq (N/M) \\ 2(T/S), & (T/2S) > (N/M) \end{cases} . \tag{11}$$

For the complete topology, the number of trips equals the sum of the number of P&D/P&D trips and the number of P&D/gateway trips:

$$J^c = J^c_{pd} + J^c_g \tag{12a}$$

$$J^c_{pd} \approx \begin{cases} (T-1)/T[(N/M)P + T^2/2S_{pd}], & T^2/2S_{pd} \geq (N/M)P \\ T(T-1)/S_{pd}, & T^2/2S_{pd} < (N/M)P \end{cases} , \tag{12b}$$

$$J^c_g \approx \begin{cases} 2[(N/M)(1-P) + T/2S_g], & T/2S_g \leq (N/M)(1-P) \\ 2T/S_g, & T/2S_g > (N/M)(1-P) \end{cases} . \tag{12c}$$

Terminal Costs

Administrative and real estate costs for operating a terminal are fixed with respect to volume, while handling costs tend to increase, at a nearly constant rate, as volume increases. In addition, sorting costs are nearly linear with respect to volume. But the sorting cost also depends on the number of categories that the shipments are divided into. As the number of categories increases, the sorting cost should increase, at a decreasing rate. In particular, increasing the number of categories by a power of k should only increase cost by a multiple of k; that is, if the initial sort has n categories, the process can be repeated k times to obtain n^k categories. For this reason, a logarithmic relationship is appropriate. (A logarithmic relationship is also supported by empirical work on choice reaction time by Hick, 1952). Altogether, the cost of operating a single terminal should behave as:

$$c_4 \approx \gamma_1 + \gamma_2(v_1+v_2) + \gamma_3\{v_1[\ell n(W_1)] + v_2[\ell n(W_2)]\} , \quad (13)$$

where: v_i = quantity of freight processed per day in direction i (i = 1 is outbound and i = 2 is inbound)

W_i = number of categories that freight is sorted into, in direction i.

Equation 13 assumes that freight entering a terminal's territory is sorted separately from freight leaving, for the obvious reason that shipments are traveling to different sets of destinations.

Star Topology Each internal shipment is processed at three terminals (N·P loads total, at point of pickup, point of delivery and gateway), and each external shipment is processed at two terminals (2N(1-P) loads, at gateway and at P&D). However, shipments leaving a P&D terminal for the gateway do not have to be sorted. (The exception is if a P&D separates shipments that have destinations in its own territory, but transportation savings often do not justify this added sorting cost.) Further, if the gateway terminal sorts shipments down to the delivery route level, then the only sorting required at the P&D terminals would be to sequence stops on delivery routes. This cost is independent of both the topology and T, and will be ignored in the analysis.

In the following cost equation, sorting at the gateway is assumed to occur in two phases. First, shipments originating in the region are sorted into the external and internal categories. The internal shipments are then combined with external shipments originating from outside the region, and sorted to the delivery route level. The external shipments originating within the region are sorted separately according to destination gateway terminal. (This sorting pattern is not optimal under all conditions, though it is quite robust.) Taking these assumptions into account, systemwide cost amounts to the following equation, which is independent of $\bar{S}$:

$$C_4^S \approx \gamma_1(T+1) + \gamma_2 N(4-P) + \gamma_3 N[\ell n(2)+\ell n(N)+(1-P)\ell n(X)] \ . \quad (14)$$

Complete Topology Triple handling is eliminated under the complete topology: all shipments are handled and sorted at two terminals, with the exception of shipments whose origin and destination fall in the same territory, which are processed just once. The pickup terminal sorts shipments into T+1 categories, and the delivery terminals sorts shipments into N/T categories (on average). At the gateway, inbound shipments are sorted into T categories and outbound shipments are sorted into X categories. Taking these conditions into account, and assuming that there is a 1/T chance that an internal shipment's origin and destination are in the same territory:

$$C_4^S \approx \gamma_1(T+1) + \gamma_2 N(4-2P-P/T) + \gamma_3 N\{\ell n(T+1) + \ell n(N/T) + (1-P)\ell n(X) + (1-P)\ell n(T)\} \quad . \quad (15)$$

Terminal Cost Discussion For either topology, handling and sorting costs are independent, or nearly independent, of the number of P&D terminals. For the complete topology, sorting and handling costs increase slightly as T increases, but this change in cost should be small relative to the increased in fixed terminal costs. Handling and sorting costs for the star topology are constant with respect to T. Both conclusions rely on the assumption that significant scale economies do not exist in freight handling or sorting.

Sorting and handling costs are also insensitive to changes in the number of stops per interterminal route. Whether routes have one or more stops, freight must still be sorted down to the terminal level. The quantity of freight handled also does not depend on the number of stops per route (though using more stops per route may allow terminals to be more compact and more efficient to operate).

COST OPTIMIZATION

The total cost is the sum of the P&D transportation cost, the interterminal transportation cost and the terminal cost. The optimal cost and optimal number of terminals are found for each topology by taking the derivative of the total cost expression ($C_1+C_2+C_3+C_4$) with respect to T, and setting the result equal to zero (the dominant cost terms are convex).

However, before T is optimized, the number of stops per interterminal route will be optimized.

Stops/Route: Star Topology

The optimal value of S depends on a tradeoff between interterminal linehaul costs and interterminal stop costs. There are two regimes to each cost function, one defined for $(T/2S) > (N/M)$ and the other defined for $(T/2S) \leq (N/M)$. In the former case, both cost terms decline as S increases. In essence, this means that it cannot be optimal to operate in the first range.

In the range $(T/2S) \leq (N/M)$, volume dependent linehaul costs increase as S increases, whereas stop costs and fixed route costs decline as S increases. The optimal value of S is:

$$S^* = \left[\frac{(T/2)[\alpha_2\sqrt{A}(.38) + \beta_1]}{(N/M)[\alpha_2\sqrt{A}(1/\sqrt{T}) + \beta_2]}\right]^{1/2} . \qquad (16)$$

If (N/M) is comparable to T, and β_1 is comparable to β_2, T would have to be at least 30 before S^* exceeds one. Because 30 is a very large value for T, multiple stop routes should only be contemplated when (N/M) is small relative to T. Generally speaking, a good approximation for S^* is $\max\{1 , (1/2)[T/(N/M)]\}$.

Optimal Number of Stops for Complete

S_g, the optimal number of stops on P&D to gateway routes, is defined by Eq. 16, except that (N/M) must be multiplied by $(1-P)$ to exclude internal shipments from the calculation.

S_{pd}, the optimal number of stops on P&D to P&D routes, is defined by a tradeoff between interterminal linehaul costs and interterminal stop costs. Similar to the star topology, it cannot be optimal to operate in the regime $T^2/2S_{pd} < (N/M)P$. Instead, the optimum occurs where the inequality direction is switched. Within this range, the optimum is:

$$S_{pd}^* = \left[\frac{(T^2/2)[\alpha_2\sqrt{A}(.51-.51/T)] + T(T-1)[\beta_1/2 - \alpha_2\sqrt{A}/T]}{(PN/M)[(T-1)/T](\beta_2 + \alpha_2\sqrt{A}/T)}\right]^{1/2} . \quad (17)$$

Though Eq. 17 is more complicated than Eq. 16, the principals are similar. A good approximation for S_{pd}^* is $\max\{1 , (1/2)[T^2/(PN/M)]\}$.

Number of Terminals

Star Topology As already mentioned, total cost is minimized when $T/2S \leq N/M$; otherwise, cost could be reduced by increasing S. Within this regime, cost is minimized at:

$$T_s^* = \left[\frac{\alpha_1\sqrt{A}(.38)N}{\alpha_2\sqrt{A}(.38)/S^* + \beta_1/S^* + \beta_2 + \gamma_1}\right]^{2/3} . \quad (18)$$

Complete Topology The optimal solution for the complete topology cannot be written in simple closed form. However, it is possible to obtain simple upper bounds on T^*.

The two major factors that limit the optimal number of terminals are the fixed terminal costs and the interterminal route costs. Sorting and handling cost are insensitive to changes in T, as is the average P&D to P&D distance.

If fixed terminal costs dominate interterminal route costs, then the optimal number of terminals is approximated by T_s^*. More generally, T_s^* is an upper bound on the optimal number of terminals for the complete topology (assuming the cost insensitivities mentioned above).

If interterminal transportation costs dominate fixed terminals costs, then the optimal number of terminals is approximated by ignoring fixed terminal costs and gateway transportation costs in the optimization. More generally, the following provides an upper bound on T_c^*:

$$T_c^* \leq \left[\frac{\alpha_1 (.38)\sqrt{A}N}{\alpha_2 \sqrt{A}(.51)/S_{pd}^* + \beta_1/S_{pd}^* + \beta_2}\right]^{2/5} . \tag{19}$$

Route costs are only dominant if the optimal number of terminals is a large number. This is because route costs increase quadratically with T, whereas fixed terminal costs only increase linearly with T.

Comparison of Topologies

To understand the relationship between optimal design and demand characteristics, a breakeven analysis was completed, comparing the star to the complete topology. In all cases, an increase in P (internal proportion) benefits the complete topology. Therefore, for values of P below its breakeven point, the star topology is preferred and for values above the breakeven complete is preferred. Figures 5 and 6 show how the breakeven point changes as N (number of P&D routes) changes for several cases:

Figure 5: M is Large

α_1	α_2	β_1	β_2	γ_1	γ_2	γ_3	X	M	A
1.00	2.00	15	10	$\begin{bmatrix} 62 \\ 250 \\ 1000 \end{bmatrix}$	5	25	25	15	1000

Figure 6: M is Small

1.50	2.00	15	10	$\begin{bmatrix} 62 \\ 250 \\ 1000 \end{bmatrix}$	10	5	25	2	1000

The differences in γ_2 and γ_3 reflect freight characteristics: with M large, P&D vehicles likely make many stops to retrieve small packages or letters. On a per vehicle basis, these would be inexpensive to handle, but expensive to sort. With M small, P&D vehicles likely make few stops to retrieve large shipments, which are expensive to handle but inexpensive to sort.

Interpretation

There is one clear incentive for adding terminals to a LAN, and that is reduction in P&D linehaul cost. Principally, this must be balanced against increased interterminal cost (due to split flow among many routes) and added terminal ownership and administrative costs. As indicated in the figures, a large number of terminals is cost justified when the system carries a large volume of freight (measured in P&D loads, N), especially when γ_1 is small and the star topology is employed. In practical terms, among common carriers, the Postal Service is likely to demand the greatest number of terminals because of the large volume, and small size, of shipments carried.

The advantage to the star topology over complete comes in the elimination of sorting at the pickup depot and from concentration of interterminal flows onto fewer links, which allows more terminals to be established. As indicated by the figures, these savings are especially important when shipments are small and spread over many P&D loads (M and N are large). The star topology is also advantageous when a large proportion of the freight is external (P is small), because insufficient internal volume may be available to justify the complete topology. Finally, star is advantageous when γ_1 is small, because

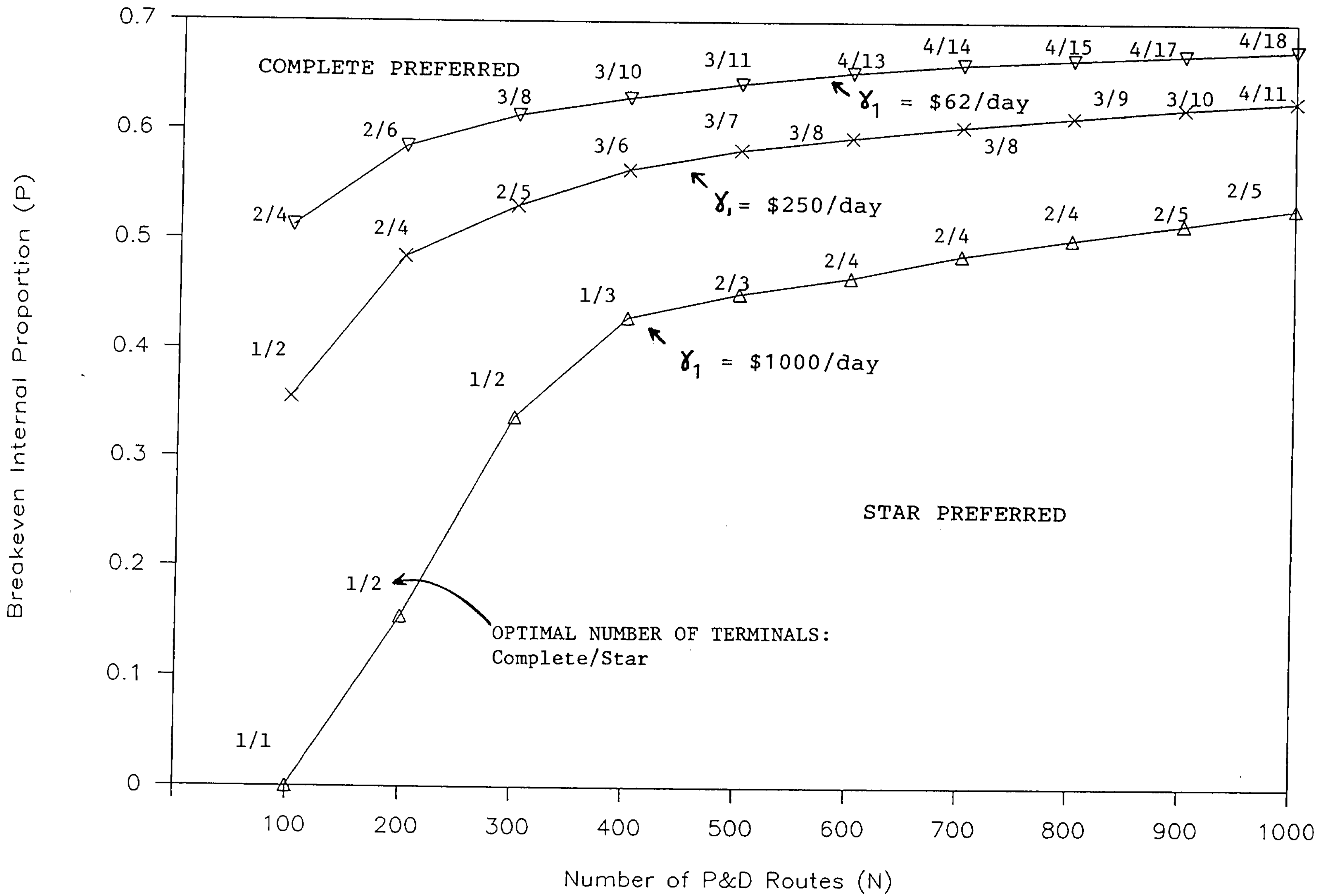

Figure 5. Breakeven Lines Between Star and Complete Topologies: Large M (interterminal capacity >> P&D capacity)

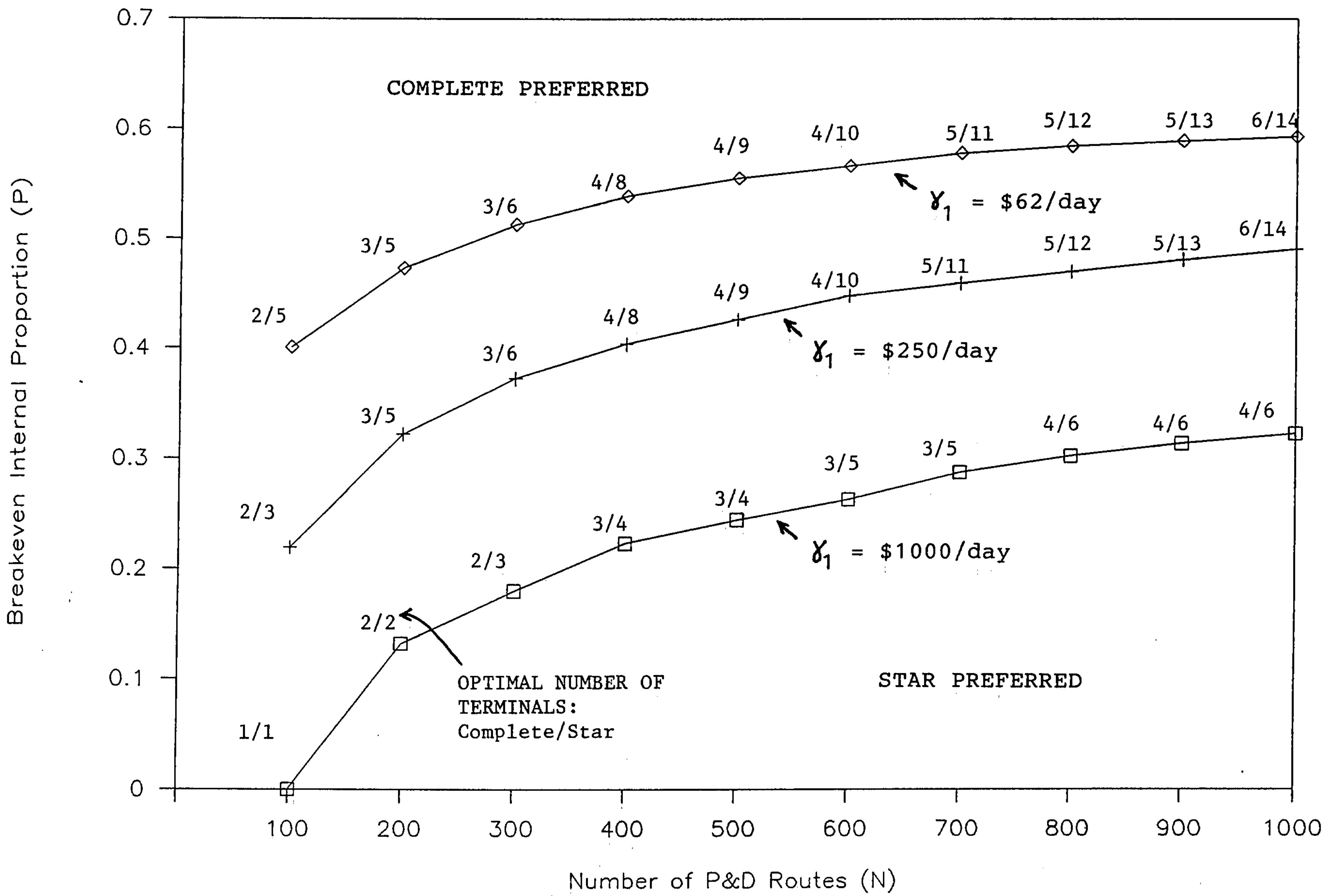

Figure 6. Breakeven Lines Between Star and Complete Topologies: Small M (interterminal capacity is twice P&D capacity)

more terminals can be built. In practical terms, the star topology is justified for nationwide systems that handle letters or small packages, such as the Postal Service or Federal Express.

The advantage of the complete over star comes in the reduction in terminal handling and from the reduction in interterminal transportation distance. These savings are especially important when shipments are large and concentrated into a small number of P&D loads (M and N small), and when a large proportion of shipments is internal (P small), in which case both handling and volume related transportation costs are significant. In practical terms, the complete topology is cost justified for LTL carriers.

One thing to note for both topologies is that the optimal number of terminals is very insensitive to changes in the region size (A). A enters as a square-root in both the numerator and denominator of Eqs. 18b, 18c, 19b and 19c, which are in turn raised to the 2/3 or 2/5 power. T^* is much more sensitive to changes in $\alpha_1 N/\alpha_2$, which represents the relative magnitudes of P&D and interterminal linehaul costs, and the ratio $\alpha_1 N/\gamma_1$, which represents the relative magnitudes of P&D linehaul cost and fixed terminal cost.

In the United States, urban areas have experienced increased congestion in recent years, which has caused travel times and travel costs to increase. Most of the delays fall within a few hours of the day, from about 6:00 to 9:00 in the morning and 3:00 to 6:00 in the evening. This is when P&D vehicles are on the road. On the other hand, interterminal runs usually occur overnight. Therefore, the ratio $\alpha_1 N/\alpha_2$ is likely to increase over time, providing an incentive for adding terminals to LANs. However, the relationship between T^* and travel time is less than linear, so changes in the number of terminals should not be dramatic.

ADJUSTMENTS TO GENERAL SOLUTION

The cost optimization identifies a minimum cost design, subject to the condition that all P&D terminals are served identically. In reality, some P&D terminal pairs should be served direct, others through the gateway terminal and still others via multiple stop routes. Unfortunately, to optimize all terminal pairs simultaneously, and account for cost interactions, is a problem of massive proportions. A reasonable heuristic, on the other hand, would be to optimize the general design, by the methods already presented, and then implement cost improvements where warranted.

There are two basic ways to adjust a solution: (1) change the route for an o-d pair, or (2) change all of the routes for a P&D terminal. The former change may only affect transportation and handling cost, whereas the later will also affect sorting cost. In addition to the two basic changes, routes can be changed for groups of stops, possibly through introduction of multiple stop tours. This section will just give a result for the first type of change (without multiple stop routes), which illustrates some of the key principals.

Change in Route for P&D Terminal Pair

If the complete topology is selected, it may be more efficient to serve some P&D terminal pairs via the gateway rather than direct. For such a change to be contemplated, it is likely the case that vehicles sent between the P&Ds are not filled to capacity, whereas vehicles sent to and from the gateway are filled to capacity (otherwise, there would be no transportation cost savings). Then:

$$\text{Savings} = 2(\alpha_2 d_3 + \beta_1 + \beta_2) - 2v[\gamma_2 + \alpha_2(d_1 + d_2)] , \qquad (20)$$

where:
v = volume sent between P&D pair, in each direction
d_3 = distance between P&D pair
d_1 = distance from first P&D to gateway
d_2 = distance from second P&D to gateway.

The first term is the saving from eliminating the direct link, and the second term is the added cost due to increased flow through the gateway. A breakeven point occurs where the savings equal zero. Based on this breakeven, the gateway route is preferred when:

$$v \leq \frac{\alpha_2 d_3 + \beta_1 + \beta_2}{\gamma_2 + \alpha_2(d_1 + d_2)} \quad . \qquad (21)$$

In words, the gateway is preferred when the volume is small and the gateway route is direct, especially if handling costs are small and fixed route costs are large.

If a switch from gateway to direct is contemplated within the star topology, then added sorting cost at the pickup terminal must be factored into the equation. This would be an added disincentive against direct routing.

DISCUSSION

The basic model presented in this paper served to highlight key issues in LAN design (summarized in Table 1). By necessity, this goal demanded simplification. In future research, model assumptions can be relaxed to investigate other scenarios. In particular, the optimal location for a P&D terminal may not fall in the center of its territory, and the optimal location of a gateway terminal may not fall in the center of the region. Optimal terminal location depends on P&D and interterminal transportation costs, as

TABLE 1. PRINCIPALS FOR LAN DESIGN

FACTORS FAVORING STAR OVER COMPLETE TOPOLOGY

Large Proportion of Shipments External
Small Shipment Sizes with Large M
Large Number of P&D Routes
Small Fixed Terminal Cost
Small Handling Cost
Large Sorting Cost

FACTORS FAVORING MULTIPLE STOP INTERTERMINAL ROUTES

Many Terminals
Small Shipment Sizes with Large M
Many P&D Routes

FACTORS FAVORING MANY P&D TERMINALS

Many P&D Routes
High P&D Distance Cost
Low Interterminal Distance Cost
Low Fixed Terminal Cost

well as land costs. If land costs and transportation costs are homogeneous across space, and interterminal costs are small relative to P&D costs, then P&D terminals should be centrally located within their territories, as assumed in this paper.

If interterminal costs are not small, then P&D terminals may be displaced somewhat from the center, toward the gateway terminal or toward other P&D terminals in the region (Campbell, 1990). Also, in urban areas, the vehicle velocity is unlikely to be homogeneous across space; it varies according to location, time and direction of travel. This is especially important in calculating P&D linehaul cost, because P&D vehicles are typically on the road during the busiest commute hours. It may be advantageous to place the P&D terminals at locations that exploit surplus road capacity (Hall and Lin, 1990). If placed near work centers, vehicles will travel in the opposite direction of commuters in the morning, as they head out to deliver, and in the opposite direction of commuters in the evening, as they return with pickups.

The gateway terminal may also be displaced from the center, possibly to save on land costs in the central city and possibly to reduce transportation costs over the wide-area-network. However, congestion delays are not an important factor in selecting gateway locations because interterminal routes are covered in the evening or overnight, when highway traffic is light.

Finally, the gateway can serve the secondary function of P&D terminal. In this case, the gateway should serve a larger territory than other P&Ds, in the same manner as a distribution center for a one-to-many network (Daganzo and Newell, 1986). The incentive for the larger territory is the cost saving in avoiding a transshipment.

All of the above factors will have some effect on topology and number of terminals. These and other issues can be studied through the same format presented here, with costs modified as appropriate.

REFERENCES

Campbell, J.F. (1990). Locating transportation terminals to serve an expanding demand. *Transportation Research*, 24B, 173-192.

Clark, C.E. (1961). The greatest of a finite set of random variables. *Operations Research*, 9, 145-162.

Daganzo, C.F. (1984) The distance traveled to visit N points with maximum of C stops per vehicle: an analytical model and an application. *Transportation Science*, 18, 331-350.

Daganzo, C.F. (1987) The break-bulk role of terminals in many-to-many logistic networks. *Operations Research*, 35, 543-555.

Daganzo, C.F. and Newell, G.F. (1986) Configuration of physical distribution networks. *Networks*, 16, 113-132.

Eilon, S., Watson-Gandy, C.D.T. and Christofides, N. (1971). *Distribution Management*, New York: Hafner Publishing Company.

Hall, R.W. (1984) Travel distance through transportation terminals on a rectangular grid. *Journal of the Operational Research Society*, 35, 1067-1078.

Hall, R.W. (1987) Comparison of strategies for routing shipments through transportation terminals. *Transportation Research*, 21A, 421-429.

Hall, R.W. and Lin, W. H. (1990) LTL trucking in Los Angeles: Congestion relief through terminal siting. University of California Transportation Center, Working Paper No. 43.

Hick, W.E. (1952) On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4, 11-26.

Prager, W. (1959) On the design of communication and transportation networks. *1st International Symposium on the Theory of Traffic Flow*. New York: Elsevier, pp. 97-104.