Concept Design and Analysis of a Linear Intermodal Freight System

PETER J. WONG, ANDREW R. GRANT, AND MASAMI SAKASITA

A conceptual linear corridor intermodal freight system was analyzed by using the computer model LINET. The train operating strategy used—freightliner—is an idealized form of a typical railroad strategy for a corridor. The problem formulation for this generic linear corridor freight system is described, and the LINET computer simulation model and the cost equations used to quantify the various trade-offs and relations between fundamental system design parameters as they affect costs and performance are presented. The results of these analyses include trade-offs associated with the most cost-effective system design, feasible system designs with sufficient capacity, minimum-cost system designs, and design for a specified level of service.

The three major components of a truck and rail intermodal freight system are (a) the local service trucks, which pick up and deliver containerized commodities within a local terminal area by using the highway and street network; (b) the line-haul rail component, whereby trains transport the containers between terminals over a rail network; and (c) the terminal component, which aggregates and transfers containers between the truck and rail components.

Little fundamental research has been done to further the understanding of the interrelations and trade-offs between intermodal freight system engineering design and operating parameters. Thus planners have difficulty making the correct decisions at both a national policy level and a detailed engineering design level to ensure the future economical and effective transport of goods that sustain the nation's economy. This is especially important due to current energy and environmental concerns. An attempt to partly fill the knowledge gap in this area is presented in this paper.

An intermodal freight system is complex; therefore, to simplify this analysis the focus of this paper is on a linear corridor system in which the terminals are in series and concentrated on the line-haul and terminal components. (Because the local service component uses trucks on the highway and road system, its performance was treated as given.) So as not to restrict unduly the range of what is feasible and potentially desirable, the study was conducted on a generic system without the current constraints of existing plants, technological limitations, and institutional restrictions.

Only a portion of the issues and trade-offs that must be understood to design an effective intermodal freight system are addressed. The basis for this paper is research findings originally documented for the Office of Systems Engineering, U.S. Department of Transportation (1,2).

STUDY FRAMEWORK

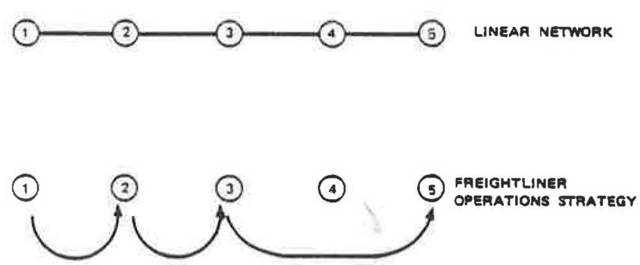
The development of a complete characterization of an intermodal system is a complex undertaking because of the many variables and degrees of interaction. Little research has been done to examine the trade-offs of the intermodal freight system. Therefore, the focus was initially on the simple line-haul system represented by the five-terminal linear network shown in Figure 1. Such a simple linear system has real-world analogs in the numerous heavy-volume freight corridors that exist in the United States. Although the linear network is simple, it provides an abundance of insights that are prerequisites for a systematic examination of a more complex two-dimensional system.

The demand in this simple linear system is characterized by the number of containers (used in a generic sense) going from each origin to each destination. A number of trains move over the line-haul segments and carry containers between terminals; a train moves at a constant speed over the line-haul segments and has a fixed maximum capacity (or size) for carrying containers. All terminals are identical and are characterized by a single processing time that is the combined time needed to load and unload containers from a train; the number of terminal platforms (or berths) determines the number of trains that can be processed simultaneously. (The number of terminal platforms refers to the number of platforms in a terminal for one direction only; it is assumed that the terminals are symmetrical.) Thus the interrelations and trade-offs among the five main engineering system design parameters listed below were studied:

1. Number of trains, N (fleet size);
2. Train speed, V (miles per hour);
3. Train capacity, C (containers);
4. Terminal processing time, P (loading and unloading a train); and
5. Number of terminal platforms, P_L (berths for loading and unloading).

The specification of the simple linear system is incomplete until the train's operating strategy on the network is specified. At one extreme is a local train strategy, in which a train stops at every terminal and picks up containers for all terminals along the route. At the other extreme is the direct-service train strategy, in which a train goes directly between two terminals without intermediate stops and carries only containers for the destination terminal. In this paper the results for an intermediate strategy called freightliner are presented. Freightliner represents an idealized version of a typical railroad operating strategy for a series of major terminals in a corridor. It is essentially a sophisticated local strategy in which stops can be skipped. This means that a train leaves the initial terminal carrying all containers going in the same direction; it stops at an intermediate terminal only if it has containers to deliver or if a specified minimum number of containers are to be picked up. Once stopped, the train will pick up additional containers going in the same direction.

Figure 1. Linear network and train operations strategy.



LINET SIMULATION MODEL AND COST EQUATIONS

A general-purpose simulation system (GPSS) computer model of the linear corridor system was developed for the analysis of trade-offs between the five system design parameters previously listed. This model is called LINET for linear network model. In the LINET model trains go from terminal 1 to 5 or from terminal 5 to 1 (Figure 1) operating in the freightliner mode. LINET is constructed so that the engineering system design parameters can be varied along with the station spacing.

Containers enter the system at an origin terminal with a request to go to a destination terminal. The specification of 24-hr demand is through an origin-destination (O-D) matrix that indicates the container volumes for each O-D pair during the 24-hr period. The actual time when a container enters the system at a specific terminal is randomized and subject to the 24-hr volume constraints imposed by the O-D matrix. The LINET model assumes that the 24-hr demand is repeated daily. The analysis presented here assumes that the system must move 2,300 containers/day.

Once a container enters the system it is not picked up until a train going in the proper direction with sufficient capacity arrives. Trains stop or skip stops according to the freightliner rules. When a train arrives at the end of the system, it is turned around to run in the opposite direction after an appropriate delay.

The LINET model has been instrumented to provide the following system evaluation statistics:

1. Train utilization factor (fraction of train capacity used);
2. Train delay time per link (train delay time to leave a link because of congestion to enter a station); this delay time could be associated with the station or the link; it was convenient for accounting purposes to associate the delay with the link;
3. Time elapsed before a container is picked up at a station;
4. Number of containers delivered during the period under study;
5. Average time containers are in the system;
6. Effective container speed through the system (total distance traveled divided by total time in the system); and
7. Total daily cost of the system.

The last item includes the capital recovery, operating, and maintenance costs. Formulas for total daily cost were developed for various categories in terms of the engineering system parameters N, V, C, P, and P_L. The formulas are given below (the units are dollars per day):

$$\text{Guideway costs} = D(395 + 0.304V^2) \quad (1)$$

$$\text{Terminal costs} = CP_L\,(157 + 133/P) + 1{,}370 \quad (2)$$

$$\text{Crew costs} = 949N \quad (3)$$

where

$$N = (DV + PV^2 - 450)/V(0.33V - 0.83) \quad (4)$$

$$\text{Fuel costs} = \{C(16 + 20U)\,[3.8V + V^2\,(0.0515 + 0.89/D)]\,(\text{train miles})\}/1{,}000{,}000 \quad (5)$$

where

$$\text{Train miles} = 2.18DC_D/CU \quad (6)$$

$$\text{Equipment capital costs} = 0.003NV^2C \quad (7)$$

$$\text{Equipment maintenance costs} = 0.02C(\text{train miles}) + 0.6(\text{fuel cost}) \quad (8)$$

In addition to the engineering system design parameters, other variables were required in the cost formulas. The variable U is the average train utilization factor (i.e., fraction of train capacity used), C_D is the average number of containers delivered during the period of interest, and D is the average distance between terminals.

The development of the cost formulas is documented elsewhere (2). The critical assumptions in the development of the cost formulas are

1. Current rail technology costs are simply extrapolated to obtain costs for advanced technology systems operating at higher train speeds;
2. The intermodal system bears the entire cost of the guideway;
3. The guideway costs increase with the square of the design train speed; and
4. Terminal processing costs increase with the reciprocal of the terminal processing time (i.e., increase inversely with terminal processing time).

As will be discussed later, other forms of guideway costs were used to test the sensitivity of the results.

MOST COST-EFFECTIVE SYSTEM DESIGN

When more money is spent on a system, improved system performance is expected. At what point is it no longer beneficial to spend more money on a system? To answer this question, a cost-effectiveness (or cost-benefit) analysis is often conducted. The result of this analysis would be the system design that provides the greatest incremental performance improvement per additional cost.

Unfortunately, cost-effectiveness is not well defined for a freight system. Consequently, for this analysis the following composite ratio was created, which can be interpreted as a measure of system cost-effectiveness:

Cost-effectiveness ratio = average effective container velocity ÷ total daily system cost.

A more accurate description of the ratio might be effectiveness-cost ratio, because the numerator is a measure of system effectiveness or performance, whereas the denominator is a measure of system costs. However, the term cost-effectiveness is more standard terminology. (The units for the cost-effectiveness ratio are miles per hour per dollar.)

The average effective container velocity is calculated by dividing the total distance of container travel by the total time a container spends in the system (i.e., total time spent in terminals and in line-haul). The result is a measure of the effective speed at which a container moves through the system, which reflects the level of service the system provides the customers.

Extensive parametric trade-off analyses were conducted by using the cost-effectiveness ratio as a figure of merit. The design speed parameter (V) was varied over a wide range to illustrate the trade-off between equipment and crew productivity and the costs associated with higher speeds. Guideway cost was a dominant component of system costs in all cases and thus mediated against high design speeds. Similarly, terminal processing technology, as embodied in a processing time parameter (P), was also varied widely. The cost component of terminal processing time was assumed to increase with the reciprocal of the terminal processing time. The results of the variation indicated a high payoff for reducing terminal

Figure 2. Cost-effectiveness ratio curve.

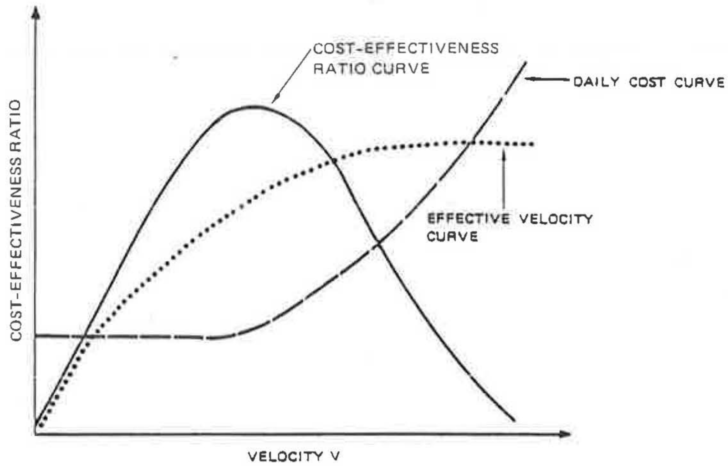
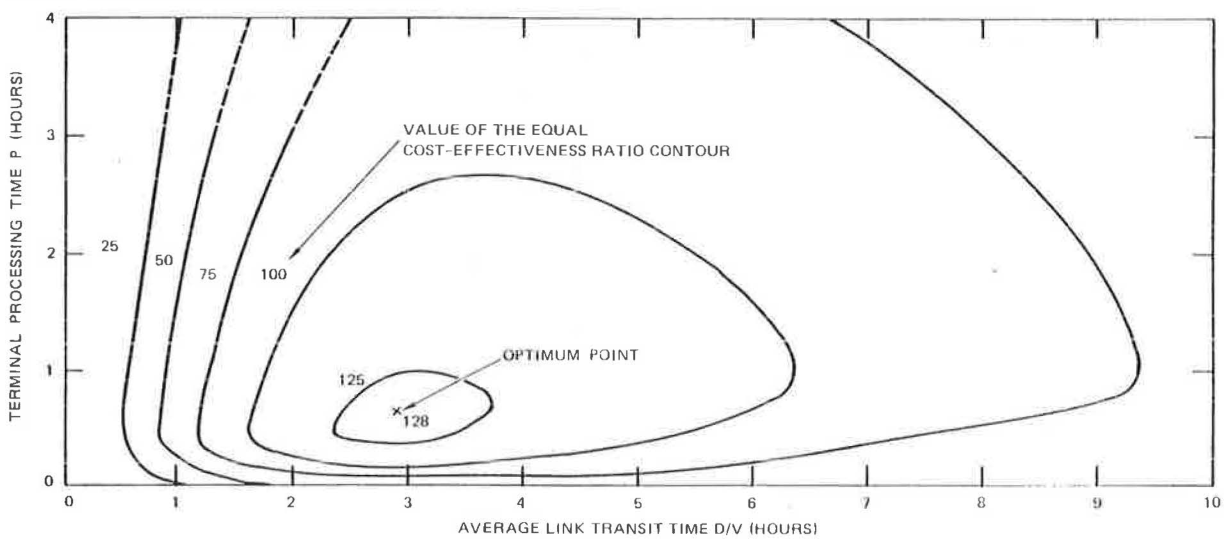


Figure 3. Contour curves of equal cost-effectiveness.



FOR PARAMETER VALUES:

C = 100 Containers

D = 108 Miles/Station

Demand = 2300 Containers/Day

1 Terminal Platform

processing time well below even the best current technologies.

To more fully understand cost-effectiveness relations, the data in Figure 2 show an overlay of three curves that represent cost, effective velocity, and cost-effectiveness ratio. (The data in Figure 2 should actually show three ordinates, but for simplification the ordinates associated with the cost and velocity curves are not shown.) The cost curve increases with V^2. The effective velocity rises rapidly with V and then levels off, which reflects the fact that increasing V adds to the velocity over the line-haul segments but not through the terminal. Thus the curve of the cost-effectiveness ratio has a maximum value. The curve shown in Figure 2 is a function of one parameter, train speed; a family of such curves and an associated set of optimum train speeds exist for various values of the other engineering parameters; e.g., number of trains, train capacity, and terminal processing time.

The data in Figure 3 show contours of equal cost-effectiveness in the two-dimensional parameter space of terminal processing time (P) and the average line-haul transit time between terminals (D/V), where D is the average distance between terminals. From this figure it appears that the contours are closed and that a distinct optimum, marked by an X, exists.

The data in Table 1 indicate that if train capacity is the independent variable and the system capacity remains the same, the optimum system configurations from a cost-effectiveness ratio are nearly identical for the 10- and 25-container capacity trains, which indicates that the optimum in terms of train size is relatively flat. Thus the optimum system configurations have a large number (30 to 60) of relatively small trains (10- to 25-container capacity) that travel at moderate speeds (45 to 60 mph). The associated optimum terminal processing time is in the range of 6 to 12 min to both load and unload a train.

The assumed cost relations play a critical role in determining the most cost-effective system design. For the cost relations and cost-effectiveness ratio criteria assumed here, however, the conclusion is that the most cost-effective engineering design for a freight system calls for a large number of small trains moving at moderate speeds; the associated terminal processing time is on the order of fractions of 1 hr.

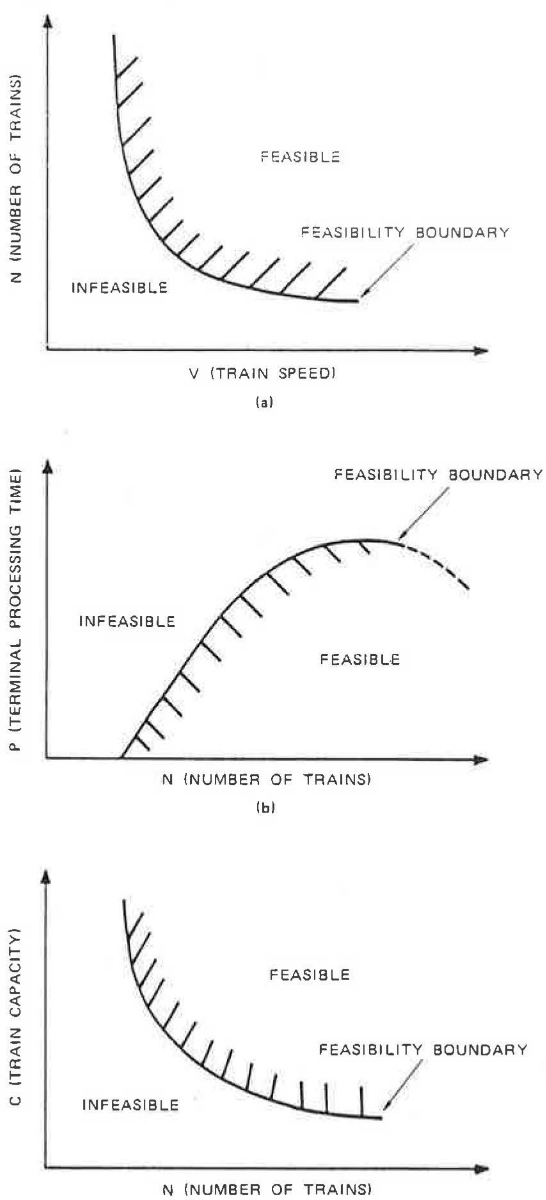
This conclusion remains valid even when the following three modifications to the assumed cost formulation are made:

1. Total guideway cost (fixed plus velocity-dependent terms) is reduced by a factor of one-tenth,
2. Only the velocity-dependent term of the guideway cost is reduced by a factor of one-tenth, and
3. A new guideway cost is formulated to be equal to a $2.00 surcharge on every dollar spent for fuel.

Table 1. Optimum sets of engineering system parameters.

	Optimum Engineering System Parameters by Train Capacity (no. of containers)				
Item	10	25	50	100	150
Cost-effectiveness ratio	175	171	153	128	119
No. of trains	59	31	19	11	7
Line-haul speed (mph)	60	47	40	37	36
Terminal processing time (hr)	0.1	0.2	0.36	0.6	0.85
Effective container speed (mph)	48	34	25	21	22
Daily cost ($000 000s)	0.27	0.20	0.17	0.16	0.18

Figure 4. Feasibility boundaries in two dimensions.



The rationalization or interpretation of these results is as follows. A specified level of system capacity can be achieved through use of a small number of high-speed trains or a larger number of smaller moderate-speed trains. The effective container velocity can be increased by accelerating (a) the speed over the line-haul segment by maintaining higher train speeds or (b) the speed through the terminals by maintaining faster terminal processing time and more frequent train departures (i.e., more trains) to reduce the container wait time for a train connection. Because the cost functions assumed in this study increase rapidly with the square of train speed (Figure 2), and because adding more trains and decreasing terminal processing time is less expensive than increasing train speeds, the most cost-effective strategy is to (a) build terminals that can process trains faster and (b) have more smaller trains moving at moderate speeds rather than fewer high-speed trains. Also, as the number of trains is increased, the arrival frequency at the terminal increases; therefore, the terminal processing for a train must be rapid in order to avoid queuing delays for service at the terminal.

FEASIBLE SYSTEM DESIGNS WITH SUFFICIENT CAPACITY

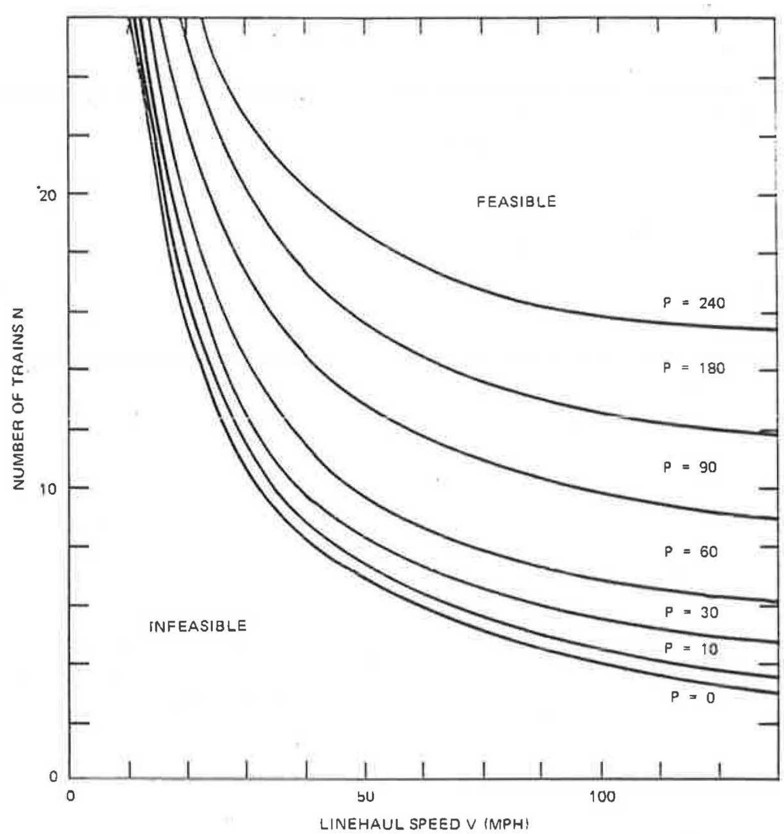
Insight into the fundamental interrelations and trade-offs among system design parameters can be obtained by studying the feasible design alternatives that have sufficient container-carrying capacity to satisfy the steady-state, 24-hr demand for container shipments. A system is defined as having sufficient capacity if essentially all of the containers are shipped within 24 hr.

The multidimensional system parameter space can be divided into two regions. In one region the system is capable of satisfying the demand, and in the other it is not. This analysis focused on the representation of this feasibility region in two dimensions. Examples of these feasibility regions in several two-dimensional parameter spaces are shown in Figure 4. The curve that separates the feasible from the infeasible region is called the feasibility boundary.

In the V versus N parameter space the feasibility boundary is hyperbolic in shape (Figure 4a). The vertical asymptote indicates that a minimum train speed is required to satisfy delivery of the containers. The horizontal asymptote indicates that a minimum number of trains is required.

In the P versus N parameter space the feasibility boundary rises with a slope to the right before leveling off. The initial rise of the curve to the right is explained by the fact that, with few trains initially, the terminal processing time must be fast to satisfy the demand. As more trains are added to the system, however, terminal processing does not have to be as fast to satisfy the delivery of the containers up to the point where the curve begins to bend to the right and level off. The curve bends because, as additional trains are added to the sys-

Figure 5. Feasibility curve family in N-V plane.



NOTE: Feasibility boundaries for linehaul speed versus number of trains for various processing times (P) in minutes

D = 108 Miles
C = 100 Containers
1 Platform/Station
Freightliner Mode
Demand = 2300/Day

Figure 6. Feasibility regions in C-N plane.

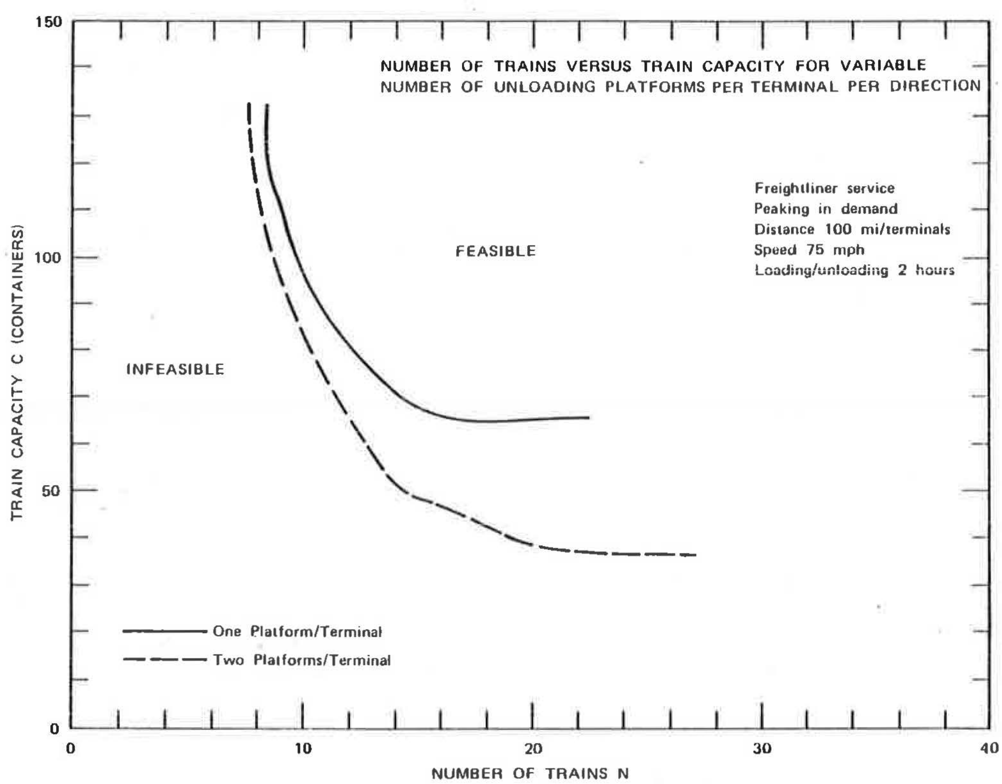
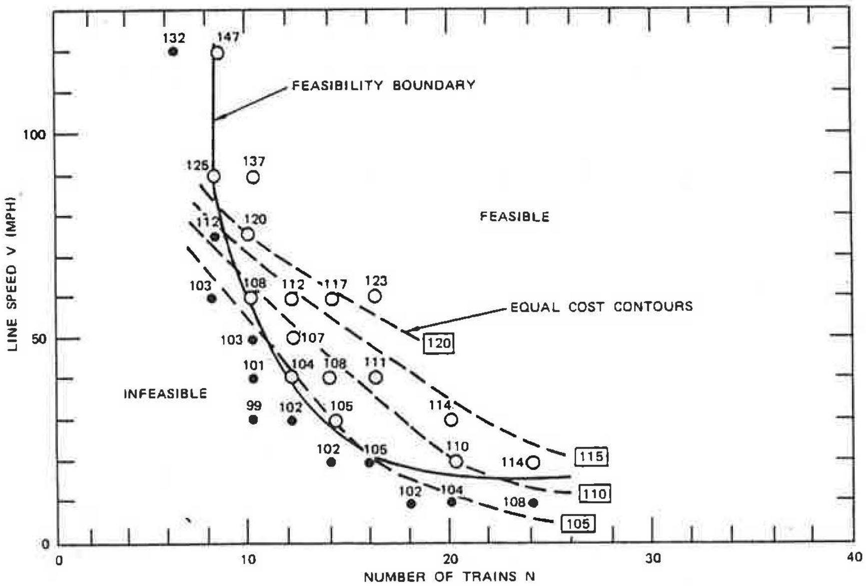


Figure 7. Feasibility curve and equal-cost contours in V-N plane.



NOTE: Numbers Reflect Total Annual Cost (Million Dollars)

tem, the terminal processing time must be sufficiently rapid to prevent queuing delays for trains waiting in the terminal to be processed. (In fact, a portion of this analysis indicates that the curve at some point begins to bend down.)

In the C versus N parameter space the feasibility boundary is again hyperbolic. The vertical asymptote indicates that a minimum number of trains are required to satisfy the demand; the horizontal asymptote indicates there is minimum train capacity.

In reality the feasibility boundary is a surface in a multidimensional parameter space. Slices of this surface in two dimensions are shown in Figure 5; the other associated parameter values are not displayed. The data in Figures 5 and 6 provide examples of how the feasibility boundary changes in the N versus V space as either the terminal processing times decrease or the number of terminal platforms increase; in both cases the feasibility regions increase.

In the N versus V parameter space the data in Figure 5 indicate the enlargement of the feasibility region as the terminal processing times decrease. As they decrease, the feasibility boundaries become a nested set of feasibility curves; therefore, feasible system designs become possible with smaller numbers of higher-speed trains as terminal processing times decrease.

In the C versus N parameter space the data in Figure 6 indicate that the feasibility boundary that assumes one terminal platform is nested inside the feasibility boundary that assumes two terminal platforms. The two-platform system can operate with a larger number of smaller trains than that possible with a one-platform system.

MINIMUM-COST SYSTEM DESIGNS

A typical example of a feasibility boundary and the associated equal-cost contours are shown in Figure 7. The cost curves are somewhat similar in shape and orientation to the feasibility curves but have less curvature. The system costs increase as the distance from the origin increases. Minimum system costs are found in the knee of the feasibility boundary at the point where the feasibility boundary is tangent to a cost curve. The costs in the knee of the feasibility boundary are fairly constant throughout the knee and near the minimum cost. Thus the knee of the feasibility curve is an area where minimum-cost designs, or near-minimum-cost designs, are achieved.

The system designs associated with points in the knee represent a considerable range of design alternatives. In the example the range of approximately equal-cost designs in the knee extends from 11 trains at 50 mph to 16 trains at 20 mph, with perhaps the least-expensive feasible solution using 13 trains at 30 mph.

The minimum-cost design is not necessarily the most cost-effective design. It is merely the least-expensive feasible solution. Points in the interior of the feasible region may provide higher cost-effectiveness even though they cost more.

DESIGN FOR SPECIFIED LEVEL OF SERVICE

One of the most important measures of the service effectiveness of a freight system is the average time a container spends in the system. A family of curves is shown in Figure 8 that represent time in the system plotted against line-haul speed for a specific combination of train capacity, demand, and interstation distance for systems along the feasibility boundary. The hyperbolic shape is clearly evident, and the general shape is typical. Two useful inferences can be made from this illustration.

1. At low speeds the time in the system rises rapidly as speed decreases, and reductions in processing time are not effective in reducing the time in the system.
2. At speeds greater than 50 mph the reverse is generally true; i.e., increased speed does not greatly reduce the time in the system. Increased processing time either increases time in the system or requires considerable increases in line-haul speed if time in the system is to be constant. At those speeds the travel time is small compared with other time components (loading and unloading time, lost time, and waiting time), and the travel time component becomes smaller as speed increases.

Figure 8. Time in system versus line-haul speed.

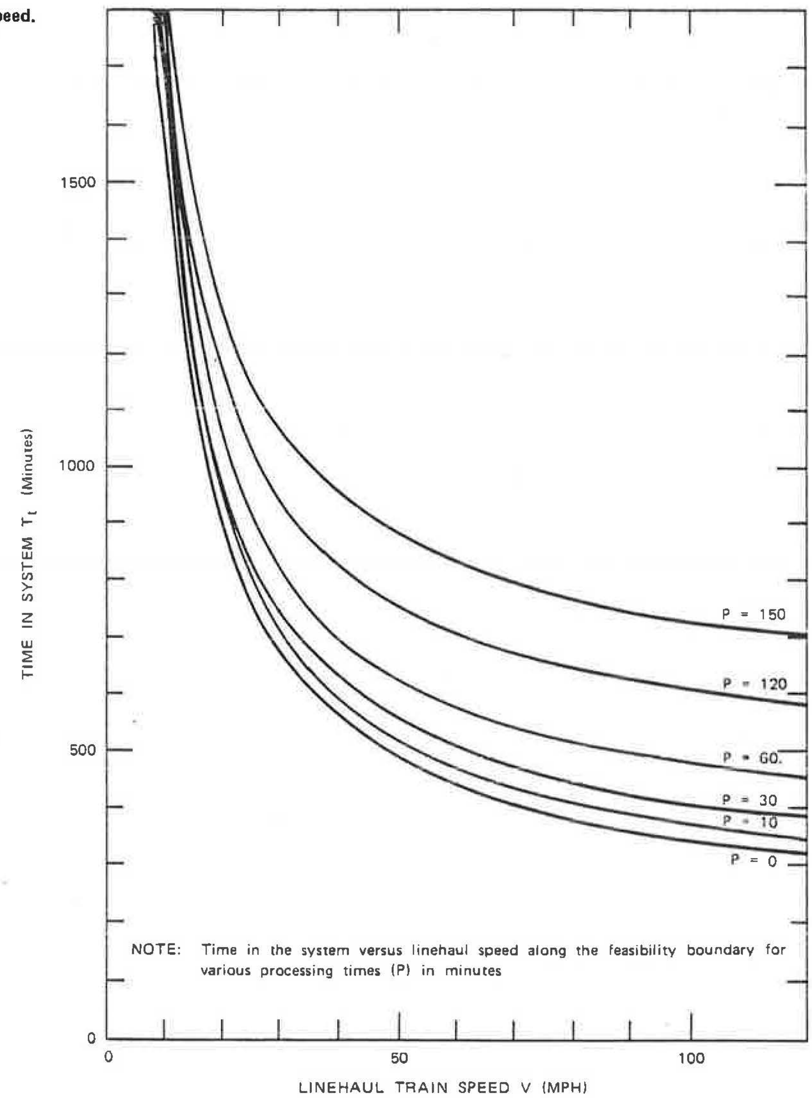
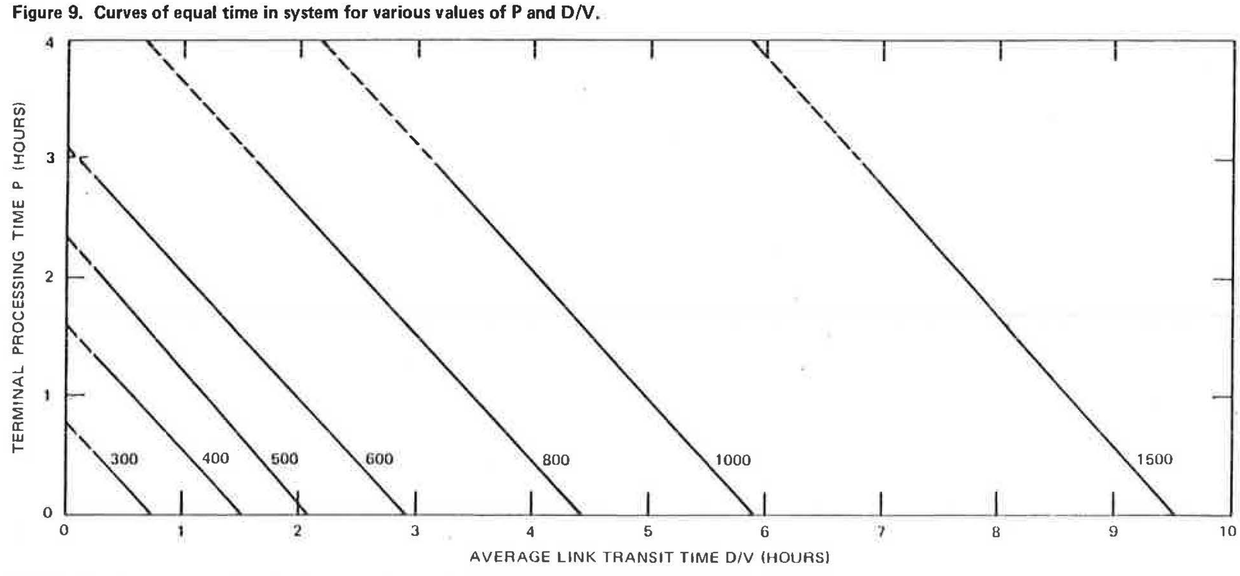


Figure 9. Curves of equal time in system for various values of P and D/V.



NOTE: Numbers represent time in the system (minutes)

C = 100 Containers
D = 108 Miles/Station
Demand = 2300/Day
1 Platform

A plot similar to Figure 8 would be useful in the initial selection of parameters for a system designed to provide a certain level of service. For instance, for a average container time in the system of 600 min, a processing time of 60 min would require a line speed of 55 mph. Reducing processing time to 30 min would reduce the required line-haul speed to 45 mph. A zero processing time would still require a line-haul speed of 38 mph. On the other hand, increasing the processing time to 120 min would require a line-haul speed in excess of 100 mph.

It is informative to plot curves of equal time in the system for various values of terminal processing time (P) and transit time across a line-haul segment D/V. (The use of the variable D/V instead of V versus P is useful because D/V and P are in the same units, i.e., time.) The data in Figure 9 show such curves for a specific combination of other system design variables. The lines in this figure are fairly straight and evenly spaced. This should not be surprising, as P and D/V are combined linearly in calculating time in the system and heavily influence the result.

The data in Figure 9 provide a means of rapidly determining the trade-off between D/V and P for any given level of service. The sections of the curves above P = 3 are unsubstantiated by LINET runs and are therefore indicated with dashed lines. It would be expected that, as the line-haul transit time (D/V) decreases, a breakdown point would occur at which the linear relation would no longer be valid. The curve should begin to bend down with decreasing D/V, which indicates that the terminal processing time (P) must decrease to avoid train queuing delays in the terminal.

The number of terminal platforms influences the size of the feasibility region (i.e., a system with 2 platforms/terminal has a larger feasibility region than a system with 1 platform/terminal). Once a system design is feasible, however, adding extra platforms to terminals has little effect on the average time a container spends in the system. Thus the number of platforms affects the ability of the system to satisfy the demand, but once the system is able to satisfy the demand, the number of platforms has little effect on system effectiveness.

ACKNOWLEDGMENT

This research was sponsored by the Office of Systems Engineering, U.S. Department of Transportation. S.C. Chu was the project technical monitor. Appreciation is expressed to A.E. Moon, R.M. Corbett, and M.A. Hackworth for their technical contributions.

REFERENCES

1. P.J. Wong, A.R. Grant, M. Sakasita, and A.E. Moon. Concept Design and Analysis of Intermodal Freight Systems--Volume I: Executive Summary. Office of Systems Engineering, U.S. Department of Transportation, Rept. DOT-RSPA-DPB-20-80-2.1, Jan. 1980.
2. P.J. Wong, A.R. Grant, M. Sakasita, and A.E. Moon. Concept Design and Analysis of Intermodal Freight Systems--Volume II: Methodology and Results. Office of Systems Engineering, U.S. Department of Transportation, Rept. DOT-RSPA-DPB-20-80-2.2, Jan. 1980.

Publication of this paper sponsored by Committee on Intermodal Freight Transport.

Network Analysis of Highway and Intermodal Rail-Highway Freight Traffic

ALAIN L. KORNHAUSER AND MICHAEL BODDEN

The analysis capabilities of the Princeton highway and intermodal rail-highway network models are described. These network models are extensions of the Princeton railroad network model and graphic information system and are based on a geocoded network representation of intermodal transfer locations and the U.S. highway system. The models contain efficient routing and traffic assignment algorithms, highway and rail cost models, and extensive network editing and computer graphic utilities. Examples of highway and intermodal routes and a graphic analysis of the railside flows of 1980 intermodal traffic based on the 1980 one percent waybill sample are presented.

Analysis of U.S. highway and intermodal (highway-rail) traffic has been difficult because precise and broad-based highway traffic data were lacking and because an efficient computer-based network representation of the U.S. highway system did not exist. The unavailability of these data is unexpected given the amount of planning and funding that has been expended on the U.S. highway system. One would have assumed that the FHWA would have sponsored the creation of such a network data system, or that the Interstate Commerce Commission (ICC) or the FHWA would have secured the authority to collect a sample of highway traffic movements similar to the 1 percent waybill sample collected for rail freight (1).

However, because the carrier portion of highway freight transportation is fragmented and some sections of highway transportation are not regulated, no national sample of origin and destination data for highway freight traffic exists. The best publicly available cross-sectional national sample of truck traffic is the 1977 Census of Transportation (2). Although beneficial, this data source is significantly inferior when compared to the rail freight waybill sample. The origin and destination data of the 1977 Census of Transportation are grossly aggregated to state levels or to metropolitan areas, and no revenue data are given. Similarly, there are little or no data available for intermodal traffic because no government agency collects it. (Because intermodal traffic is deregulated, there may not exist a public need to know.)

The rail freight waybill sample only reports rail interchange locations, and not the ultimate highway