Incorporation of Operational Decision Making in Intermodal Terminal Simulation Models

DOUGLAS P. SMITH

The model structure outlined in this paper provides a framework for the analysis and improvement of certain terminal operating procedures. The foundation of the model is a procedure for forecasting and updating the volumes of trailers to be handled. The short-term uncertainty relating to outbound trailer volumes can be one of the major causes of terminal inefficiency, particularly with respect to hitch use. This uncertainty is incorporated into the model structure and is used in the assignment of railcars to hitches. A combination of automatic and interactive methods are used by the simulator to allocate terminal resources. These resources include loading equipment, tracks, railcars, and switching facilities. This allocation process simulates the management component of the terminal. The physical component is represented by a series of queues, buffers, and processors, each with specified capabilities and availability. Unloading activities, the gate, and storage are not included. Results that indicate the accuracy and potential applicability of the model are not yet available. Testing is being done by using Canadian National Railways' Brampton Terminal, which is located on the northwest corner of Toronto.

Simulation models have been touted as a tool that can aid in the development and planning of intermodal terminals and systems. The complexity of most intermodal operations makes it difficult to evaluate alternatives by using simple analytic methods, but it is reasonable to assume that a well-developed set of simulation models will allow the intermodal operator to test a variety of system configurations quickly and at low cost. Simulations are appropriate because of the time-varying nature of terminal activities and the intensity of peaks. The characteristics of a model currently being developed to perform detailed analyses of the loading operations in intermodal terminals are discussed in this paper.

TERMINAL OPERATIONS

In order to identify characteristics of typical terminal operations, many intermodal terminals were visited during summer 1982. These covered a wide range of sizes, layouts, and operation policies. Attributes of some of these terminals are given in Table 1. Although their physical characteristics may vary widely, analysis of the operations at these terminals reveals a number of consistencies in both the work-load pattern and the methods used to handle the work load.

Most terminals have two distinctive types of peak, one recurring on a daily basis and the other recurring weekly. The daily peaks follow from a terminal being in the load mode in the evening and the unload mode in the morning. Most trailers are loaded by customers during the day and delivered to the railway in the late afternoon and early evening; outbound train schedules reflect this pattern. Similarly, customers want to have their trailers available to unload during the day; thus, the early morning period is characterized by train arrivals and unloadings.

Although these patterns are generally true, other

Table 1. Characteristics of representative intermodal terminals.

Terminal	Railroad	Apron Tracks	Car Spots[a]	Loading Method	Parking Spaces	General Comments[b]
South Kearney	Consolidated Rail Corporation (Conrail)	5	153	5 side-lifts, 1 crane	1,800	TOFC and COFC[c]; high volume
47th Street	Conrail	3	91	1 side-lift, 1 crane	700	TOFC; high volume
West Springfield	Conrail	2	52	2 side-lifts	310	TOFC; medium volume
Beacon Park	Conrail	4	82	3 side-lifts	550	TOFC and COFC[c]; medium volume
Detroit	Norfolk and Western (N&W)	10	50	Circus	200	TOFC; low volume
		1	5	1 side-lift	200 TEU	COFC; low volume
Calumet	N&W	2	79	3 side-lifts	1,200	TOFC and COFC; medium volume
Luther	N&W	3	82	2 side-lifts	850	TOFC and COFC; medium volume
Ogden	Burlington Northern	2	52	3 cranes	600	TOFC; high volume
		1	25	2 side-lifts		COFC
Chicago	Missouri Pacific (MP) and Louisville and Nashville	10	166	3 cranes	1,000	TOFC and COFC[c]; high volume
St. Louis	MP	9	51	Circus	400	TOFC; low volume
		1	5	Rail-mounted crane		COFC; low volume
Chicago	Illinois Central Gulf	4	140	3 cranes	1,000	TOFC and COFC[c]; high volume
Corwith	Santa Fe	5	200	6 cranes	4,200	TOFC; very high volume
		1	12	1 side-lift		COFC
Detroit	Grand Trunk Western	2	48	2 cranes	500	TOFC; low volume
Detroit	Detroit, Toledo, and Ironton	7	38	Circus	500	TOFC; low volume
Chicago	Soo Line	3	35	2 side-lifts	120 200 TEU	TOFC and COFC; low volume
Alexandria	Southern	2	38	2 cranes (rail mounted)	300	TOFC; medium volume
Montreal	Canadian Pacific (CP)	4	57	7 side-lifts	3,000 TEU	COFC; medium volume
Toronto	Canadian National (CN Rail)	3	90	2 cranes	2,000	TOFC; medium volume
Montreal	CN Rail	10	40	Circus	260	TOFC; medium volume
		2	40	1 crane, 9 side-lifts	3,600 TEU	COFC; medium volume

Note: TEU = twenty-foot equivalent units.

[a]Figures for car spots are based on 89-ft cars.
[b]TOFC = trailer-on-flatcar and COFC = container-on-flatcar. Volumes are divided as follows: low = 0-200/day (load and unload), medium = 200-500, high = 500-1,000, and very high = 1,000 or more. These volumes are based on typical heavy days and are based, for the most part, on estimates rather than actual operating records.
[c]Indicates no ground storage for containers.

factors (such as multiple daily departures) will result in some variation about the daily norm. The weekly pattern is characterized by high levels of storage early in the week and high loadings toward the end of the week. This reflects a shipper tendency to move higher volumes toward the end of the week, which results in higher loading volumes on Thursdays and Fridays. Trailers that arrive at a terminal over a weekend are not likely to be picked up until the following Monday, which results in higher storage requirements during the early part of the week. These regular cycles simplify the analysis of individual terminals because they allow one to focus on particular periods during the day or week.

The reported decision-making process at the terminal level was also consistent over the terminals visited. When asked how they made operational decisions, most of the operators interviewed indicated that they would "play it by ear." They elaborated on this by saying they had a general idea what the demand pattern for a day (or week) would be, but that there was too much uncertainty to make a fixed set of resource allocations at a very early stage. Over-the-road arrivals of intermodal trailers are not controlled by the terminal nor is complete information on future trailer arrivals available, so all decisions are based on estimates of the daily volumes. An initial set of decisions is made and updated as the day progresses. Consistent with the scheduled timing of different services, these updates will be used for decisions with an ever-changing set of possible alternatives. An example of this change could be the feasibility of switching at different points in a loading schedule.

These characteristics of intermodal terminal decisions demonstrate the critical importance of human factors and local management in terminal operations. The uncertain environment of short-term decision-making activities requires carefully designed decision support systems together with appropriate operations policies, particularly with respect to the loading component of terminal operations. It is the loading component of the terminal that is affected strongly by complexity, and for this reason the current analysis focuses on loading. This emphasis is justified by the relative importance of the loading function to both terminal and overall system performance.

Examination of a terminal in the context of the overall rail network indicates that the level of effectiveness of that terminal in delivering service depends on the ability of the loading component to block trains appropriately and to assemble them quickly. The unloading component is important with respect to making trailers available to customers, but it is a relatively simple procedure with no blocking or hitch use issues as well as marginally faster cycle times, and therefore it is less likely to affect system performance. Other areas such as the gate, hostling, and storage are important for the support they give to loading. The lack of sufficient support could easily be the limiting factor for a specific terminal.

LOADING DECISIONS

The loading process for outbound trailers involves four general groups of decisions: (a) the assignment of apron tracks to specific trains, (b) the assignment of railcars to blocks for loading, (c) the determination of switching requirements, and (d) the determination of loading and unloading sequences. These are outlined below.

Track assignment refers to the selection of a specific track or tracks for assignment to each train. Track assignment is based on train length, expected track availability, and, in mechanized terminals, crane movement restrictions. Local conditions, such as the proximity to storage areas, the location of internal road crossings, and the physical characteristics of the apron, may also favor specific track assignments.

Cars are assigned to outbound blocks through the selection of appropriate strings of empty cars and their allocation to specific destinations. Anticipated volumes are the major decision factor. The assignment is normally done iteratively and can be reassessed as trailers arrive throughout the loading period. Blocks are located relative to one another in a way that facilitates train makeup.

Switching is required on outbound trains when there is a need to add railcars to the original allocation or to switch those for which there is nothing to load. Additional switching will be required if it is not possible to properly block the train during loading.

Loading sequence refers to the assignment of trailers to specific hitches on railcars. Sequence assignments are normally made so as to optimize some measure of hitch use; these assignments are made either in the gate office during check-in or by the crew during loading.

These decision groups are either preset as standard practice or are made by terminal staff on an informal basis. In many terminals, track or block allocations will not vary from day to day. This standard allocation stems from an earlier decision on operation procedures and may be adjusted, given a significant change in circumstances. Switching will commonly be done on a scheduled basis, but extra switches may be requested as required. The scheduled switch is part of current practice, but the decision for an extra switch is normally based on an informal assessment of the current situation. Hitch-assignment decisions are made continuously as trailers arrive at the terminal; this decision making is done on an informal basis.

MODEL STRUCTURE

The translation of informal decision rules into a form that can be used by a computer simulation can be done in two ways. The first is to allow interaction between the computer and the operator. This type of simulation, known as "man in the loop," has been recommended for the simulation of intermodal terminals (1) and is used by CN Rail to test and develop designs for classification yards (2). The computer is used primarily for bookeeping purposes, and all decisions are made by the operator in the same manner that they would be made in the terminal. This method effectively removes the requirement of specifying decision rules in machine formats, but it is disadvantageous in terms of simulation time and cost. The second method is to develop a structured set of rules that closely approximate the observed decision-making process and can be coded into a computer algorithm. These rules will usually involve selecting the decision that is optimal according to some predetermined criterion.

Decision making in the computer model is achieved through a combination of interactive and automatic methods. Those decisions that are repetitive are made automatically consistent with a predetermined set of rules, and those that are seldom repeated are handled by an experienced operator. The decisions, their criteria, and the methods used for each are described below.

Track assignment is handled via the interactive interface. These decisions are made at the start-up of a simulation and at pauses in the simulation (also known as interrupts), which are generated whenever the status of a track changes. Examples of status changes include completion of unloading or loading. The operator assigns tracks on the basis of anticipated volumes, the relation between apron tracks and storage areas, and the physical characteristics of individual tracks.

Block assignment is handled via the interactive interface, with some updates made automatically and others interactively. A preliminary block assignment is made during track assignment, and the relative location of blocks is a function of train makeup considerations and expected volumes. Initially, only the starting point and the loading direction (or directions) for a block are set. The finishing point remains flexible, which reflects the uncertainty in volumes. Automatic decisions would include the release of sets of cars for loading in a block. Essentially, this means that the adjustment of block assignments will reflect changes in the expected volume due to variations in the trailer arrival pattern. If volumes are higher than expected, it may be necessary to reallocate railcars and possibly require a change to either the block or track assignment. This decision is made interactively, and the required simulation interrupt is generated when there is a conflict in the automatic updating of block assignments.

Switching decisions are handled interactively, while the switch itself is handled automatically. The factors leading to a switch include a requirement to spot cars for loading and for postloading train makeup. Any of the simulation interrupts used to make decisions on track or block allocation can be used for switching and, in addition, any indication of future railcar shortages will generate an interrupt. This reflects the advantage inherent in being able to schedule switches early rather than waiting until the last minute when it may be physically impossible to load the newly placed cars before cutoff.

Hitch allocation is done automatically in concordance with a specified set of rules. Three basic rules are used. The first is "first suitable hitch," in which a trailer will be assigned to the first space within which it will fit. A 40-ft trailer, for example, could be placed in either a 40- or a 45-ft position, but not a 27-ft position.

A second rule is "best hitch," in which a trailer is placed on the best hitch available. By using this rule, the 40-ft trailer would be placed on the first 40-ft position or, if none were available, on the first 45-ft position.

The final rule is "minimize excess train length." Hitch positions would be assigned according to expectations of volume and trailer mix. If a relatively high proportion of 40-ft trailers were expected, the optimum allocation could have some of them placed on 45-ft hitches. The determination of the expected increase in train length is based on the probability associated with specific trailer arrival events.

A simple example of the approach concerns a situation where nine railcars each have one 40-ft trailer loaded on the 40-ft hitch at 1 hr to cutoff. A tenth 40-ft trailer arrives and the current decision concerns whether to put it on the first 45-ft spot or load it on the tenth railcar in the 40-ft position. Based on past arrival patterns, the probability of more 45-ft trailers arriving in the last hour is as given in Table 2. (For simplicity, we assume that there will be no more 40-ft trailers.) The table shows that it is known with certainty that at least five 45-ft trailers are coming, so a minimum of five spots can be saved with no risk of penalty. The expected penalty associated with only five arrivals is calculated by determining the empty spaces that would remain and multiplying their total length by the probability of the event "five more trailers." A similar calculation is performed for each of the other arrival events with nonzero probability, and the sum of these products gives the expected excess train length associated with each decision. In this example, loading the 40-ft trailer in the 45-ft spot offers a clear advantage, even though an automatic 5-ft penalty is incurred.

These calculations are not suitable for a terminal clerk to perform each time a trailer arrives, but they are a reasonable representation of the type of intuitive reasoning an individual would make. Essentially, the individual would recognize that the likelihood of a large number of trailer arrivals is not high enough to warrant reserving many more positions. In the actual terminal simulations, the arrival events are much more complicated than the example, and frequently include compound events such as six 45-ft trailers, three 40-ft trailers, and three 27-ft trailers. Rather than summing over six possibilities, as is the case in the example, it may be necessary to consider hundreds of possibilities as is done in the simulation model.

The major requirements for the implementation of the minimum excess train length hitch-assignment rule are the probabilities associated with the various arrival events. These are determined by the analysis of historical information on trailer arrivals. Trailers will be grouped by destination, trailer length and weight, departure time, day of week, and plan (1, 2, 3, and so on); and the consistency and predictability of their arrival patterns will be determined. Ideally, these patterns would differ only in terms of timing and magnitude. This would greatly simplify the forecasting and data-collection tasks for a specific simulation. It is expected, however, that specific variations in pattern will be associated with different departure times and with plan 2 traffic.

Departure time or cutoff time may affect arrival patterns because of its relation to the times wher shippers make trailers available. Most trailers will become available from mid-afternoon through the evening after being loaded by the shipper during the day. Clearly, shippers with multishift operations can delay or advance trailer releases with greater

Table 2. Sample hitch-assignment calculations.

Expected 45-ft Units	Probability	Excess Space	
		Load Car 10	Load 45-ft Spot
0-4	0		
5	0.2	5 * 45 = 225 ft * 0.2 = 45	3 * 45 = 135 ft * 0.2 = 27
6	0.3	4 * 45 = 180 ft * 0.3 = 54	2 * 45 = 90 ft * 0.3 = 27
7	0.2	3 * 45 = 135 ft * 0.2 = 27	45 ft * 0.2 = 9
8	0.1	2 * 45 = 90 ft * 0.1 = 9	0
9	0.1	45 ft * 0.1 = 9	40 ft * 0.1 = 4
10	0.1	0	2 * 40 = 80 ft * 0.1 = 8

Note: The total expected excess space for "Load Car 10" and "Load 45-ft Spot" is 144 and 75, respectively, The latter column incurs a 5-ft penalty, which brings the minimum up to 80.

freedom than can single-shift operations, but they form only part of the market. It is unlikely, therefore, that the arrival pattern for a 2200 cutoff can be represented by a 3-hr shift in the pattern for a 1900 cutoff. Similarly, it is not likely that plan 2 traffic, which is railway controlled, will have a pattern that matches non-rail-controlled traffic. The impact of finite pickup and delivery resources would be the primary reason for this variation.

Given that a consistent set of arrival patterns has been identified, it is possible to develop estimates that describe the probabilities of certain trailer arrival events and use these probabilities to identify optimal hitch assignments. Depending on the nature of the patterns, it may be possible to update the estimates as the day proceeds. The update of volumes for a specific train may depend on the volume of earlier arrivals for that train or possibly arrivals for a benchmark train. If, for example, a 1500 departure was expected to have 60 trailers and 80 turned up, it may be reasonable to increase the estimates for later trains by some factor. Similarly, plan 2 traffic may be a useful indicator of overall volumes because its magnitudes are known earlier in the day. The accuracy of the updates will depend on the consistency of arrival patterns and the relations between them. If these patterns turn out to be essentially random, then updating will not improve performance, but it should be recognized that the use of probabilities in hitch allocation would still result in long-run optimality.

There is a range of methods that can use additional information about a process to update the estimate of the final result. These range from simple look-up tables to complex techniques used in feedback control systems. Bayesian updating is used in this model; it essentially takes an initial estimate of trailer arrival probabilities, adds the information, and then produces an adjusted estimate of these probabilities. This adjustment is intended to approximate the intuitive updating done by terminal staff as the day progresses.

The decision-making algorithms described above will simulate the management portion of the terminal. The physical component will be handled in a manner similar to standard simulation models, which represent terminals by a series of queues, buffers, and processors (e.g., lines of trucks, parking lots, and cranes) with specified capabilities. Expected throughput is determined by calculating the expected availability of terminal resources that have known processing rates. Availability is a function of delays, which includes, for example, switching interference during respots or train makeup, nonproductive crane travel for the purpose of track changes, or waiting time during a changeover from loading to unloading. Where the forecast demand exceeds the short-term capability for processing, an interrupt will be generated so that more resources can be allocated if this is feasible.

The construction of a highly detailed terminal model is a large endeavor. To reduce the overall effort required, this model will focus on the train-loading and makeup activities, whereas the gate and parking will be considered as external factors. The trailer arrival pattern at the apron will be assumed to be the same as that at the gate. This assumption is reasonable in many situations, but it should be examined carefully, particularly where hostling requirements are severe or gate delays are highly variable. Similarly, the impact of off-loading requirements on loading can be ignored where a terminal follows a morning unload and afternoon and evening load cycle. If this is not the case, the analysis must become more complex.

The programming languages used for the model are BASIC and Assembler and all programming has been done by using an IBM personal computer with 512 K of memory. Actual memory requirements for the finished version will be less.

The model is being developed by using the CN Rail TOFC terminal in Brampton, which is located at the northwest edge of Toronto. This facility serves the southern Ontario market. The Brampton Intermodal Terminal (BIT) has a number of advantages that favor the development of this type of model. The terminal follows a simple morning-load and afternoon-unload pattern; there are no space restrictions in the facility; and each of the three apron tracks is of similar size and accessibility. Hostling requirements are minimal, and gate delays are not a factor in the loading operation. BIT uses two overhead cranes. Three of the four daily trains are blocked by destination and respots may be required during the loading period. The need to load 8 to 10 major blocks over three tracks creates conflicting demands on the cranes and requires appropriate blocking patterns.

CN Rail operates a much wider variety of railcar types and carries a more complex mix of trailer sizes than other North American railroads. Hitch use is extremely important to terminal operations, and the assignment rules will have many more alternatives than would be the case with 89-ft cars and either 40- or 45-ft trailers. Experience gained in this analysis may provide valuable insights into what may happen in the United States, given a mixture of 89-ft and multiplatform articulated cars together with 40-, 45-, 48-, and possibly 27-ft trailers in the system.

POTENTIAL APPLICATIONS

Three possible applications for a model of the type described are presented here: (a) the development of a simple decision support system for clerical staff in loading operations, (b) the identification of optimal blocking and switching strategies, and (c) the evaluation of alternative railcar fleets.

There is a simple relation between the minimize excess train length hitch-allocation method used in the model and the simple best-hitch or first-hitch rules. Depending on the expected trailer arrival patterns, the minimization of train length could cause switching between the two simpler rules. A diagram similar to Figure 1 could help the clerk in this switching process. The vertical axis represents the time remaining until cutoff and the hori-

Figure 1. Points of equivalent effectiveness of best-hitch and first-hitch rules.

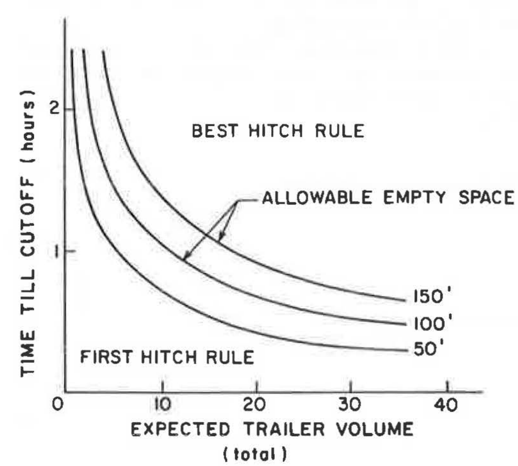
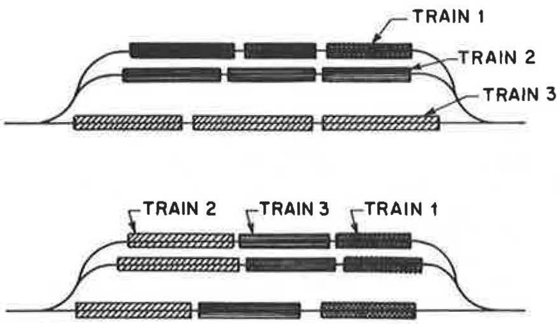


Figure 2. Alternative blocking arrangements for apron tracks.



zontal axis represents the total volume expected for that train or block. The curved lines on the graph represent the railcar length that can be skipped in order to load an arriving trailer on the best hitch available.

For example, if 20 trailers were expected, and 48 min remained until cutoff, it would be feasible to leave up to 110 ft of the train empty in anticipation of future trailer arrivals in order to use the best-hitch rule. If selection of the best hitch required leaving more than 110 ft, then the first-hitch rule would be used. The simulation model would be used to determine the nature of the trade-off for each terminal, as well as for each block and trailer type if this level of detail was necessary. In addition, the potential benefit of this decision aid could be evaluated. This would depend on the consistency of trailer arrival patterns and the complexity of the trailer and railcar fleet.

A second application of this model is related to the development and testing of alternative arrangements for the loading and assembly of trains. Figure 2 shows two possible methods for the loading of three trains, each of which has three blocks. Three tracks are available in the terminal. In the first method, each train is loaded on a single track, which can result in either empty cars or insufficient cars, both of which require switching. The other method assigns one block per track but requires that all tracks be shut down during switching. The relative advantage of a method depends on volume characteristics, schedule timings, and switching, and it could be determined by repeated simulation.

A final application of the model could be to determine the impact on hitch use of changes in the character of the trailer and railcar fleet. This would involve changing the characteristics of the operating environment of the model, which includes the trailer arrival distributions and the strings of cars that are available for loading. The loading activity could then be simulated to determine the impact of these changes and the results used in the evaluation of the effectiveness of various additions to the current railcar fleet.

These three examples indicate the potential of the model for providing decision support in situations that require repetitive short-term decisions, medium-term operating policy decisions, and long-term capital investment decisions. In each case, the evaluation is based on a detailed analysis of the situation that exists at a terminal during the loading phase.

SUMMARY

The model structure that has been outlined in this paper provides a framework for the analysis and improvement of certain terminal operating procedures. The foundation of the model is a procedure for forecasting and updating the volumes of trailers to be handled. The short-term uncertainty relating to outbound trailer volumes can be one of the major causes of terminal inefficiency, particularly with respect to hitch use. This uncertainty is incorporated into the model structure and is used in the assignment of railcars to hitches.

ACKNOWLEDGMENT

The research reported in this paper was supported by a grant from Transport Canada through the Universities Transport Program. Additional scholarship support was provided by DeLCan through the Roads and Transportation Association of Canada scholarship program. Data used in the testing of the model were provided by the Department of Operational Research, CN Rail. In addition, countless railroad personnel in terminals and headquarters spent many hours describing their operations and showing me their facilities. It is their efforts that have provided an abundance of raw material.

Draft versions of this paper were read by Bruce Hutchinson of the University of Waterloo, John Nichol of CN Rail, and John Newland of CP Intermodal Services. Their comments, together with those of a number of referees, have added substantially to the quality of the text. Of course, any errors or omissions are my responsibility alone.

REFERENCES

1. S.E. Shladover and K. Wilkie. Intermodal Freight Terminal Simulator: Requirements Definition. FRA, Rept. FRA/ORD-81, 1981. NTIS: PB 81-247652.
2. G.P. Engelberg and C.R. Yager. Canadian National Trims Cost of Yard Expansion Programs Using Computer Simulation. Canadian National Railways, Montreal, Quebec, Canada, Draft Rept., 1982.