REPORT NO. DOT-TSC-FRA-77-25

PB 279 860

USE OF COMPUTER SIMULATION FOR THE ANALYSIS OF RAILROAD OPERATIONS IN THE ST. LOUIS TERMINAL AREA

E. William Merriam

Bolt, Beranek, and Newman, Inc.
50 Moulton Street
Cambridge MA 02138

NOVEMBER 1977
FINAL REPORT

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL RAILROAD ADMINISTRATION
Office of Federal Assistance
Washington DC 20590

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

Technical Report Documentation Page

1. Report No. DOT-TSC-FRA-77-25	2. Government Accession No.
3. Recipient's Catalog No.	PB279860
4. Title and Subtitle	USE OF COMPUTER SIMULATION FOR THE ANALYSIS OF RAILROAD OPERATIONS IN THE ST. LOUIS TERMINAL AREA
5. Report Date	November 1977
6. Performing Organization Code	
7. Author's	E. William Merriam
8. Performing Organization Report No.	DOT-TSC-FRA-77-25
9. Performing Organization Name and Address	Bolt, Beranek, and Newman, Inc.* 50 Moulton Street Cambridge MA 02138
10. Work Unit No. (TRAIS)	RR739/R8302
11. Contract or Grant No	TSC-13305
12. Sponsoring Agency Name and Address	U.S. Department of Transportation Federal Railroad Administration Office of Federal Assistance Washington DC 20590
13. Type of Report and Period Covered	Final Report Dec 76 - May 77
14. Sponsoring Agency Code	
15. Supplementary Notes	*Under contract to: U.S. Department of Transportation Transportation Systems Center Kendall Square Cambridge MA 02142

16. Abstract

This report discusses the computer simulation methodology, its uses and limitations, and its applicability to the analysis of alternative railroad terminal restructuring plans. Included is a detailed discussion of the AAR Simulation System, an overview of twelve other railroad simulations, and an analysis of how they or other simulation systems might aid the restructuring project being conducted by the railroads in St. Louis. Included is critical analysis of what "validation" of simulation means and what it does and does not imply. Also discussed is the meaning of the terms "network" (as in network simulation) and "levels of detail." Simulation builders and railroaders view these terms differently, which often results in disappointment with the results of supposedly "successful" simulation ventures. The importance of user familiarity with both the simulation system and railroad problems is stressed. A major conclusion reached is that none of the existing network simulations is suitable for detailed analysis of railroad terminal areas. Development of a simulation system incorporating a new approach for performing such analysis is within the state-of-the-art and is recommended.

17. Key Words	Simulation, Railroad, Terminal, Freight Operations
18. Distribution Statement	DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	
Unclassified	Unclassified	78	A05-A01

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This assessment of the use of computer simulation for the analysis of operations in railroad terminal areas was performed for the Transportation Systems Center of the U.S. Department of Transportation. It was performed as part of the Rail Systems Analysis project (PPA RR 727), and had its genesis in the Federal Railroad Administration's Office of Federal Assistance as a result of a request by Mr. Richard Crisafulli of that office.

The author wishes to acknowledge the helpful guidance provided by Mr. Kenneth Troup and Ms. Laura Baker of TSC as well as Mr. Crisafulli of FRA. Most of the data in this report dealing with the St. Louis restructuring project have been supplied by Mr. Chandler Lewis of CONSAD Research Corporation. The opinions expressed in this report, however, and any errors or omissions are solely the responsibility of the author.

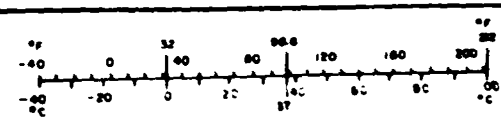
METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

EXECUTIVE SUMMARY

The Federal Railroad Administration (FRA) is involved in a major joint project with the railroads in St. Louis to develop a plan for restructuring the configuration, operations, and/or corporate arrangements in the St. Louis terminal. The complexities of operations there, plus past use of computer simulation techniques in the analysis of rail operations, prompted the FRA to seek information on the applicability of simulation systems to the restructuring of the St. Louis terminal. Since the Association of American Railroads Simulation System had been employed during a 1972 study of St. Louis, the FRA also wished to know if it or any of the results obtained from it would be applicable to the current restructuring effort. This report thus discusses the simulation methodology in general, its uses and limitations in railroading, and its specific applicability to the analysis of alternative terminal restructuring plans. The report includes an analysis of how existing railroad simulation systems might aid in the study underway in St. Louis.

The major train movement activities which take place in a terminal area are reviewed in the report in order to show their interactive and often interferring nature. Changes in these activities or in the terminal's physical limitations may be analyzed in isolation, but determining the effect of such changes on overall terminal operations is of major

importance. Computer simulation is a way of helping with this difficult task.

The use of and complications involved in the simulation of railroad operations are reviewed. Some of the problems addressed include problem representation, the level of detail represented, validation, and the interpretation of results. It is particularly important that the individual who interprets the results of the simulation is intimately familiar with the details of both how the simulation works and of the railroad-related problems being addressed. This has been a common failing in the past use of simulation. Additionally, substantial costs for development, operation, and data collection must be expected regardless of whether an existing simulation system is used or a new one is built. Even with these negative aspects, though, simulation can have important application to railroading. Historically, the simulation methodology has been successfully employed to analyze railroad operations at a global level. The more detailed interaction among railroads in terminal areas appears to be another area of potential application for simulation.

All of the existing simulation systems are similar in the way they model railroad operations and produce outputs, with some differences in application as a function of the original purpose for each system. Misunderstanding about the nature of network simulation systems often results because the track configuration is not represented in detail in existing systems and many other features of railroad operations are only indirectly represented. In terminal operations analysis, these track configurations and other features are critical. As a result, none of the existing network simulation systems is viable for terminal area analysis, and none is judged to be appropriate for the kinds of analyses required in the St. Louis restructuring project. This is due primarily to the fact that the existing systems were not designed for terminal analysis. Though successfully used in more global railroad analysis, these systems do not model the interactions among railroads which are critical in the terminal area.

The data requirements for a meaningful simulation of the St. Louis terminal area are far in excess of data currently available in the restructuring project. The material collected by Parsons Brinckerhoff in the first study is out of date and cannot be used. The more recent CONSAD data are useful for several types of simulation as discussed in the next paragraph, but more has to be collected before simulation could begin. In any case, the outstanding data requirements are substantial.

Four areas of terminal analysis are potential applications for simulation. Rather than being four applications for a single simulation system, the areas represent four different types of simulation requirements. The areas are:

1) Overall evaluation of the terminal and development of terminal design parameters. This is the Phase I activity being performed for the FRA and the St. Louis railroads by CONSAD Research Corporation using a manual analysis technique with limited data collection. No simulation system exists which is appropriate for the kinds of problems that are presented in this phase primarily because current systems represent most of the important terminal characteristics indirectly and at a level of abstraction that is too far from reality. A simulation system developed for the applications described in 3 and 4 below would incorporate features that would satisfy the requirements of overall terminal design. If subsequent terminal analyses are conducted, the use of that simulation to evaluate and develop the overall parameters of the terminal is recommended. However, it is too late to use such a simulation system in the St. Louis Phase I study.

2) Detailed design of yards. All of the four yard simulation systems reviewed could be successfully employed for such analysis. The more limited nature of yard simulation reduces the possibility of misunderstanding indirectly represented facts, and makes the data requirements and operating costs less than for the larger network systems. Therefore, use of a yard simulation system should be seriously considered whenever the details of yard design or redesign are being studied.

3) Detailed junction and corridor design. Because current simulation systems do not directly represent track and certain other features, they are not appropriate for conducting the type of detailed simulation which is needed for the design to be conducted in Phase II of the St. Louis restructuring project. Development of a simulation system incorporating a new approach for performing such analysis would be of value to the conduct of Phase II. It is recommended that such development take place within the constraints of the Phase II schedule, project funding limitations, and data availability from the railroads. The FRA contractor for Phase II should give serious consideration to incorporation of such a simulation into the study. Design and implementation of the simulation system must be carefully performed, and high development and data costs, on the order of $400,000, can be expected.

4) Overall detailed terminal evaluation and design. This area represents a combination of the simulation applications above. All yard and connecting trackage would be explicitly simulated as would train operation. No simulation exists for this purpose, nor is development of such a simulation at this time recommended. It is likely, however, that the system described in 3) above could be built so that it can be expanded into the broader, more detailed analysis, if required.

1. INTRODUCTION

Over the past several years, there has been considerable interest in restructuring the railroad operations of terminal areas in general, and of the St. Louis terminal area in particular, so as to increase efficiency, lower costs, and improve the environment both sociologically and ecologically. In 1974, the East/West Gateway Coordinating Council and the joint venture of Parsons, Brinckerhoff, Grotz and Eric Hill, under Federal Railroad Administration (FRA) sponsorship, completed a comprehensive study of the St. Louis terminal area (1). Included in the study was the evaluation of three major alternative reconfiguration plans, using the Association of American Railroads Simulation System (AARSS) (2,3). For various reasons, none of the three alternatives was implemented, and the FRA currently is sponsoring a new study of the St. Louis area which is being performed by CONSAD Research Corporation and a Technical Advisory Committee (TAC) of railroad personnel from St. Louis (4,5). This study (called "Phase I") will develop overall design parameters for the terminal, will provide data on one or more alternative configurations which the TAC and FRA can review, and will provide a Statement of Work for Phase II in which the detailed engineering design and evaluation of the alternative(s) chosen will take place.

The TAC decided not to use computer simulation during the Phase I study, primarily because of the severe time constraints for its completion. Since Phase I is mainly concerned with defining overall terminal parameters and choosing one or two possible global configurations from a small set of alternatives, it was felt that the necessary analytical processes could be accomplished using simple manual methods. However, the question has arisen as to what sort of simulation, if any, should be used in Phase II since it is here that the final configuration will be specified and where it detailed design will take place. This question breaks down into several parts, as follows:

1) Is the AARSS used in the previous St. Louis study, applicable to the Phase II effort?
2) Are there other existing simulation systems which are applicable?
3) Should a simulation system be developed specifically for the St. Louis terminal effort and, if so, what should its characteristics be?
4) If the answer to any of the above questions is "yes", what level of funding is required for implementation?
5) What implications do the answers to the above questions have in considering the use of simulation for railroad terminal areas in general?

The purpose of this report is to provide detailed answers to these and similar questions concerning railroad terminal area simulation. Before such issues can be meaningfully discussed, it is necessary to understand the terminal area environment to which the simulation will be applied and the kinds of questions which are to be answered when studying possible reorganization. For this reason, Section 2 discusses railroad terminal areas and the problems associated with their efficient operations. (Notice that the emphasis here is on operations, since that is the area in which systems such as the AARSS are applicable, as opposed to areas such as community impact, organizational structures, and the myriad of other factors that must be balanced when making reorganization decisions.) With this understanding of the railroad terminal environment, Section 3 then discusses the simulation methodology. Details of the AARSS and other systems are discussed in Section 4, and Section 5 discusses the applicability of these systems to the St. Louis Restructuring project and to terminal areas in general.

2. TERMINAL AREAS AND RESTRUCTURING

2.1 Why Terminal Restructuring?

The typical railroad terminal area consists of interconnected trackage serving a variety of facilities including industrial areas, classification yards, storage yards, repair facilities, and stations. The current configuration of any terminal area is the result of over a century of evolution and change as commodity markets have come and gone, as railroads have merged, and as competition has forced individual railroads to create facilities and adopt particular operating procedures. As a result, terminal areas today have a large amount of abandoned trackage, parallel rights-of-way, and different facilities performing similar functions. Lower profitability has caused individual railroads to seek ways to lower costs while maintaining or improving the level of service to their customers. Thus, focus has been placed upon major mergers between heretofore competing railroad companies and upon joint operating agreements between companies. Since terminal operations represent a sizeable cost to most railroads and since the kinds of facilities required in a terminal are nearly identical from one railroad to another, it is only natural that an attempt be made to consolidate many different competing terminal facilities into a small number of cooperative facilities and to abandon those which are not needed. Such is the case in the St. Louis terminal area. Here, seventeen competing railroads are cooperating to develop a restructuring plan which will improve each of their operations as well as the operations of the terminal as a whole, and will, at the same time, reduce the associated costs. The remainder of this section addresses the operation of terminal areas in general, but draws upon the St. Louis restructuring project for specific examples and illustrations.

2.2 Terminal Operations

The major activities which take place within a terminal area are:

1) Run through trains: Trains enter the terminal area and leave it again with little or no change in the train's consist. Engines, cabooses, and crews may or may not be changed, and crew changes often occur with the train in motion. Safety inspections may be made, depending on the distance the train has traveled since its last inspection.
2) Inbound road trains: Trains originate at some distant point and arrive at a yard within the terminal area, with few or no consist changes having taken place along the way.
3) Outbound road trains: Trains are made up at a yard within the terminal area and go, with few or no changes in consist, to some point outside the terminal area.

4) Local trains: Trains originate, terminate, or both within the terminal area and make frequent stops to pick up or set out cars at the various industries along their routes.

5) Transfer trains: Trains run between two yards of the same railroad, both of which are within the terminal area.

6) Interchange trains: Trains deliver cars from one railroad to another. Usually these trains originate and terminate at yards, but it is also common for them to operate between lesser points, sometimes simply into or out of a designated siding.

7) Car classification: Trains arriving at a yard are broken up and the cars are reassembled into trains going toward the car's ultimate destination.

8) Car and locomotive cleaning and repair: Regularly scheduled maintenance and repairs necessary to make a car road-worthy are performed.

9) Car holding: Empty cars, cars waiting for waybills, maintenance-of-way equipment, etc. are stored until they are needed or can be moved.

With all of these activities occurring simultaneously, there are many situations in which work designed to accomplish one task can interfere with another. Furthermore, anticipating conflict situations is difficult since the individual railroad companies operating within the terminal area almost always dispatch trains independently of one another. Also, individual yardmasters often do not know of arriving trains until they are only a few miles from the yard, and even then they usually don't know the train's consist until it is inside the yard. For these and many other reasons, it is important that the trackage within a terminal area be designed for efficient operation and maximum flexibility. However, as terminals have grown over the years and more and more trackage has been squeezed into tighter spaces, operating efficiency has suffered and terminals, such as St. Louis, that used to be smooth-flowing gateways have become major bottlenecks.

2.3 Impediments to Efficient Operation

2.3.1 Yard and Corridor Capacity Limitations

One of the reasons for clogged terminals is that the yards and the trackage connecting the yards (called "corridors" in the remainder of this report) do not have the capacity to accommodate the large volume of traffic that must flow through the terminal.

Corridor capacity is a function of the number of tracks in the corridor, the location and frequency of passing tracks, the number and complexity of junctions that must be passed through, and speed restrictions. (The latter are due to grade and curve limitations, type of signalling systems, quality of track and roadbed, and proximity to conflicting traffic.) When a given corridor's capacity is exceeded or operations over it are slowed down or halted (e.g. due to a derailment), other corridors are affected because traffic cannot flow out of them into the affected corridor. This process can be repeated until a sizeable portion of the terminal is tied up. Of course, one way to avoid such huge tie-ups is to hold trains in yards until the way is clear. However, if yard capacity is exceeded, it will be necessary to hold trains out of the yards, further tying up the corridors. If capacity is so thoroughly exceeded that this occurs, there probably is no solution but to hold trains outside of the terminal area. Then, there is the problem that the entire railroad system will become clogged because of the congested terminal; but fortunately, such events usually only happen during bad weather or other emergencies (e.g. the massive grain export crisis in the midwest in 1973).

Yard capacity is determined by the number of tracks available to serve such activities as receiving, departure, make up, classification, maintenance, cleaning, and holding. The interconnection of these various tracks is also important. A well-designed yard will allow operations to take place simultaneously in all of its areas, and will also allow one under-utilized facility to temporarily be used to accommodate overflow from another. Important considerations for classification yards include whether flat or hump switching is done, how many leads or humps are available, how the approaches to the yard tracks are arranged, and how much interference is caused by simultaneous operations (e.g.

switching one classification track while pulling another).

When yards and corridors are designed, the above issues and more are taken into account. However, because of changing operating policies, new developments (such as larger and heavier cars) and higher volumes, terminal trackage often tends to operate beyond its capacity limitations. The St. Louis terminal is no exception to this, and major emphasis is being placed in the restructuring plan to insure that capacity will be available to meet present and future traffic volumes.

2.3.2 Junctions

Whenever one track meets or crosses another, there is a possibility for conflict between two trains such that one will have to stop while the other passes. Such conflict situations are aggravated in terminal areas because there are often many crossings or junction points within a short distance of one another; there is a high density of trains operating through these areas; the trains are often long (on the order of 100 or more cars); and the trains usually operate at slow speeds.

In St. Louis, there are several junctions that are particular problem areas because of their complexity and high volume of traffic. In one of them, CP Junction, the switching lead of Madison Yard crosses three other tracks. Thus, whenever long trains are being made up or otherwise switched in the yard, these three tracks will be impassable.

This situation, combined with one other long train, could completely tie up this junction which in turn would cause back-up throughout the terminal. Since the current St. Louis restructuring plan proposes that Madison Yard be expanded, solving this interference problem with CP Junction is a major priority.

Valley Junction and Q Tower are two other junctions which are located along several of the busiest corridors in the terminal. As with CP Junction, bottlenecks can occur causing traffic to back up into other areas of the terminal. Some of the problems might be solved by a better track configuration, by coordinating dispatching to the approaches to the junction through some central authority, or by relocating certain facilities so that it is no longer necessary to have so many trains going through the junction. In reality, some combination of them may be required, but more study is needed before such a determination can be made.

2.3.3 Bridges

Bridges and tunnels represent potential bottleneck situations. This is because of the inability to relocate them as traffic patterns change (due to the prohibitive expense involved), and because of the small number that are available to handle a large volume of traffic.

In St. Louis, there are only three railroad bridges crossing the Mississippi River. All of the east-west traffic funnels to the approaches to these bridges and then fans out on the other side. This results in a high density of traffic at several junctions as well as along certain corridors close to the bridges. It is not yet clear what impact the restructuring plan will have on this problem, but further study by CONSAD is underway.

2.3.4 Highway Interference

Grade crossings are a major problem for railroads, both from a safety and operational point of view. Often local ordinances dictate the length of time a road can be blocked by a train, and this in turn limits the length of slow moving trains. Some towns even have ordinances which limit the hours during which roads can be blocked at all, thus imposing severe limitations on the railroad operation. A further problem is that if it should become necessary for a train to stop for any length of time while spanning a grade crossing, it may have to be split to allow traffic to pass. A delay in starting would then be encountered due to the recoupling operation. These problems are compounded when several roads cross the same track at close intervals. There are at least two such serious grade crossing areas in St. Louis and fourteen others which require attention. Part of the restructuring plan already deals with some of these, while further work is required before they all can be taken into account.

Railroad tracks in streets are also a problem both for rail and auto traffic. Such an arrangement is necessary in order to service industries crammed into high-value land. (Usually this situation arises in waterfront and dock areas.) It also occurs when no other access route is available to certain important facilities.

Operating problems are caused by the mixing of auto and train traffic so that congestion of either affects the other. Also, a car parked so as to leave too little clearance for trains will halt operations until it can be moved. In St. Louis, a particularly bothersome track is one running down the middle of a street which provides access to a yard. Unfortunately, the only other access to that yard is over a bridge, with all of the attendant problems mentioned above. Thus, it is not clear what can be done about this track except to alter traffic flows so that fewer movements over it will be required. Of course, the reverse could also happen, and more movements may be required. This is one area that will be further analyzed as the restructuring plan develops.

2.4 Measures of Terminal Efficiency

It is clear that each of the items mentioned in the preceding section has an impact on overall terminal operations, but it is not clear just what that impact may be. For instance, if trains crossing the switching lead mentioned in Section 2.3.2 normally experience a delay of 15 minutes, we might be tempted to conclude that eliminating the delay would increase the efficiency of the terminal by that amount multiplied by the number of trains delayed. Though this is probably the case, such a change may have unexpected repercussions. Consider the situation where the crossing tracks feed a very small yard which cannot handle many incoming trains at once. It very well may be that the 15 minute average delay is just what is needed to regulate traffic so that this yard can operate efficiently. Moreover, the functioning of this yard may have significant consequences for the rest of the terminal.

The preceding example is not intended to discourage the solution of local problems, but rather to point out that changes made in one part of the terminal can have an impact far removed from where the change takes place. For this reason, it is necessary to look at terminal operations from an overall as well as a local point of view. Since it is usually easy to assess the local costs and effects of changes, the remainder of this section will deal with the problem of determining their effects on the terminal as a whole and specifically with various measures that can be used to assess these changes.

One of the problems in discussing overall terminal efficiency is that there are many definitions of what it means to run a terminal efficiently. It is probably the case that no one definition is correct in itself, but rather that the key to an efficiently operating terminal is the balancing of many parameters. This balancing often reflects policy decisions that in turn reflect goals to be achieved. For this reason, the balance desired must be decided by those operating the terminal. Among the considerations that must be taken into account when making such decisions are:

1) Car throughput: The number of cars that can be moved through the terminal in a given period of time.

2) Car transit time: The length of time taken for a single car to move through the terminal. The mean, standard deviation, and maximum transit times are all of interest. (An interesting operational problem is created in St. Louis due to the large transit time of piggyback cars. In these cases, the loads are taken off the flat cars at the point of arrival in the terminal area and are driven to the departure point and reloaded. This operation apparently can be done by truck faster than by train. As well as resulting in lost revenues, it requires that each of the railroads maintain complete piggyback facilities which are expensive in terms of capital outlay and operating costs).

3) Car miles traveled: The smaller this number, the less the cost of operating the car fleet, primarily because of reduced maintenance costs.

4) Per diem costs: An indirect measure of how long cars stay in the terminal.

5) Vandalism and theft losses: These losses usually occur while cars are not moving or are spotted at out-of-the-way places. Such costs can be very large and should be taken into account when planning terminal operations.

6) Crew costs: A complex function of hours on duty and various union work rules.

7) Engine costs: The fixed cost of the engine plus operating costs per hour.

8) Paperwork costs: Handling waybills, cars without waybills, etc.

Unfortunately, most of these considerations conflict with one another. For instance, reducing the transit time may require taking a longer, less congested route, thereby increasing the car miles traveled. One of the most difficult tasks in planning terminal reorganization is to take into account all these different factors that contribute to terminal performance. Computer simulation is often suggested as the right tool to help with this difficult task. But just as there are different kinds of hammers for different jobs, so too are there different kinds of simulations. Whether any of the existing simulations can handle the task at hand is the subject of the remainder of this report.

3. THE SIMULATION METHODOLOGY

A natural question arises as to whether there is any way to verify that various terminal restructuring ideas will work before committing huge sums of money to their implementation. In principle, computer simulation techniques offer the opportunity to do just that, but they must be applied carefully and with a full understanding of their operation and limitations. The first part of this section discusses what computer simulation is and how it might aid in the design of a railroad terminal area. The last part discusses the simulation methodology in more detail as well as some problems and limitations associated with it.

3.1 Simulation's Potential

Simulation provides a method whereby the physical facilities, traffic, and operational decisions of a railroad can be represented in a computer and the railroad can be "operated" using that information as a basis. The facilities, traffic, and/or decision rules can be altered as much or as little as desired and the new configuration can be operated. Thus, many different strategies and physical layouts can be tried, their results compared, and an operational evaluation can be made without altering the current railroad operations.

In simulating a terminal area, it would be desirable to include many of the items mentioned in Sections 2.2 and 2.3. Ideally, the geometry of yards and connecting tracks would be represented including the details of the junctions. Further, such details as highway grade crossings and rules which dictate when and for how long they can be blocked, areas of high vandalism, and the location of piggyback and other important industrial facilities would be included. Then, trains would be "run" over this geometry. They would pass over corridors, be made up, broken down, and reclassified in yards, and perform run-through and switching movements. In short, represented within the computer would be all of the significant details of a running terminal. From time-to-time, reports would be printed indicating the status of yards and other facilities, and at the end of the simulated time period, summary reports would analyze throughput, transit time, and all of the other parameters mentioned in Section 2.4.

In a truly ideal simulation system, it would be possible to watch all this as it is happening and even to influence the progress of the simulation if on-the-spot decisions needed to be made. When problem areas would develop, or more detailed analysis of a given region would be desired, it would be possible to "zoom-in" on that one area and get more detailed information about it. Such a system would allow flexibility in the design and analysis of terminals in that it would be relatively easy to assess the impact of proposed changes on overall terminal operations, and to experiment with changes until a desirable operating pattern is achieved. It is even possible that the simulation could be used later on to aid the day-to-day operations of the terminal by helping to predict problem areas before they actually occur.

3.2 A Realistic Look at Simulation

The preceding scenario is within the scope of current technolc ', but such grand tools do not come without some strings attached. In order to be effective the simulations must be thoroughly understood by those using them, and the costs of their development and operation are quite large. The remainder of this section elaborates these points.

3.2.1 Some Potential Problem Areas

Simulation is not a problem-solving system designed to give objective answers to specific questions. Rather, it is an experimental methodology designed to give insights into what is usually a very complex and incompletely specified problem. It is a tool and, like all tools, it requires a skilled user to realize its potential. Unlike normal tools however, simulation will produce seemingly good results without requiring skill on the part of the user. It is the interpretation of these results that requires skill in both the details of how the simulation system works and in the subject matter to which it is applied; for it is the user, not the simulation system, who is going to generate the solution to whatever the problem may be.

3.2.1.1 The Validation Problem

The first step often employed when using the simulation methodolcgy is to show that the simulation system being used is capable of faithfully reproducing the results that are actually observed in the real world. This step is called "validation", a word which implies that once the system has successfully passed this phase, the results produced by it will be valid. Unfortunately, no such implication can be drawn. It is not even necessarily true that after validation the simulation is an accurate representation of the real world. Moreover, it is not possible to be sure that its inaccuracy is limited to the difference between the simulated and observed measures of actual performance. In fact, about the only thing that can be said is that if the system does not perform well in the validation phase, then it probably cannot simulate alternative configurations of the same general situation.

Consider, for example, the case where a simulation is being done of a single track corridor connecting two yards. The single track has several passing sidings along it and crosses two other tracks. Schematically, it may look like Figure 3-1. The question being considered is how to increase the capacity of that trackage.

The validation phase might proceed by collecting data on the current operations along the corridor including histories of actual train movements, average delay times associated with the various stretches of track and the crossings, etc. The geometry of the corridor would be represented and operating rules would be established to

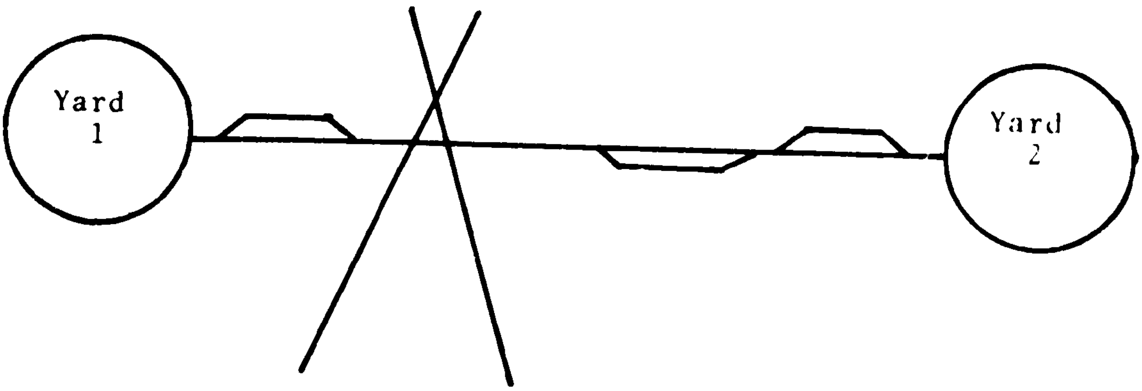


Figure 3-1 - EXAMPLE OF A NETWORK TO BE SIMULATED

coincide with those followed by the actual trains. Such rules would indicate what classes of trains have priority over others and delay times would be apportioned on this basis. Once these data are collected and fed into the simulation, results might be produced that show the average time for a train to traverse the corridor, the average delay encountered by the trains, and the amount of traffic handled per unit period of time. These data would then be compared with actual corridor data and, if the two sets are close, then the simulation would be considered to be valid.

The problem that arises is what it means for the two sets of data to be "close". It is easy to say when the sets are much too far apart. This usually arises when there is some gross error in the simulation. Such an error is usually easily detectable. However, when the data sets are close, that closeness may hide several subtle errors which mostly cancel out during the validation phase, but which may grossly distort the simulation of alternatives. In our example, this situation could easily arise if the train delay times for the crossings and the corridors were individually incorrect, but when combined, produce nearly the correct total delay time. Any simulated alternative which uses those delay times individually would then produce meaningless results, even though the validation simulation seemed to work.

3.2.1.2 Problem Representation and Interpretation of Results

The way a problem is represented in the simulation has a significant impact on the kind of results that can be obtained. Thus, in the example above, it may be desired to determine what effect a double track would have between say, Yard 2 and the crossing tracks. Suppose one decided to represent a double track by changing the delay times associated with the trackage between the two end points. In this case, the delay time for opposing trains might be reduced, while the delay times for passing and conforming trains would remain the same. The simulation could then be run and the effect of the reduced delay on the overall corridor determined.

Notice, however, that the results produced would not say anything about double trackage. They would merely indicate what would happen to the operations in the overall corridor if that particular link were changed in such a way that the given delay times were achieved. It was the user who had the double trackage idea in mind when he developed the delay time parameters. There very well may be other ways to achieve the same delay times, and it is even possible that double trackage will not achieve them. Thus, with this type of representation, it would be false to conclude anything about specific track arrangements. Instead, only a more general statement about delay times can be made, with the detailed track design left until later.

While this might seem like an obvious point, it is very often overlooked. While working with a simulation, it is very easy to believe that the parameter values chosen really do represent exactly that situation which is currently in mind. Unfortunately, this is more often false than true, and many incorrect conclusions have been reached because those interpreting the results have either failed to recognize this error in their own thinking, or because they were unaware of how the simulation actually handled its data. This latter case usually arises when the person doing the interpretation is different from the person who set up the simulation. The latter may have assured the former, for example, that that link was simulated as double track, but such a statement may have vastly different meaning to the two parties concerned. This is the reason for the earlier statement that the person interpreting the results must have intimate knowledge of both the simulation system and the subject matter to which it is applied.

3.2.1.3 Levels of Detail and Data Requirements

The kinds of questions to be answered by a simulation dictate the kinds of data that must be collected. It is usually the case that more data are required by a simulation if more detail is to be provided. It is, therefore, necessary to balance the requirements for detailed information with the cost of gathering the input data.

As an example of what is meant here, consider the illustration of the previous section. There, the general type of question being asked was: "What is the overall effect on the corridor if the parameters of a specified section of track are changed by a certain amount?" A simulation to answer this question would require a fairly simple geometric representation of the corridor, such as is shown in Figure 3-2. Here, the yards, crossings, and connecting trackage are represented as simple entities

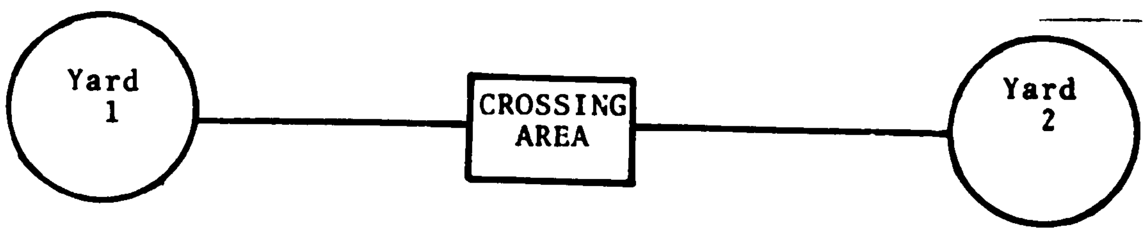


Figure 3-2 - ABSTRACT REPRESENTATION OF EXAMPLE TRACKAGE

connected in a linear manner. Each of these entities has some parameters associated with it such as the maximum and minimum time to traverse the section and average delay time associated with the section. Additional data required by this simulation would be a list of trains including their departure point, departure time, and class, and operating rules which specify when and how the delay times are to be apportioned to the various trains.

Consider now the data required to answer the question: "What is the overall effect on the corridor if the configuration of a specified section of track is changed in a certain way?" In this case, the list of trains would remain the same, but the geometry would have to be represented in all of the detail of Figure 3-1. Also, the operating rules would have to be more complex in order to specify which train is to take sidings and under what conditions, whether trains must wait for other trains, etc.

Thus, a significantly larger amount of data is required for the case where a configuration is being evaluated than for the case where the effects of parameter changes are being studied. Notice, however, that the data required in the configuration change case is much less an abstraction from reality than is that of the parameter evaluation case. This slightly offsets the fact that more data is required since it may be an easier type of data to collect. Also, the simulation of the new configuration will be performing operations that will be more closely akin to those actually taking place along the corridor and it will provide detailed answers beyond the ability of the simulation which only uses parameter values to determine its results. For instance, in addition to testing the specified track configuration, the delay times can be generated for that configuration so that many different configurations can be compared.

The purpose of this section has been to illustrate that the amount of detail simulated and the kinds of questions that can be answered are closely related to the amount (as well as to the kinds) of data required. Also, the issues dealt with by the more detailed simulations will be more closely related to the real-world issues. This latter advantage is highly desirable, not only because it can provide more realistic simulations, but also because the problems of validation, representation, and interpretation are easier to deal with when what is happening inside the simulation is a close representation of what is happening on the railroad.

3.2.2 Cost Considerations

The costs associated with simulation fall roughly into the following categories:

a) Development of the simulation system.

b) Collection of data.

c) Operation of the system.

It is difficult to state which of these categories consumes the most money because that is a function of how sophisticated the system is, how many times it is to be used, and the computing environment in which it is to be run. However, in any case, the cost of simulation is substantial and is a prime factor which has limited its use.

Since simulation is a technique designed to help with problems that are too complex to be dealt with by conventional analytical techniques, the decision of whether to use it or not involves balancing the costs of simulation against the benefits to be received from it. This is a difficult task, however, since both the costs and benefits are hard to predict, and they are very much a function of how a particular project is organized.

3.2.2.1 Use an Existing System or Develop a New One?

Consider the case where an existing simulation package is to be used. Already the cost has been significantly reduced, but so might the benefit to be received. This is because the act of developing a simulation system is much more than a computer programming task (although it is unfortunately often thought of as such). Instead, it is during the development phase that issues are decided regarding such things as the kinds of questions to be answered, the method of representing the railroad data, the levels of detail to be provided, and the type of human interface to be employed. All of these items are extremely important to the success of a simulation venture, and the need to consider them in detail is itself a benefit since it forces the user to have a clear understanding of his problem domain.

One can imagine using an existing system and reaping the benefit of this analysis by doing the analysis first and then looking for a system which provides the necessary capabilities. This is in fact what is usually done, but unfortunately there is often no existing system that meets all of the criteria thus formed. Therefore, it is necessary to use an existing system in a non-optimal manner or to develop a new one from scratch. The former approach usually requires forcing the problem into a framework in which it doesn't really fit; the latter approach is expensive and time consuming. Hopefully, as more and more systems are developed and disseminated, the likelihood of finding one that meets individual needs will be greater, and development costs will be reduced. Until then, the prospective simulation user would be wise to plan for substantial development costs even if an existing system is to be used.

3.2.2.2 Data Collection

Railroad applications of simulation tend to deal with a large amount of data. This is a direct result of the fact that railroads handle a huge number of cars in a multitude of trains between a wide variety of origins and destinations. Further, there are a myriad of operations that are performed by the railroads such as interchange, classification, and cleaning which often account for more time and expense than the actual road-haul operations.

Thus, in order to faithfully capture railroad operations, a prospective user of simulation methods must be prepared to provide a huge volume of data. Depending on the type of simulation being done, these data will have to be gathered by direct observation, by analysis of railroad records, by conversations with operating personnel, by intuition, or by a combination of all of these. Furthermore, as different alternatives are tried, some of the data such as track geometry and train routings will have to be changed. In a well designed simulation system, these changes can be accomplished with a minimum of effort, but in others, the effort required to simulate alternatives may approach that of the original validation simulation. Once again, the simulation user must thoroughly understand the system being used and must plan for substantial costs associated with the data collection phase.

3.2.2.3 System Operation

Once the system has been developed and the data collected, the simulation can be run and the results analyzed. The major expense during this phase is the cost of computer time, which can be considerable depending on the complexity of the simulation. Fortunately, this cost is usually easy to identify in advance and is rarely overlooked. However, often the number of simulation runs desired is underestimated, especially considering the experimental nature of the simulation methodology. As a result, much money and effort can be spent in developing a simulation, only to have it go unused because too little was budgeted for its operation.

Another expense that might be incurred during the operational phase is that of additional computer programming required to present simulation results in ways required by the specific problem. In a sense, this is an extension of the development effort, but it is impossible to predict in advance all of the types of reports and summaries that may be desired. For this reason, many simulations provide a means whereby the user has access to a trace of events happening while the simulation is being run. By writing appropriate programs, these data can be analyzed in a wide variety of ways and can be presented in forms which are tailored to the user's needs.

3.2.3 Is it Worth All of the Problems?

The above sections have purposely emphasized the negative aspects of simulation partly to compensate for the often inflated claims of simulation proponents, partly to provide a basis for understanding the so-called simulation "failures" of the past, but mostly to provide a realistic setting within which to evaluate the details of simulation systems. If, after reviewing all of the negative aspects of simulation, it is still possible for a prospective user to justify its use, then it is probably the case that that venture will be more successful than would one that only emphasized the positive aspects.

Before leaving this section, let us briefly list some of the positive benefits of simulation:

a) Simulation is an experimental methodology that is often the only way to handle complex problems. The requirements imposed by the methodology upon the user underscore the fact that the purpose of simulation is to help generate intuitions from which problem solutions can be proposed. Such solutions are usually easy to develop, but the intuitions that let one know that the proposed solution is correct are often difficult to come by. It is sometimes the case that the mere act of developing the simulation generates these insights and that the actual running of it becomes unimportant or anticlimatic.

b) The act of putting a simulation together forces an attention to detail that may not otherwise be present in a project. This often leads to alternatives and solutions that may not have otherwise been considered.

c) Not withstanding the discussion in section 3.2.1.1 on the validation problem, it is in fact possible to convince oneself that a particular simulation accurately reflects the operating characteristics of a given situation. In such cases, it is indeed possible to "operate" the railroad in advance to determine the characteristics of the proposed changes.

While it is true that the above things are expensive and difficult to accomplish, their value is likely to be great. Thus, the simulation methodology is still a viable way to attack the many problems involving complex railroad operations.

4. EXISTING SIMULATION SYSTEMS

4.1 Network Simulation Systems

This section discusses the so-called "network" simulation systems. The AARSS is probably the best known of these systems and it will be discussed in considerable detail. This system is representative of the other network simulation systems and is of special interest because it was previously applied to the St. Louis terminal area and remains the only system known to this author to be applied to the global evaluation of any terminal area.

Before discussing the AARSS, it is necessary to clear up two issues about network simulation systems that have been the subject of much misunderstanding. These involve the meaning behind the word "network" and the phrase "levels of detail." Unfortunately, these mean something very different to simulation builders than they do to railroaders. Misunderstandings come about because the simulation builder's meaning is an abstraction of the railroader's meaning. Thus, entire conversations can take place with each party in full agreement with the other, only to learn later that they weren't communicating at all.

For a simulation builder, "network" refers to a specific abstract structure used to represent certain kinds of data. In "Network Simulation Systems," it is the track configuration (i.e. the railroad "network") that is represented by this abstraction. Thus, the railroad trackage in Figure 4-1

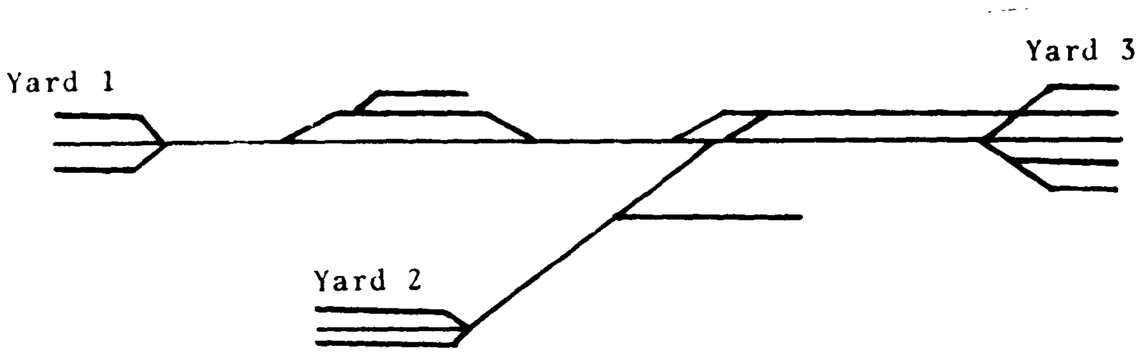


Figure 4-1

would be represented by the network in Figure 4-2.

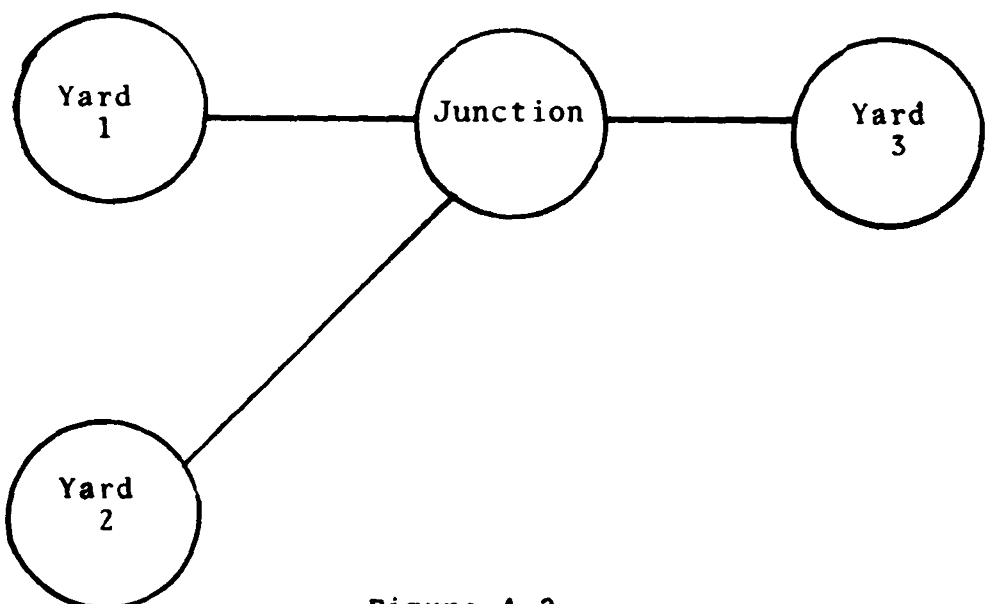


Figure 4-2

Notice that there is no explicit representation of the yard trackage, the passing track, the spur tracks, the double or single track segments, or the junction trackage. Nor is there any indication that it is impossible for a train to take the route: YARD 1 - JUNCTION - YARD 2 (without backing up, at least). Thus, the network representation used in network simulation systems corresponds only in a very limited way to the actual railroad trackage. Perhaps a better and less confusing term would be "connectivity diagram" since what is really represented is that the yards and junction are somehow connected to each other, while the exact details remain unspecified. While this report will continue to use the term "network", an attempt will be made to reduce its ambiguity by inserting the appropriate context.

As to the "level of detail" misunderstanding, it is difficult to read almost anything that has ever been written on network simulation without seeing a sentence which reads something like "...can be represented to any level of detail desired by the user." To railroaders, this means that the trackage shown above might be represented as it is drawn or perhaps in any of the ways in Figure 4-3.

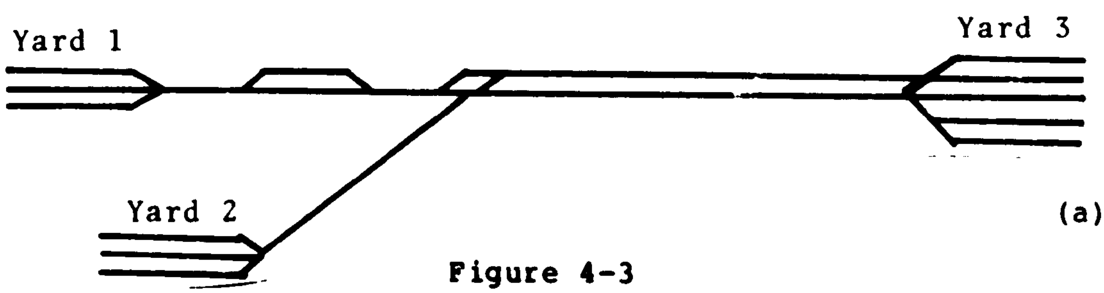
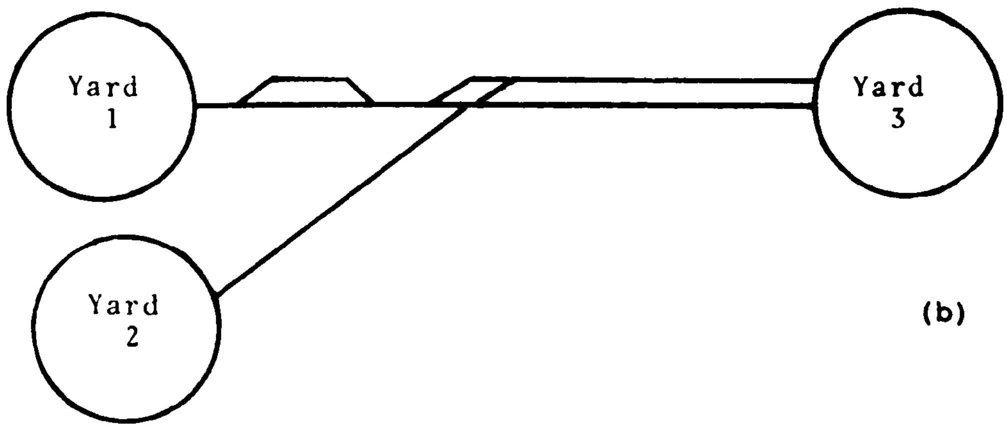
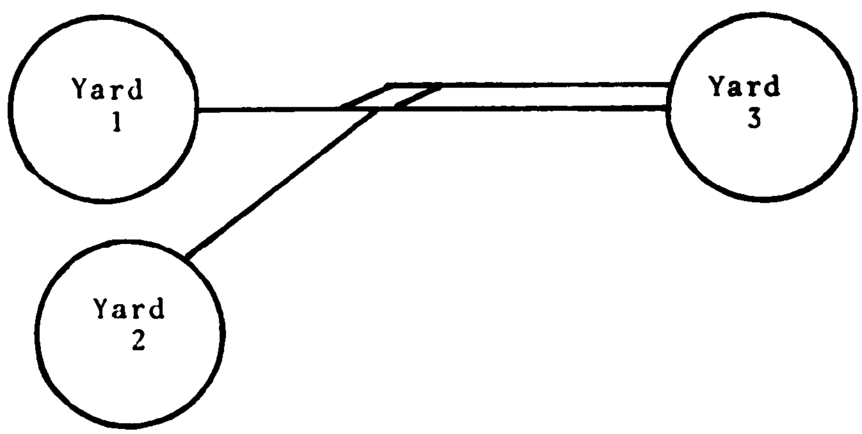


Figure 4-3



(b)



(c)

Figure 4-3 (cont.)

Or it may be represented by some other combination of track segments, depending on what is considered important. This is not, however, what a simulation builder means by "levels of detail", since we have already seen that all of these variations would be represented by the same network structure as shown in Figure 4-2. Instead, what is meant is that the nodes (the circles) or the links (the connecting lines) can have a variety of information associated with them, that the specification of <u>that</u> information is optional, and that, when specified, it can take one of several forms. For example, numbers may be assigned to the links which represent the average delay time that will be associated with a train if it should meet a train traveling in the opposite direction along the same link. The amount of delay may be specified as a constant number, as a function of some variable such as train class, as a random number between some limits, or it may not have to be specified at all in which case a default value will be assumed. It is these different possible ways of specifying the parameters of the links and nodes that are referred to by the phrase "levels of detail."

4.1.1 The AAR Simulation System

This section presents a brief overview of the AAR Network Simulation System (6,7) including a discussion of the kinds of data required by the system and how the system uses these data to produce its output. While the description is necessarily brief and incomplete, it attempts to give enough of the details of the system so that a critical review of its usefulness will have meaning.

4.1.1.1 Input data

4.1.1.1.1 The Network

The railroad network is represented in the AARSS as an abstract network consisting of nodes (the circles) and links (the lines) as discussed above. Nodes usually correspond to yards or industrial areas, but they can also correspond to junctions, industrial sidings, passing sidings, interchange points, or any other kind of geographical area. However, it must be emphasized again that these nodes simply represent places, not functions. Thus, if a node represents a junction, it is only representing it as the place in the schematic diagram of the railroad where it exists. No notion of the function of a junction is to be implied; only its presence can be inferred. Parameters associated with the nodes imply their function and, as will be seen in the next section, the node parameters are limited to specifying functions which typically take place in yards.

Links correspond to the trackage connecting the nodes. Note that there is one and only one link between each node, regardless of how many tracks actually go between the nodes or what the configuration of those tracks may be. Associated with each link are three values that represent the average amount of delay time experienced by trains due to meets, passes, and conforms(*) that occur on the link. The system assumes that any number of trains can occupy a link at a time and that passing sidings are available at any point along the link.

4.1.1.1.2 Facilities, Resources, and Jobs

Each node can have various attributes associated with it that indicate the kind of functions that can be performed at that node. Facilities are areas that trains and cars can occupy, such as receiving, classification, and departure yards, maintenance tracks, holding tracks, etc. Here again, the configuration of the facility trackage is not known to the system· only the capacity (in number of cars) of the entire facility is known. Thus, a one-track facility which can hold 100 cars is simulated the same as a 5-track facility where each track holds 20 cars.

Resources are things or people that are needed to perform jobs (defined below). Examples are: engines, inspection crews, loading docks, hump leads, etc. They are defined by specifying the number of units of the resource that are available.

-----

*Meet: Two trains passing each other while going in opposite directions.
Pass: One train overtaking another while going in the same direction.
Conform: One train staying a relatively fixed distance behind another while going in the same direction.

Jobs are units of work that may be performed at the node and, in terms of the simulation, represent time delays in the movement of trains or strings of cars through the node. Associated with each job is a list of facilities and resources required by it, and a linear function which specifies the time required to perform the job. Examples of possible jobs are receiving yard inspection, car cleaning, classification, and departure inspection.

4.1.1.1.3 Traffic

The traffic in the system consists of cuts of cars that are to be moved from an origin to a destination. Each cut has associated with it an origin time, origin node, destination node, traffic class, and number of cars. When the specified origin time is reached, the cut of cars "appears" at the origin node and becomes a candidate for inclusion in the consists of trains heading toward the destination node. Whether a cut of cars is actually included in a train's consist is a function of the class of the cut and whether the tonnage or car limits of the train have been exceeded (see below).

4.1.1.1.4 Trains

Trains are defined in the system by specifying a task list and a priority. The task list is a list of all of the things that the train is to do, and includes its route and schedule. The task list specifies the class and number of cars it is to pick up or deliver at each node, and the jobs that are to be performed at the node on the train or designated cuts of cars. Included in the part of the task list that contains the train's route is an indication of the tonnage, number of cars, and running time limits for each link along its route. The actual running time of the train on a given link is computed by interpolating between the minimum and maximum running time limits, using the actual consist of the train to compute the tonnage. This time then represents the minimum time that the train can traverse a given link. Depending on train priority, the delays mentioned in Section 4.1.1.1.1 are added to this time to arrive at the actual running time.

4.1.1.2 A Simulated Train Run

To illustrate how the above information is used by the simulation, this section will describe how a typical train is "run" through its task list.

Trains always originate at a node and may or may not already have traffic assigned to them. At the scheduled departure time, the train and its consist become disassociated from the node and are considered to be traveling on the link. While traveling on this link, nothing happens to the train's consist, but it is possible that delays will occur. As mentioned in Section 4.1.1.1.4 above, when the train is initially put on a link, it has associated with it a minimum link transit time. That is, it is not possible for the train to arrive at the next node before this amount of time has elapsed. As long as there are no other trains on the link, this train will indeed traverse the link in that minimum time. If other trains are present, the delay times associated with that link (see Section 4.1.1.1.1) will be added to the time of whatever train (or trains) have lower priorities. For example, if two trains are on the link at the same time heading in opposite directions, the meet delay time associated with the link will be added to the transit time of the train of the lowest priority. This new transit time is then used to determine the time at which the train is removed from the link and handed over to the next node for processing.

A typical scenario for a train arriving at a node would be that it drops off or picks up one or more cuts of cars, undergoes some jobs such as departure inspection and air test, and then either waits until its scheduled departure time or departs immediately depending on the specification in its task list. At the departure time, the train becomes associated with the next link and processing continues as above, with the train traversing links and nodes until it reaches its destination node.

At the nodes, it is not possible for a given job to be started until the required resources and facilities are available to it. Because there are only a limited number of these, it is possible that considerable delays might take place. Resources are assigned on a first-come first-served basis and cannot be reserved except in cases where a lower priority train would delay a higher priority train. Also, all jobs must be performed before a train can depart and they must be performed in the order specified in the train's task list. Thus, if a node becomes congested, considerable delays can result.

The cars that are set off at each node are processed in such a way as to reflect the classification policy defined for the node. Specifically, the various cuts of cars are put into classification groups based on their destination and traffic class. It is these groups (after a suitable delay time) that then become the candidate cuts for other trains to pick up at the node.

4.1.1.3 Output Produced

Seven reports are produced by the system, providing information on:

1) The time it takes cars to get between the various origin-destination nodes.
2) The history of the number of cars contained in the classification groups at the various nodes. This includes the number of cars set off and taken by trains, as well as those originating at the node.
3) The history of trains as they progress along their routes. Deviation from schedule as well as changes in consist are shown.

4) For each or any link, statistics on the number of meets, passes, and conforms that occurred, and the average transit and delay times.

5) The average times spent by trains in each node along with the number of train arrivals, departures, originations, and terminations.

6) The history of operations at the nodes including arrival and departure times of trains, their deviation from schedule, the amount of time gained or lost (with respect to the schedule) through the nodes, and details of the job processing at the node.

7) Utilization statistics of the resources and facilities at the nodes including percentage utilization and total hours used.

In addition to the seven reports produced, the system produces a file of information which contains a history of events during the simulated time period. This file forms a data base which user-written programs can access to produce reports tailored to a specific need.

4.1.2 Other Network Simulation Systems

The AAR Simulation System is only one of many network simulations built over the years, and a natural question arises as to whether there are others that may be better suited to the terminal area environment. After studying several other systems, it has become apparent that all of the systems are basically the same. That is: they all use the same sort of link-node representation; they all make similar assumptions about link delay times and capacity; and they all require basically the same type of train, traffic, and network definition input. Perhaps the major difference in processing is the way in which the systems handle the nodes and the ease with which nodes of different types can be defined. The biggest difference of all, however, is in the ease of understanding the documentation, a factor which could play a major role in the success of a simulation venture. In this regard, the documentation for the AAR Simulation System is especially difficult to use, as the information necessary to understand most details is scattered throughout the manual.

Since the other systems are so similar to the AAR System, they are simply listed here with their references, rather than discussed in detail.

1) C&O/B&O Mini-network (8,9).
2) Missouri Pacific CARS Simulation System (10).
3) Frisco Simulation (11).
4) British Railways Model (12).
5) Canadian National Network Model (13).
6) SCL/L&N Simulations (14,15).
7) SRI Model (16).

The Canadian National Network Model bears special mention because it has several features that make it stand out from the others. Its documentation is very good; it is well written, easy to follow, and appears to be complete. The system has a wide variety of node types and it models yards in greater detail and in a more realistic fashion than do the other systems. This feature tends to reduce the data input burden on the user. (It has the disadvantage that it is less flexible in the event that the user should want features in the yards not explicitly provided for by the system designers. It is hard to imagine what such features would be, however.) Finally, the logic used to assign the meet, pass, and conform delays is somewhat better than in the other systems. That is, delays can be assigned to both trains rather than to just the lowest priority one, and criteria can be specified which control when a pass is to take place (rather than assuming that it will always occur).

One other system was studied that initially appeared to be applicable. This was the Railcar Network Model, (17) developed at Queen's University (Ontario) for use on the Canadian railroads. Upon further study, it was determined that this model operates at an even higher level of abstraction than the systems cited above and does not handle individual train movements. Thus, it was considered even less applicable to the terminal environment and will not be elaborated upon here.

4.2 Yard Simulation Systems

In addition to the above network simulation systems, the following yard simulations were analyzed:

1) Battelle Terminal II (18).
2) SCL/L&N Classification Yard Models (15,21).
3) SRI Yard Model (16,19).
4) New York Central Model (20).

It was determined that these models operate in essentially the same way as the nodes of a network simulation. In most cases, however, the amount of detail included in these strictly yard simulations is greater than in their network system counterparts. Even so, all of the caveats and assumptions that go along with the network simulation systems (the validation problem, interpretation of results, problem representation, levels of detail, data requirements, etc.) also apply to the yard systems.

5. APPLYING SIMULATION TO THE ST. LOUIS RESTRUCTURING PROJECT AND OTHER TERMINAL AREAS

This chapter identifies those areas where simulation could be useful in the current St. Louis restructuring effort, and specifies the general kinds of questions that are likely to be asked in each. This is followed by a discussion of the availability of the data that would be needed to support the simulation of the various areas identified. Finally, an assessment is made detailing how existing or possible future simulation systems should be used in the St. Louis terminal area study and in terminal area analysis in general.

5.1 Status of the St. Louis Restructuring Project

Most of the effort of the restructuring project has so far been devoted to collecting data about the physical track arrangements and about train and car movements throughout the terminal area. These data have already been used to calculate current corridor capacities and will be used throughout the project as the need arises.

The currently proposed restructuring plan consists of the abandonment or relocation of a large amount of industry trackage and yard facilities, the upgrading of three large yards and several major corridors, and the addition of a new corridor. An alternative plan which calls for the upgrading of only two yards has also been proposed. Currently these plans are being evaluated by the railroads, using the following guidelines that have been set down by the Technical Advisory Committee (23):

1) The key objective of the restructuring plan development is the reduction of car transit time through the terminal area.
2) Operations should be concentrated where possible to reduce the number of corridors.
3) Corridor crossing movements should be eliminated.
4) Corridor track conditions should be upgraded to improve operating speeds.
5) A by-pass should be developed to reduce congestion at Valley Junction.
6) Operations across the three bridges are to be evaluated to reduce delay and increase capacity.
7) Increased through movements will be considered where feasible.

Additional guidelines are being developed by the individual railroads and the proposed plans are being evaluated. The results of this evaluation will be an outline of the operating desires of the railroads from which train movements will be assigned to the yards and corridors. From these movements, estimates of corridor densities and transit times will be made. With this information in hand, changes will be made in the plan and the entire process iterated upon until agreement is reached that the plan is feasible and acceptable to all of the railroads concerned.

5.2 Types of Analyses in Which Simulation Can Help

This section suggests a paradigm in which the simulation methodology can be used to develop operating parameters of the yards, corridors, and junctions such that desired values of the measures of overall terminal efficiency (outlined in Section 2.4) can be attained. Then, when these operating parameters have been decided upon, simulation can help with the detailed design of the yard, junction, and corridor trackage. After that is done, simulation can then be used to verify that these detailed designs, when combined into the overall terminal, will in fact generate operations that achieve the originally defined goals.

5.2.1 Overall Terminal Evaluation

As mentioned in Section 5.1, the TAC has set the reduction of transit time through the terminal as the key objective of the restructuring plan. At the same time, the other criteria listed should also be met if the restructuring plan is to be acceptable. In order to accomplish these goals, it is necessary to identify those areas (yards, junctions, corridors, etc.) that have a negative impact on the transit time or other factors, and to make changes to eliminate or improve them. Because of the complex interactions that take place in a terminal area, it is difficult to assess how any one change will affect overall operations, transit time, or transit time reliability. When more than one change is being considered, the complexity of the evaluation increases even more.

All of this would be no problem if there were a limited number of possible changes and enough money available to implement all of them. However, the normal case is that there are a large number of changes that could be made with only enough money available to implement a few of them. (In St. Louis, for example, there are many routes that could be upgraded to major corridors, but only a small portion of the possibilities can actually be selected for improvement.) The question then reduces to finding the configurations of changes which have the best operational characteristics for the amount of money to be spent. As it may also be that the amount of money available is a function of the benefits to be derived, it is desirable to be able to develop a list of several configurations from which one can be chosen.

5.2.1.1 Developing Design Parameters

It is in the evaluation of these alternative configurations that simulation can be helpful. Such evaluation could take one of the two different forms mentioned in Section 3.2.1.3. That is: a configuration could be proposed where the parameters of its sub-units would be specified, but where the design details would be left until later; or, a configuration could be specified with all of its detail. The first type of simulation could be used during the initial planning phase for the evaluation of global alternatives. For instance, a typical question that might be asked of such a simulation is: "What would be the overall effect if a new corridor of a given capacity were provided between two specified points?" In this context, the simulation might be run several times, each time changing the capacity parameters and using the results to determine how the capacity of that corridor affects global transit time and other operating figures. Using this and other pieces of knowledge gained in a similar way, the overall specification of the terminal area would emerge.

5.2.1.2 Working Out The Details

The second type of simulation would be of use during the phase of the project when detailed track configurations are being worked out. For instance, after the parameters of a junction have been determined and after initial detailed design of that junction has been performed (see Section 5.2.2.2), it would be desirable to see how that design would influence the overall transit time, given current and projected traffic patterns. In this context, the kinds of changes that would be made from one simulation to the next would involve the arrangement of the actual junction trackage. It may even be that the junction would have different load characteristics if a route through it were altered at the same time that, say, a route out of a nearby yard was altered. Thus, the mode of operation of this type of simulation would be such that many changes of the kind illustrated above would be made and transit time figures as well as other measures would be evaluated for each of them. Such a process would allow a designer to experiment with different designs to see what kinds of results are produced from each. As good localized designs are found, the overall terminal design would gradually take shape in such a way that as each change is made, its overall effect on the terminal can be monitored.

Notice that this simulation would take place at a very detailed level, so that it would be possible to get results from it that were unavailable from the previous ones. For example, data could be generated on the lengths of times grade crossings are blocked, the amount of time trains spend stopped in high-vandalism areas, and the detailed costs of operation.

5.2.2 Local Design

5.2.2.1 Yards

The proposed reconfiguration plan calls for the expansion of two or three yards to provide major receiving and classification facilities for the terminal. Simulation can play an important role in their design by allowing just about every important detail to be specified and manipulated. Typical questions might relate to the number of classification tracks, humping strategies, hump speeds, location of various facilities, strategies for assigning switch engines to various tasks, yard operating strategies, etc.

Another area where simulation would be useful would be in the determination of how the proposed yard performs under saturated conditions. Thus, it would be possible to determine the degradation in performance that could be expected as receiving, classification, departure, and holding tracks fill to various levels.

5.2.2.2 Junctions and Corridor Segments

As part of the reconfiguration project, junctions will have changes made to them and different levels of traffic will be routed through them. Just as with yards, it is desirable to be able to try out different specific configurations and to develop various measures of the limits of the junction's operation. Also, since the performance of a given junction is a function of the performance of nearby junctions and corridors, it is desirable to include these in the evaluation. Bridges are a good example of where this type of analysis is essential. Here, it is not really the traffic over the bridge that is of direct interest, but rather how the various routings to and from the bridge combine to produce the levels of traffic and resultant congestion around its approaches.

Simulation can help in this type of design work by providing a method whereby detailed changes in track configuration can be made and the resulting effect on a given traffic pattern can be observed. Alternatively, the track configuration can be fixed and the traffic volume and routing can be manipulated to determine such factors as the saturation points of the various routes and delay times as a function of volume.

5.3 Availability of Data for Simulations

As mentioned in Section 3.2.2.2, railroad simulations require a large amount of data that is often expensive to collect. Already, though, a significant amount of data has been collected relating to the St. Louis terminal area as a result of both the previous and current restructuring study efforts. The next two sections summarize what these data are and analyze their usefulness to the kinds of simulations outlined in Section 5.2 above. The last two sections discuss current plans for collecting more data and the possible need for data even beyond that.

5.3.1 Parsons Brinckerhoff Data

During the original Parsons Brinckerhoff study of the St. Louis area, detailed information was gathered for a 15-day period and used to drive the AAR Simulation System. Unfortunately, conditions have changed sufficiently so that those data no longer accurately reflect terminal operations. For example, some yards and corridors no longer exist, and the advent of CONRAIL has significantly changed the operations over the trackage that remains. Thus, while useful for the earlier study, these data are not useful for any current or future simulation of the St. Louis area.

5.3.2 CONSAD Data Already Collected

As a part of the current St. Louis terminal reconfiguration project, CONSAD Research Corporation has collected a large amount of information mostly relating to physical track arrangements and train and car movements throughout the terminal area. These data include:

a) An overall network map.

b) Corridor maps with slightly more detail than the network map.

c) Very detailed junction maps.

d) Train movements - Including routes traveled, average number of cars per train, running frequency, schedule, etc.

e) Corridor densities - Aggregation of the train movements overlayed onto the network map.

f) Representative amount of interchange traffic by carrier, detailed as traffic that goes directly from one railroad to another, and traffic that is handled by the two terminal railroads in the area.

g) Bridge delay statistics.

Most of these data would be useful in the four types of simulations discussed in Section 5.2 above. Some would be directly usable as input to the simulation, and some would be useful in the validation phase. These uses are summarized in Table 5-1.

	OVERALL PARAMETER DETERMINATION	DETAILED YARD DESIGN	DETAILED CORRIDOR & JUNCTION DESIGN	DETAILED OVERALL TERMINAL DESIGN
Network Map	I		I	I
Corridor Maps			I	I
Junction Maps			I	I
Train Movements	I	I	I	I
Corridor Densities	V		V	V
Interchange Traffic	I	I	I	I
Bridge Delays	I		V	V

I = useful as input to the simulation.
V = useful during the validation phase.

Table 5-1 - USEFULNESS OF EXISTING DATA

5.3.3 CONSAD Data Yet to be Collected

Most of the CONSAD data collection effort is over, but there are two areas where more work is planned. The first of these involves gathering estimates from the individual railroads concerning the growth of interchange traffic. Such estimates are essential for evaluating the alternative configurations since actual implementation of any reconfiguration plan is likely to be five to ten years in the future.

The second area under consideration for more data collection involves the refinement of the car information associated with train movements. So far, only the average number of cars per train is known. The refinement would include the minimum number, maximum number, distribution, and variation of the number of cars which make up each train. Such information is needed for any of the four types of simulation under consideration.

5.3.4 Additional Data Requirements

Whereas all of the data mentioned above is useful in the various simulations, much more is needed before an actual simulation could be run. Table 5-2 summarizes some important kinds of information needed for each of the simulation types.

The smallest amount of additional data is needed for the yard simulation. The overall parameter determination and detailed corridor and junction design simulations require about the same amount of additional data, but the nature of that data is somewhat different for the two. The most extensive additional data requirement is for the detailed overall terminal design simulation, even though Table 5-2 shows this to be about the same as that required for the detailed corridor and junction design. The reason for this is that the latter simulation is meant to answer a more restricted set of questions than is the former, thus requiring a smaller amount of data. In all cases, other than the yard simulation, extensive additional data must be collected before a simulation can be run.

5.4 Assessment of Existing Simulation Systems and Recommendations for New Systems

The four potential simulation applications defined in Section 5.2 are discussed below, along with an assessment of how existing or possible future simulation systems meet the requirements of the particular application.

	OVERALL PARAMETER DETERMINATION	DETAILED YARD DESIGN	DETAILED CORRIDOR & JUNCTION DESICN	DETAILED OVERALL TERMINAL DESIGN
Routing statistics of individual cars	X		X	X
How trains handle classes of cars	X		X	X
Train connection data	X		X	X
More detailed yard facility data	X	X	X	X
Yard and junction delay times	X			
Entry time of trains into terminal area	X		X	X
Degree of variability of all of the above	X		X	X
Corridor meet, pass, and conform times	X			
Slightly more detailed network map			X	X
Detailed operating rules & strategies			X	X
Detailed yard diagrams		X		X
Detailed yard operating strategies		X	X	X

Table 5-2 - ADDITIONAL DATA REQUIREMENTS

5.4.1 Developing Overall Terminal Design Parameters

Currently the TAC is weighing alternative plans and developing overall design parameters for the terminal. It is in this area where it is most tempting to use one of the network simulation systems described in Section 4.1. Unfortunately, none of these network simulation systems is appropriate to the kind of problems that are presented in the St. Louis restructuring project. This is primarily because the simulation systems represent most of the information that is important for terminal operations only indirectly and at a level of abstraction that is too far from reality to be of use in terminal planning. For example, junctions and passing trucks are of critical importance in terminal operations. If they become clogged or are improperly placed, terminal operations can be considerably impaired. The current network simulation systems are unable to represent these situations realistically and, as a result, are unable to detect problem areas. For example, in actual operations, delays do not always accrue to the lowest priority trains, because passing sidings are not available at any given point along the railroad.

Also, much of the information required by the systems as input is exactly the information that is sought. The requirement for meet, pass, and conform delay times is one example of this. For instance, when planning a new track configuration, these times are unknown. If the new configuration has no system-wide effects, then it might be possible to predict the delay times. But it is precisely because such changes do have system-wide effects that a simulation is desired in the first place.

The above comments should not be construed as a condemnation of the various network simulation systems, since none of them has ever claimed to be applicable to terminal areas. In fact, quite the opposite is true, as is illustrated by the following quotation from the introduction to the AAR Network Simulation System users guide (Ref. 6 - Pg. 8):

> "In developing the network model, it was considered that the upper limits of detail should be something less than that of a terminal model... If a user felt that such a high level of detail was necessary, he could use the network model first to get a feel for interaction effects and then use a component model for a more detailed simulation."

Unfortunately, none of the network models is likely to provide this "feel for interaction effects" because of their lack of faithfulness in reproducing most of the significant features of the terminal area environment. Further, no other systems seem to exist which can reproduce these features. Thus, the decision by the TAC not to use an existing simulation system during the Phase I study (where the overall terminal design parameters are being specified) was a wise one.

5.4.2 Detailed Design of Yards

As mentioned in Section 4.2, the various yard simulations suffer from many of the same problems as the network simulations. There is one very important difference between the two however, and that difference involves the scope of the problem being attacked. In the case of the entire terminal area, it is impossible for one person to have a clear understanding of its operating subtleties, whereas a yard is constrained enough so that one person can easily understand most of its operations. Therefore, it is less likely that false assumptions will be made with the yard simulations about delay and job processing times or about other parameters that influence the simulation. For example, there is usually only a fixed number of engines working in a yard at one time, so that the major variables are the rate of arrival and departure of trains. Contrast this to the terminal area as a whole where a large and unpredictable number of trains operate at any one time and numerous routing possibilities exist.

Since the magnitude of the yard simulation problem is smaller, less human and computer time is required for each run of the simulation. Therefore, more alternatives are likely to be tried and any anomalous cases are likely to show up as unexplained deviations of some variable. Thus, simply because of the lesser magnitude of the problem, it is likely that yard simulations will embody more reasonable representations of the real world than will terminal area simulations, even though they both rely on the same simulation techniques (i.e., that of implying the real-world structure through parameter values, rather than by explicit representation of that structure).

All of the yard simulations studied appear to be viable candidates and are likely to produce the benefits discussed in Section 5.2.2.1 above. Since the amount of data required for yard simulations is not great, and since their running time is relatively low, they should be seriously considered for use whenever the details of yard redesign are being studied.

5.4.3 Detailed Design of Junctions and Corridor Segments

The primary problem with the network simulation systems is that they are unable to represent trackage at all, and it is the details of the trackage that often constrain terminal operations. The framework suggested above in the quotation from the AARSS user's guide is appropriate, except for one major flaw. That is, that no "component model for a more detailed simulation" exists that can be used after the user has gotten a "feel for the interaction effects." This lack of a detailed component simulation is curious since almost every model reviewed makes the same point as the AARSS user's guide. Clearly it is not the lack of a need for such a system that has prevented it from existing. Instead, the reason seems imbedded in the history of railroad simulation which began with the Allman model (22) and has continued virtually unchanged until the present time. A basic tenet of the Allman approach is that much can be gained by applying some fairly simple analytic abstractions to the railroad simulation situation. Indeed, this is the case, and for simulating railroad networks at, say, the inter-city level, the existing network simulation systems provide a good basis. But, when this analytic approach is pushed too far (as into terminal areas, for instance), it fails because the abstractions are too remote from reality. Further, the simulation methodology has gotten a reputation for being unable to handle great detail, probably because such detailed simulations are viewed in the context of the current analytic techniques, a situation which is clearly hopeless. But it is just this type of detailed simulation that is needed in the St. Louis terminal area, especially for the evaluation of corridors and junctions.(*)

Thus, in order to provide good terminal area simulations, it is necessary to adopt a new approach to handling the kind of detail required by them. Specifically, it must be possible to represent the trackage of the terminal in detail -- right down to the last crossover, if necessary. Further, it must be possible to simulate trains running over that trackage so that it is known what sections of track are occupied and when. Only in this way are junction and passing siding conflicts going to be detectable and correctable. Finally, it must be possible to "run" the railroad in the computer in a manner

-----

*This opinion has also been expressed by Mr. Chandler Lewis of CONSAD Research during personal discussions.

similar to actual operations. The benefit of this is that the simulation will finally be able to provide the delay times, conflicts, etc. rather than requiring the user to supply them as input, and that is, after all, what is really desired of a simulation.

In addition to the characteristics mentioned above, the simulation system must be easy to change and run, and its output must be easy to analyze. All of the existing systems suffer to some extent from the problem that extensive changes must be made to the input data for what on the surface seems to be a simple configuration change. Further, it is an often repeated complaint that the output reports provide so much detail that it is difficult to determine the real effect of changes. And finally, if the simulation is really to be an experimental tool, it must run fast enough and easily enough so that many runs can be conveniently made.

Such a simulation system is within the current state of the art of computer technology and, given its apparent potential benefits, efforts should be directed toward its development. Such an advanced system will probably cost from three to five-hundred thousand dollars to develop, but, when complete, should pay for itself through ease of running, through ease of modification, and most importantly, through the data that it will provide.

5.4.4 Overall Terminal Evaluation Considering All Details

The system just described can be viewed as the combination of a good yard simulation (such as one of those mentioned in section 4.2), and a new simulation which can handle details of the corridor and junction trackage. The major difference between that system and one which would consider all terminal details is that, in the latter, the yard and industrial trackage would be explicitly simulated as would the trains operating over that trackage. Since no experience yet exists with simulations of even the simpler variety, it is not recommended to develop such a detailed simulation at this time. However, such a system would be a natural extension of the one outlined in Section 5.4.3. and that system should be built with such growth in mind.

6. REFERENCES

1. East-West Gateway Coordinating Council, Comprehensive Areawide Railroad Consolidation and Relocation Study, St. Louis Region; Executive Summary and five volumes; Parsons, Brinckerhoff, Grotz and Eric Hill; St. Louis, Mo; June 1974.

2. Hoover, T., & Minger, W.K., "Computer Simulation of a High-Volume Rail Gateway," Proceedings - Sixteenth Annual Meeting Transportation Research Forum, XVI, (1), 1975, 139-147.

3. Landow, H.T., "The Cost and Service Effects of Alternative Terminal Reorganization Plans," Proceedings - Sixteenth Annual Meeting Transportation Research Forum XVI, (1), 1975, 133-138.

4. FRA Memorandum and Statement of Work, St. Louis Terminal Restructuring Project, April 29, 1976.

5. CONSAD Research Corporation, Proposal to Accomplish an Analysis of the St. Louis Terminal Restructuring Plan, a technical proposal prepared for the US DOT, FRA, June 7, 1976.

6. Midwest Research Institute, The AAR Network Simulation System - A Tool for the Analysis of Railroad Network Operations, February, 1971.

7. Minger, W.K., "The AAR Network Simulation Model," Railway Management Review, 72, (2), A101-A119.

8. Drucker, R.W., et al. "A Mini-network Computer Simulation Model for Railroad Planning," Rail International, Nov.-Dec., 1973.

9. C&O/B&O Railroad. "Chicago-Buffalo Service Study - Computer Simulation," Internal Project Paper, October, 1971.

10. Yoakum, R.L., & Beaumont, L.H., "A Railroad Car Scheduling System Incorporating Simulation," Missouri Pacific Railroad, St. Louis, Internal Project Paper, Undated.

11. Bellman, J.A., "Railroad Network Model," Proceedings - Second International Symposium on the Use of Cybernetics on the Railways, Montreal, October, 1967.

12. Coates, P.J., & Clark, P.W., "A Simulation Method Applicable to Railway Operation," Proceedings - Second International Symposium on the Use of Cybernetics on the Railways, Montreal, October, 1967.

13. Canadian National Railways, The CN Network Model-User's Guide, Operational Research Branch, R&D Dept., November, 1971.

14. Louisville & Nashville Railroad Company, Progress Report - Network Simulation Model of Traffic Flow, October 3, 1970.

15, IBM. Simulation on the Family Lines System, New York, 1974.

16. Siddiqee, W., et al., Blocking and Train Operations Planning, Final Report - SRI Project 3759, October, 1975.

17. Petersen, E.R., & Fullerton, H.V., The Railcar Network Model, Canadian Institute of Guided Ground Transport, Queens University at Kingston, Ontario, Canada, June, 1975.

18. Nippert, D.A., "Simulation of Terminal Operations," Simulation of Railroad Operations, October, 1966.

19. Tuan, P.L., & Procter, H.S., A Railroad Classification Yard Simulation Model, SRI Internal Project Paper, Undated.

20. Nadel, R.H., & Rovner, E.M., "The Use of a Computer Simulation Model for Classification Yard Design," Proceedings - Second International Symposium on the Use of Cybernetics on the Railways, Montreal, October, 1967.

21. King, L.D., Simulation of a Railroad Classification Yard, L&N Railroad Company, Internal Project Report, Undated.

22. Allman, W.P., "A Network Model of Freight System Operations," Simulation of Railroad Operations, October, 1966.

23. St. Louis Gateway Terminal Project Technical Advisory Meeting Minutes, December 1, 1976, p.2.

125 Copies