MASS--A MAIL SERVICE SIMULATION

Paul L. Tuan and David S. Nee*
Stanford Research Institute
Menlo Park, California

Abstract

This model was developed for the U. S. Post Office Department for the purpose of evaluating the relative merits of alternative mail processing, handling, and transportation plans. It may be used to evaluate proposed mail sorting and routing schemes, manpower allocations, new mail processing equipment, transportation scheduling, and mail volume fluctuations.

1. INTRODUCTION

The purpose of this simulation program is to provide the U. S. Post Office Department with a computer model for evaluating the relative merits of alternative nonpriority mail** processing, handling, and transportation plans in terms of service time*** and utilization of resources.

The nonpriority mail constitute approximately 80% of the total mail volume by cubic feet. Congestion situations created by these mail in postal facilities or transportation terminals can seriously impair the effective delivery of mail, and under certain conditions, it can cause total stoppage of mail flow in a particular area. Therefore, it is important that the Post Office Department must possess and utilize appropriate analytical tools with which to assess feasible alternative sorting and routing plans, containerization programs, transportation scheduling, and material handling procedures, etc., under varied mail volumes and compositions. MASS is one of the tools designed to fulfill this purpose.

Simulation techniques are employed on only those problems which are not amenable to analytical queuing solutions. These problems usually have the following features:

(1) They involve large postal networks which contain numerous service points and transportation links.

(2) The arrivals of mail are arbitrarily distributed.

(3) Transportation schedules and mail dispatch times are arbitrary discrete functions.

* Now with Honeywell, Inc., EDP Division.

** Nonpriority mail includes second class nontime-sensitive, third class and fourth class mail. Sometimes they are also called "nonpreferential" or "bulk" mail.

*** Service time is defined as the difference between the time when the mail enters the originating facility and the time when the mail reaches its destination.

Because of the compatibility of network characteristics between GPSS and postal systems, GPSS/360 was selected as the programming language. This research received its support from the U.S. Post Office, Systems Engineering Branch of the Bureau of Research and Engineering. Acknowledgment is made to Mr. Paul Whiting (now with the COMSIS Corporation of Santa Clara, California) and Mrs. Diane Jennings for their contribution in computer programming, and to Professor Robert Oakford of Stanford University for his advice on the development of the postal model.

2. CHARACTERISTICS OF THE MODEL

The main characteristics of MASS relate to a "job-resource" relationship, where jobs (i.e., groups of mail) enter a postal network at certain time intervals, and require the application of units of resources in a prescribed sequence. A postal network is normally composed of elements such as mail processing facilities, transportation terminals, transportation links, processing stations, and materials-handling links. It also includes source nodes and sink nodes where mail enters and leaves the network. The resources consist of mail processing and handling crews, transportation vehicles, mail processing and handling equipment, and so forth.

The simulation process requires the generation of batches* of mail at each of the source nodes and the routing of these mail batches through the network, from node to node, in a prescribed sequence. Each batch of mail will eventually complete its journey through the network and will leave the network via its designated sink node. With the exception of source and sink nodes, a node in MASS can be any of the following:

(1) A post office

(2) A processing station (e.g., primary sort, secondary sort, canceling, culling, loading, unloading)

(3) A container loading and dispatch point

(4) A motor vehicle (including Flexivan types) loading and dispatch point (with a departure schedule)

(5) A gateway

(6) A train depot (or terminal)

(7) A transportation link

(8) A materials handling link

(9) A combination of any of the above items.

The details of a network simulation can vary from an intra-postal facility level (individual processing stations within a postal facility) to an inter-postal facility level, or include both levels in the same simulation.

For each category of mail to be simulated, a routing list must be given. Figure 1 shows a simplified postal network in which three source nodes (S_1, S_2, and S_3) and four sink nodes (T_1, T_2, T_3, T_4) are displayed. If there were six categories of mail to be simulated (denote them M_1, M_2, M_3, M_4, M_5, and M_6), then the routing lists might look as shown at the top of Figure 1.

The network nodes constitute the permanent components of the model. In the MASS program each node is a type of MACRO. The present version of MASS provides 15 different types of MACROs. Each MACRO has a number of "dummy arguments" that need the assignment of proper values before a simulation.

Mail batches are temporary components. The attributes of each batch are defined by 11 parameters (e.g., origin point, mail category number, container type assigned). In the simulation, each batch is routed through a prescribed set of nodes in accordance with a "transfer function." For example, observe from Figure 1 that each batch under category M_4 has five ordered nodes on its path--S_2, X_7, X_5, X_6, T_3. Thus, in the transfer function for M_4, five consecutive integers (step numbers) beginning with 1 are given as the arguments of the transfer function. The "values" of the function are the names of the five nodes (MACROS) corresponding to these steps. Each time a batch of mail completes its activity at a node, the simulation program will automatically increment the step integer value by one, identifying the next MACRO to which the simulation activity will be transferred. Each mail category has a unique transfer function, and each batch of mail carries a transfer function number as one of its 11 parameters.

* A "batch" represents a group of mail, with a predetermined size (e.g., 200 parcels, 1000 letters), as a single indivisible unit throughout the simulation.

3. MEASURES OF EFFECTIVENESS AND APPLICATION AREAS

The measures of effectiveness of a postal plan are:

(1) The total transit time of mail between given pairs of nodes

(2) The queue statistics associated with each mail-processing, handling, and transportation facility

(3) The utilization statistics associated with each processing, handling, and transportation facility.

MASS may be used to evaluate proposed sorting and routing schemes, manpower allocations, new mail-processing equipment, transportation scheduling, mail volume changes, etc. The postal network can be based on either an interfacility flow where each facility may represent an aggregated complex of smaller facilities, or an intrafacility flow where each individual processing station within a building can be simulated. It is also possible to mix both facility levels in the same simulation.

4. A CASE STUDY WITH MASS

4.1 INTRODUCTION

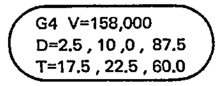
MASS has been applied to simulate a New York/Chicago postal network. One of the studies is the evaluation of a new parcel post proposal in which the fourth class parcel post from Chicago Main Post Office to Manhattan, New York, would be moved in large containers instead of the traditional mail sacks. It was conjectured that a containerized operation would bypass many intermediate mail handlings, hence, it may reduce total transit time, queue size at critical points, and possibly cost. The mail flows under the two alternative systems (containerized system vs sack system) are shown in Figure 2.

4.2 BLOCK SYMBOLS AND NETWORK IDENTITY

The block symbols and data displayed therein as shown in Figure 2 are explained below:

4.2.1 Mail Volume Generation Block

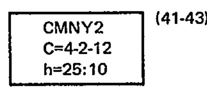
This block contains the location number of the generation mode. Associated with each generation block, but not necessarily inside the block, one may specify the mean daily volume (V), the percentages of the mail of each category (D), and the percentages of the daily mail for each of the three tours (T). This block symbol is:



Thus, the above example indicates that at Generation Node 4 the mean daily volume is 158,000 units of mail, and 2.5%, 10%, 0%, and 87.5% of the mail are of categories 1, 2, 3, and 4, respectively. Some 17.5% of the mail arrives during the first tour (i.e., an 8-hour shift, from midnight to 8 a.m.), and 22.5% and 60.0% arrive during the second and third tours, respectively.

4.2.2 Mail Service Block

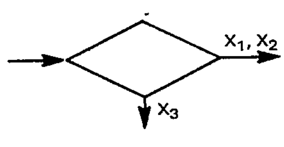
This block contains the mnemonic name of the service facility, the service capacity per tour (C), and the mean and standard deviation of the service time (h).



In this example, CMNY2 denotes the secondary New York sort work station at the Chicago MPO. The number of mail handlers assigned to this work station for each of the three tours is 4, 2, and 12, respectively. The mean service time per transaction by each of the mail handlers is 25 simulation time units, in this case minutes, with a standard deviation of 10% of the mean of the normal distribution. The numbers in parentheses indicate the work station numbers, one for each tour.

4.2.3 Mail Routing Block

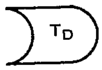
This block is used to indicate the routing of the mail and it is represented by:



In this particular example, mail of types X_1 and X_2 is routed to a different destination from that of X_3.

4.2.4 Departure Schedule Block

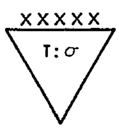
This block is used to represent the departure schedule of the postal vehicles. The symbol used is:



where T_D is the departure time expressed in terms of the 24-hour clock.

4.2.5 Time Delay Block

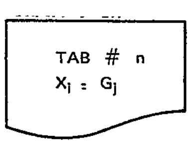
This block is used to represent the transit time of mail between two points in the postal network. For example, it could represent the travel time of a parcel on a conveyor belt between two work stations or it could represent the train travel time between two cities. The symbol for this block is:



where T is the mean transit time, σ is the standard deviation of percent of the mean, and xxxxx is the mnemonic name of the block.

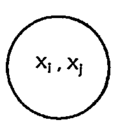
4.2.6 Tabulate Block

This block is not used to describe the flow of the mail, but to indicate the tabulation, at any desired point, of the transit time statistics for specified categories of mail, X_i, that originated at a specified generation node, G_j. The symbol used is:



4.2.7 Sink Block

This block is used to indicate the departure of certain categories of mail from the simulated postal network. Thus,



indicates the departure of the X_i and X_j categories of mail.

Each of the work stations and mail-transit blocks is labeled with a five-character mnemonic name. The first two characters identify the postal facility (e.g., CM implies Chicago Main Post Office), and the last three characters signify the operation of the block.

4.3 SYSTEM PARAMETERS

4.3.1 Mail Volume Generation

Mail enters the network at the generation blocks only. At each generation block, one or more of the categories of mail listed in the following tabulation is generated:

Category	Mail Destination
A	Manhattan
B	New York Connections
C	Others--New York Truck Terminal
D	Others

Mail volume generation is based on a normal distribution function with a given mean and 20% relative standard deviation. Saturday mail volume is assumed to be one-third of the weekday daily volume. No mail is generated on Sunday. The mean daily volume (number of batches), the percentage distribution of the mail for each of the three tours, and the percentage distribution for each category of mail are listed in Tables 1(a) and 1(b).

Table 1(a)

MAIL VOLUME GENERATED BY TOUR

Generation Block No.	Daily Mean Volume (No. of Batches*)	Tour (%) 1	2	3
1	135	23	11	66
2	118	50	30	20
3	270	5	54	41
4	788	5	65	30
5	520	20	40	40
6	775	20	40	40

* 1 batch = 200 parcels. Thus on the average, approximately 135 x 0.23 x 0.25 = 7.8 batches of Category A mail will be generated at G_1 of the eastbound network during the first tour.

Table 1(b)

MAIL VOLUME GENERATED BY MAIL CATEGORY

Generation Block No.	Daily Mean Volume (No. of Batches*)	Mail Category (%) A	B	C	D
1	135	25	10	30	35
2	118	25	10	30	35
3	270	3	1	3	93
4	788	58	17	0	25
5	520	100	0	0	0
6	775	100	0	0	0

* 1 batch = 200 parcels. Thus, on the average, approximately 135 x 0.23 x 0.25 = 7.8 batches of Category A mail will be generated at G_1 of the eastbound network during the first tour.

4.3.2 Service and Service Discipline

The service time at each of the work stations is assumed to be normally distributed. The service discipline is FIFO throughout the networks.

4.3.3 Transportation

The transportation between the two gateways (Chicago and New York City) is by train only. The travel time of the train is approximately 24 hours.

4.4 CONDITIONS OF THE SIMULATION EXPERIMENTS

The simulation of each of the two postal systems is for 8 days of operations, from Monday through Monday. At the end of each tour, the system statistics are printed out. With the 8-day simulation duration, it is possible to observe the weekly variation in the performance of each work station.

Three separate simulation runs were made for the Containerization System. These runs were made to investigate the sensitivity of the transit times of the containerized mail under varying combinations of container loading crew sizes at the container loading station at the Chicago MPO, CMCL1, and dispatch times of the containers. The three simulation runs are called Run A, Run B, and Run C of the Containerized System, and the conditions for the three runs are as follows:

Condition	A	B	C
Crew Size	1	2	2
Dispatch Time*	1000	1000	0400

4.5 RESULTS AND DISCUSSIONS**

4.5.1 Transit Times

Transit times statistics presented in this section were gathered at the end of the 8-day simulation runs. Statistical analyses were performed with the t and F tests. They both showed satisfactory results.

The mean and standard deviation of transit times of the "Manhattan parcel post" together with the sample size for the sack system and Runs A through C of the containerized system are listed in Table 2. The transit time is between the New York Secondary Sort Station at Chicago MPO, CMNY2, and the City Separation at New York Truck Terminal, NTCTY, in the case of the sack system, and the City Distribution Station at NY FDR Station in the case of the containerized system.

Table 2

TRANSIT TIME STATISTICS OF EASTBOUND SYSTEMS

System	Transit Times (hours) Mean	Standard Deviation	Sample Size (No. of batches)
Sack	56.6	4.3	172
Container-Run A	62.6	11.8	138
Container-Run B	58.8	10.3	174
Container-Run C	52.5	5.3	186

The mean transit time of the sack system is somewhere between the times for Run B and Run C of the containerized system. Also, the standard deviation in time of the sack system is somewhat comparable to that of Run C of the containerized system, but is much smaller than that of Runs A and B of the containerized system. Figure 3 shows the histograms of the transit times of the "Manhattan parcel post" for the sack system and Run C of the containerized system. The transit times of the sack system are much more normally distributed than those of the containerized system. In fact, the transit times of the containerized system

* Time is expressed in 24-hour clock.

** For the purpose of this paper, only the results of transit times are presented in sufficient detail.

seem to have two modes, one at about 51 hours and the other at 63 hours. Figure 4 shows the transit-time histograms of Runs A, B, and C of the containerized system. It is seen that the spread of the transit times increases from Run C to Run A. The increase is most obvious between Runs B and C. The only difference between Runs B and C is the container dispatch time; 1000 in the case of Run B and 0400 in the case of Run C. Changing the dispatch time from 1000 to 0400 allows a larger proportion of dispatched containers to make the train departure schedule.

Figure 5 shows the transit-time histograms, in terms of percent, of the New York Connection parcel post (Category B mail) for Runs A, B, and C of the containerized system. The processing of Category B mail is not influenced by the changes in crew size and dispatch time schedule of the Category A mail. Thus, the variation in the mean transit times of the Category B mail is probably caused by the random variation of the simulation experiments.

4.5.2 Work Station Statistics

The utilization factors of the work stations at the Chicago MPO and Chicago Truck Terminal of the sack system are in general higher than those of the containerized system. This is understandable since the work stations in the sack system process all parcel post while the work stations in the containerized system process only the noncontainerized parcel post. An exception is the first tour utilization factor of FDCTY in the containerized system, which is significantly higher than that of the sack system. This is because of the fact that FDCTY is normally processing mail at the saturation level of about 94% during the third tour. When the containerized mail from Chicago MPO arrives at FDCTY during the third tour, FDCTY is simply saturated. The overflow mail is worked off during the first tour of the next day.

An examination of the delay-time and queue-length statistics indicated that the following work stations have relatively long delay times and large queues: FDUN1, NTPRS, and FDCTY in the containerized system, and NTUN1 in the sack system.

Figures 6(a) and 6(b) compare the utilization factors (by tour) of two work stations with similar missions under two separately simulated systems--NTUN1 (New York Truck Terminal Flexivan Unloading) under the sack system and FDUN1 (FDR Station Flexivan Unloading--Containers) under the containerized system. Both work stations have the function of unloading from Flexivans the parcel post originated from Chicago. However, NTUN1 unloads sacks directly from Flexivans, whereas FDUN1 unloads only Flexivans with containers (approximately 400 parcel post per container). The volume arriving at NTUN1 is higher than that of FDUN1 since the latter work station handles only containerized parcel post from New York Secondary Sort in the Chicago MPO, whereas the former work station handles New York bound parcel post from all generating points in the Chicago MPO.

It takes more than a day for the first group of mail to reach New York from Chicago (the simulations begin from a "cold-start" condition--i.e., no mail in the system at time zero), hence there is virtually no workload during the first two days of the simulation. Also, the effect of the weekend reduced mail volume (one-third of normal volume on Saturday, and no mail generated on Sunday) does not reach these two New York work stations until the second tour of the eighth day (Monday).

In the simulation, the Chicago to New York train arrives in New York at 1450 daily. On the average, it takes three hours for the Flexivan to travel from the depot to unloading platforms. It is therefore reasonable that Figures 6(a) and 6(b) show that the utilization at NTUN1 is on the increase during Tour 3 (1600-2400) and reaches a peak during the following tour (Tour 1). On the other hand, the unloading of Flexivans at FDUN1 (container unloading) is relatively faster and it has less mail to unload than NTUN1; hence, at FDUN1, Tour 3 is the only tour during which work occurs and is also completed.

4.5.3 Block Statistics

At the end of each simulation interval, "block statistics" are printed out. These statistics are useful in studying the performance of the system under investigation. For example, the block statistics under the containerized system reveal that on the second tour of the first day, two transactions were ready for ramping and that they were ramped and departed with the train. On the second tour of the second day, 30 transactions arrived at the ramping station; only 18 of the 30 transactions succeeded in making the train. On the third day, 20 transactions arrived at the ramping station, and the train carried 22 transactions. Since there were 12 transactions left over from the previous day, only 10 of the current 20 transactions made the train. By the same process of deduction it can be shown that after the first six days, the transactions that arrived at the ramping station during the seventh and eighth days did not arrive at TAB5 at the end of the simulation; 54 transactions out of 138 did not make the train. Thus, the transit time for these 54 transactions on the average would be about one day (1440 minutes) longer than that of the other 84 transactions. Hence, there are two modes, which are approximately 1440 minutes apart.

The 12 transactions (batches) that missed the train on the second tour of the second day arrived at CMCL1 no later than the first tour of the second day. If it had not been for the container departure schedule, 1000, they could have been ramped before the train departure schedule and made the train. The same is true for all 54 delayed transactions. Consequently, had it not been for the container departure schedule, the distribution of the transit times of the containerized mail would have been single mode with a mean transit time in the neighborhood of 50 hours.

This analysis of the cause of bimodal distribution of the transit times of containerized mail is a typical example of the use of the block statistics to study the performance of the system.

5. DATA COLLECTION AND MODEL VALIDATION

There are several well established data reporting systems in the Post Office Department. To name a few, the Cost Ascertainment reporting system, the Work Measurement System, the originating mail volume reporting system, and data supplied by business mailers, etc. In addition, the simulation team had performed site visits to pertinent postal facilities to construct mail flow diagrams, and to collect additional information which was not included in the established reporting systems.

To validate the MASS model, a sample of 500 parcels between various points in Chicago and New York City was used. The days and times of the mailings were designed to permit reasonable estimates of transit times between points of interest. The data derived from this field test were then used to test the reasonableness of simulation runs covering the same network. The test period covered is from February 12 to 25, 1968. The parcel post volumes at the originating postal facilities during the test period are considered normal and no unusual conditions existed which would cause undue delay to parcel post mailings. The testing data collected reflect the sorting and dispatching schemes of the test period only, however some inferences can be made to other time periods.

Parcels were selected at random during distribution at the originating postal facilities (Chicago Main Post Office and Chicago South Suburban Facility). POD forms 3754 (Parcel Post Test Record) were attached to the parcels. The forms were detached from the parcels with delivery information furnished by clerks during distribution in the incoming parcel post unit, by parcel post carriers, and in a few cases by postal patrons when delivery was effected. In essence, the information obtained from each sample includes the following:

(1) Time and date recorded at the originating postal facility

(2) Time and date recorded at the destination postal facility

(3) Actual or expected date of delivery to the patron

(4) Identification of facilities

(5) State of sortation of the parcel at the time of parcel selection.

From the above raw data it was possible to develop parcel post transit time statistics for eight pairs of major facilities between New York City and Chicago. These statistics were used to confirm and calibrate simulation model structure and input parameters. However, it has been noted that the transit times derived from simulations are invariably smaller than those deduced from the field test. This was due to the fact that certain events were not included in the simulation. For example, the "re-cycling" of parcel post in a facility from mishandling was purposely left out of the model due to lack of data. For similar reasons the idling of mail storage cars in rail yards was also omitted. Thus, we obtained results which are idealized in the sense that any delay due to recycle or rail yard idling will only degrade the performance measure predicted. In these cases, we recognize that the simulated results are upper bounds on performance.

6. CONCLUSION

MASS permits detailed exploration of alternative proposals for processing, handling, and transportation of mail. It is relatively inexpensive. As an example of economy of operation, simulation of a network with 40 service points, covering eight days of operations with an average daily volume of 520,000 pieces of parcel post, required approximately 10 minutes on an IBM 360/65.

Recommendations for future work include the incorporation of a dynamic manpower scheduling scheme, the computation of direct operating costs, and the simulation of equipment breakdowns.

REFERENCES

1. "Cost Ascertainment Studies and Tests at First- and Second-Class Post Offices," Series F-7 (Rev.), Post Office Department, Washington, D.C.
2. "General Purpose Simulation System/360 Users Manual," H20-0326-2, International Business Machines Corporation, 1968.
3. David S. Nee and Paul L. Tuan, "A Mail Service Simulation (MASS)," SRI Research Memorandum ORD-RM 6748-5, January 1969.
4. Paul L. Tuan, "A Postal Network Simulation," SRI Research Memorandum ORD-RM 6748-1, July 1968.
5. Paul L. Tuan, "Parcel Post Time Test and Evaluation," Appendix C of SRI Interim Technical Report 6748-1, August 1968.
6. "Work Measurement System," Series M-20, M-33, Post Office Department, Washington, D.C.

BIOGRAPHIES

Paul L. Tuan is a Senior Operations Analyst with the Stanford Research Institute. Previously he was the Director of Computer Center, University of Utah, and the Detroit District Technical Manager of the Burroughs Corporation. He has a B.S. degree from the University of Shanghai; M.B.A. from the University of Denver; M.S. and ENGINEER from Stanford University.

David S. Nee has been an Operations Analyst with the Stanford Research Institute. Prior to that he was a system manager with the Marchant Research, Inc. He has a B.S. degree (physics) from St. John's University, Shanghai; M.A. (mathematics) from Bowling Green State University; B.S. (E.E.) from the University of California at Berkeley.

M_1: S_1, X_1, X_2, X_3, X_4, T_1

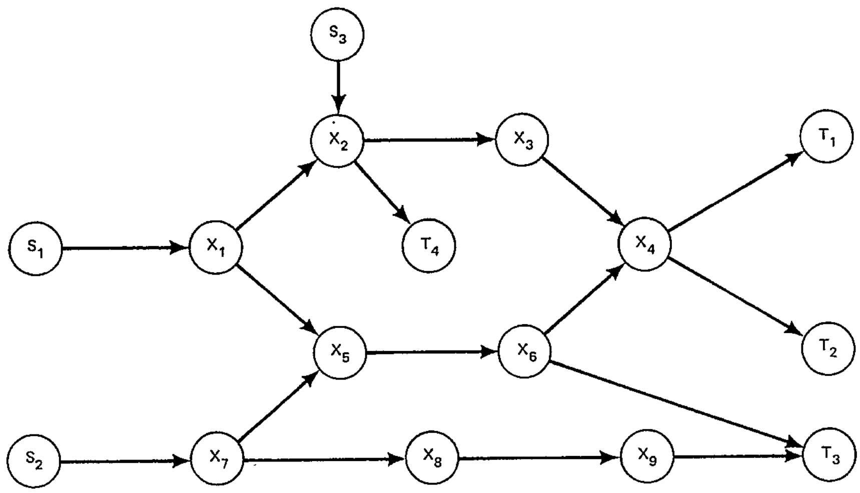
M_2: S_1, X_1, X_5, X_6, X_4, T_2

M_3: S_2, X_7, X_5, X_6, X_4, T_1

M_4: S_2, X_7, X_5, X_6, T_3

M_5: S_2, X_7, X_8, X_9, T_3

M_6^*: S_3, X_2, T_4



* M_6 illustrates a "competing" type mail flow that affects the utilization condition of X_2, but otherwise is not pertinent to the simulation. This type of mail will bot be simulated beyond certain nodes and is terminated at a sink node (e.g., T_4) where transit time statistics are not tabulated.

FIG. 1 A HYPOTHETICAL POSTAL NETWORK

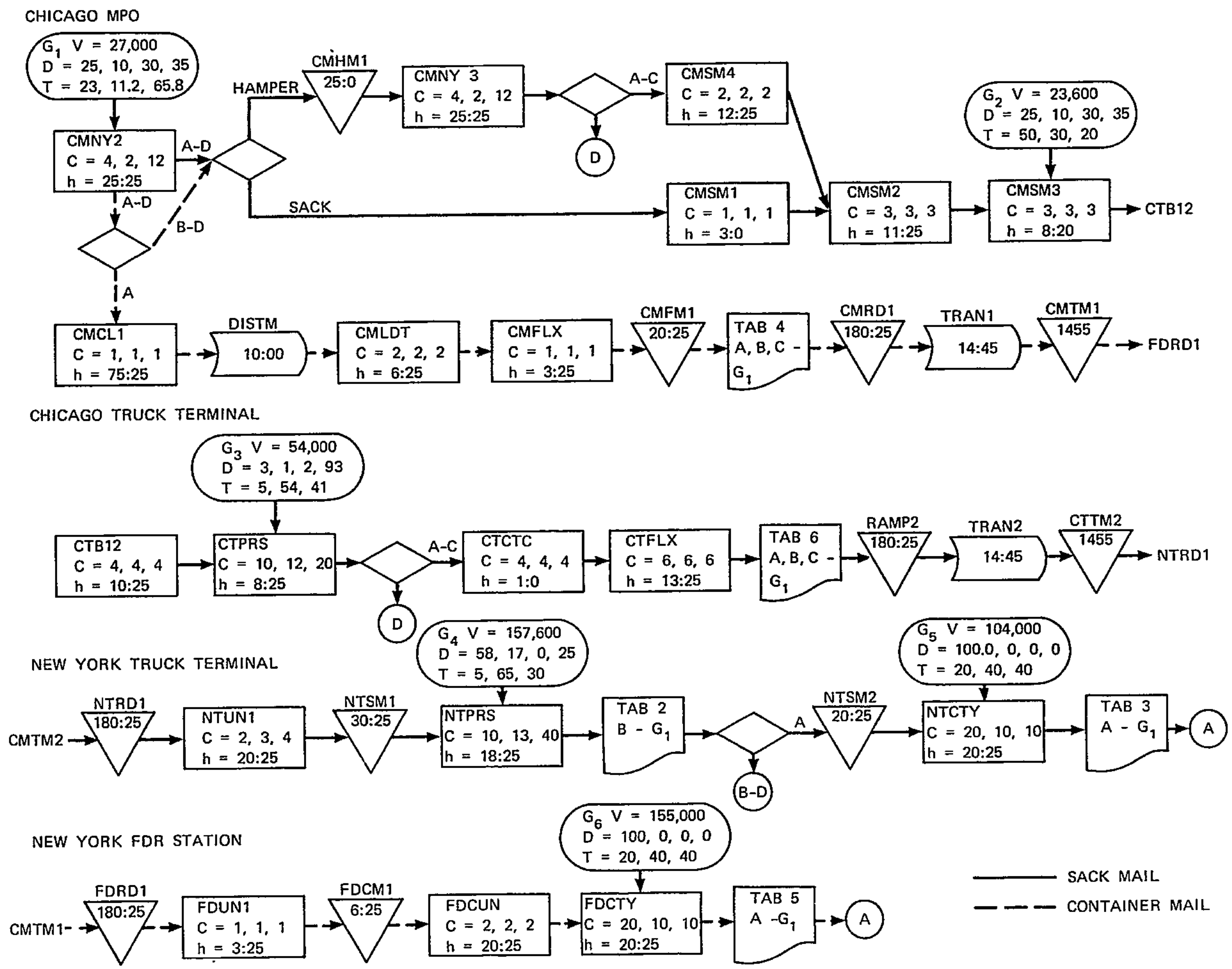


FIG. 2 EASTBOUND SACK AND CONTAINER SYSTEM

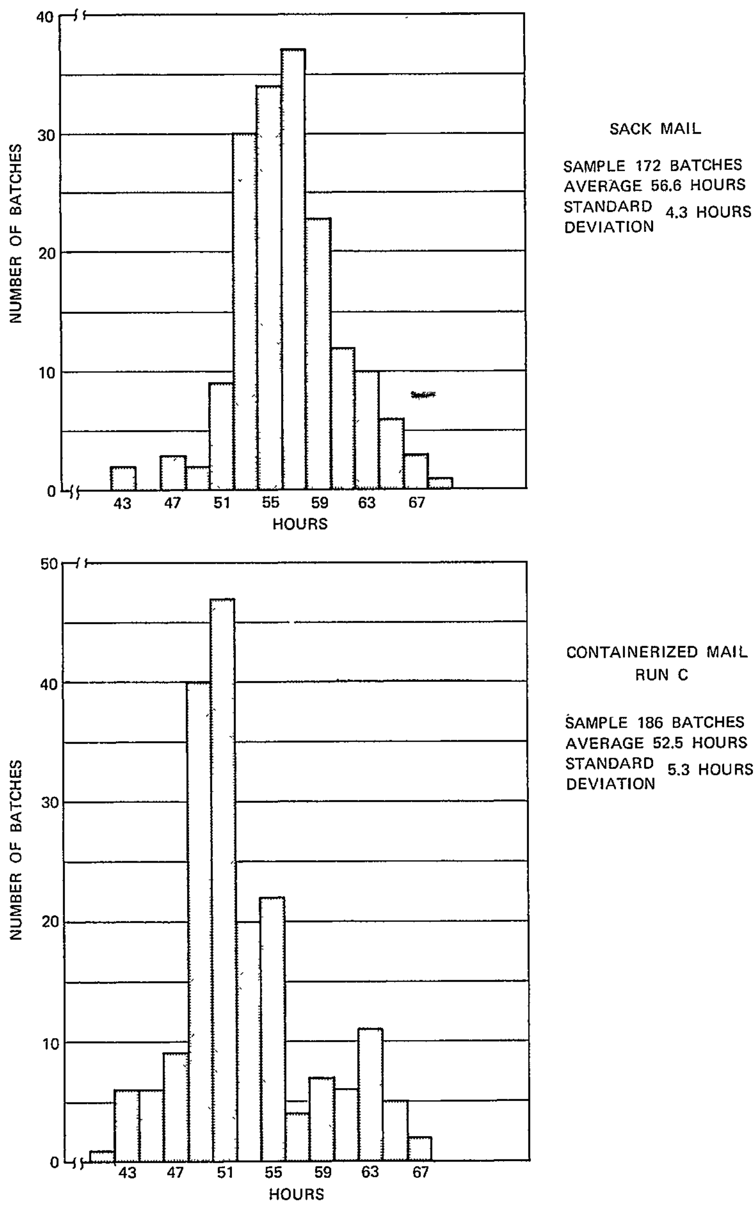


FIG. 3 TRANSIT TIMES -- MANHATTAN PARCEL POST -- BETWEEN CHICAGO MPO AND NEW YORK

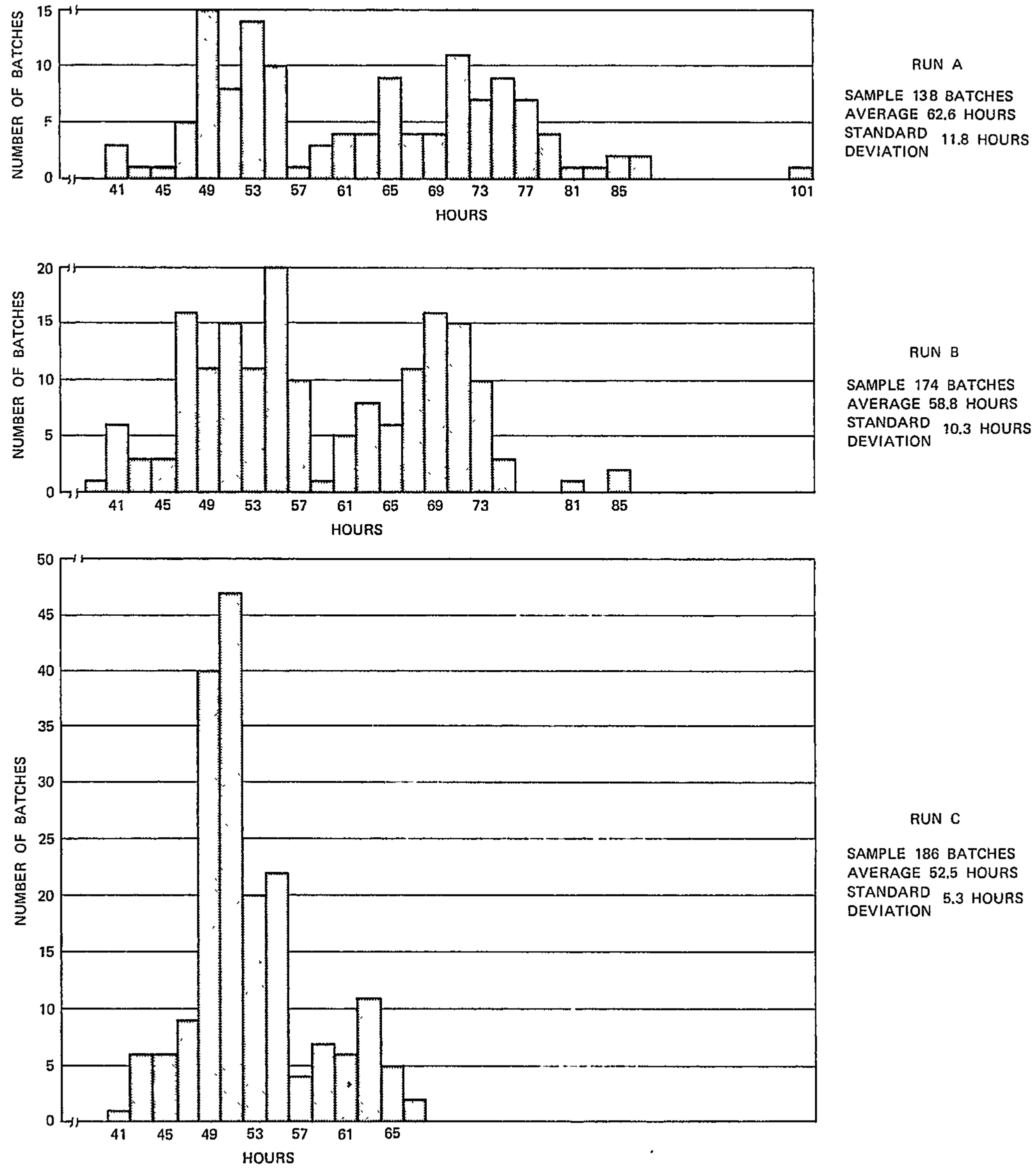


FIG. 4 TRANSIT TIME -- CONTAINERIZED PARCEL POST -- BETWEEN CHICAGO AND MANHATTAN

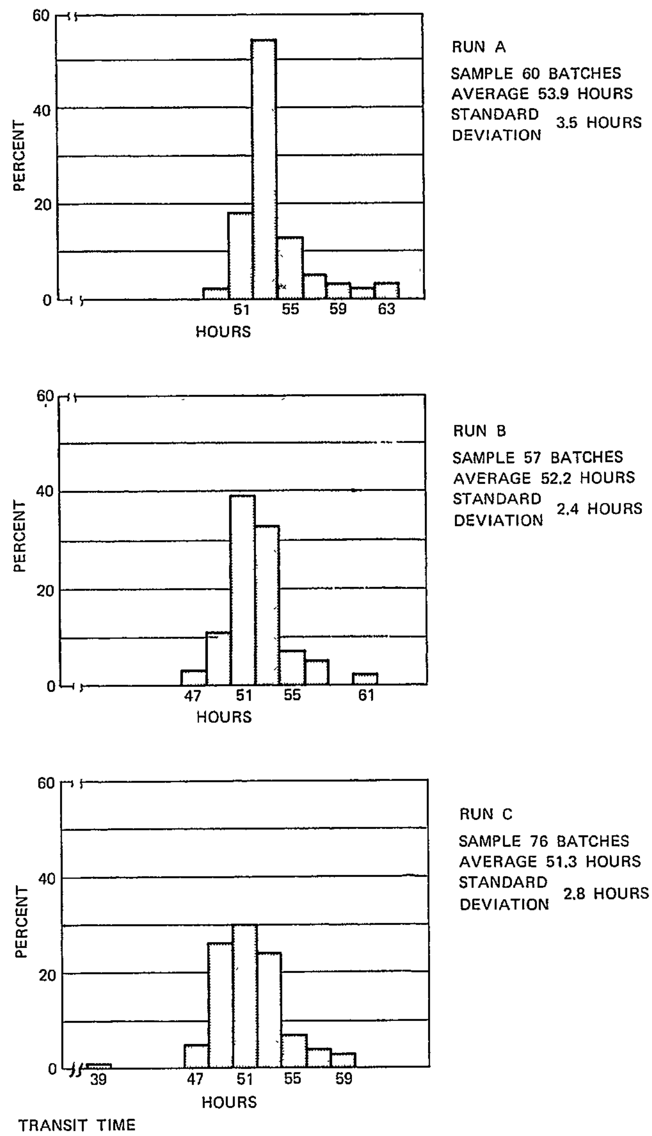


FIG. 5 TRANSIT TIME -- EASTBOUND CONTAINERIZED SYSTEM -- SACK MAIL

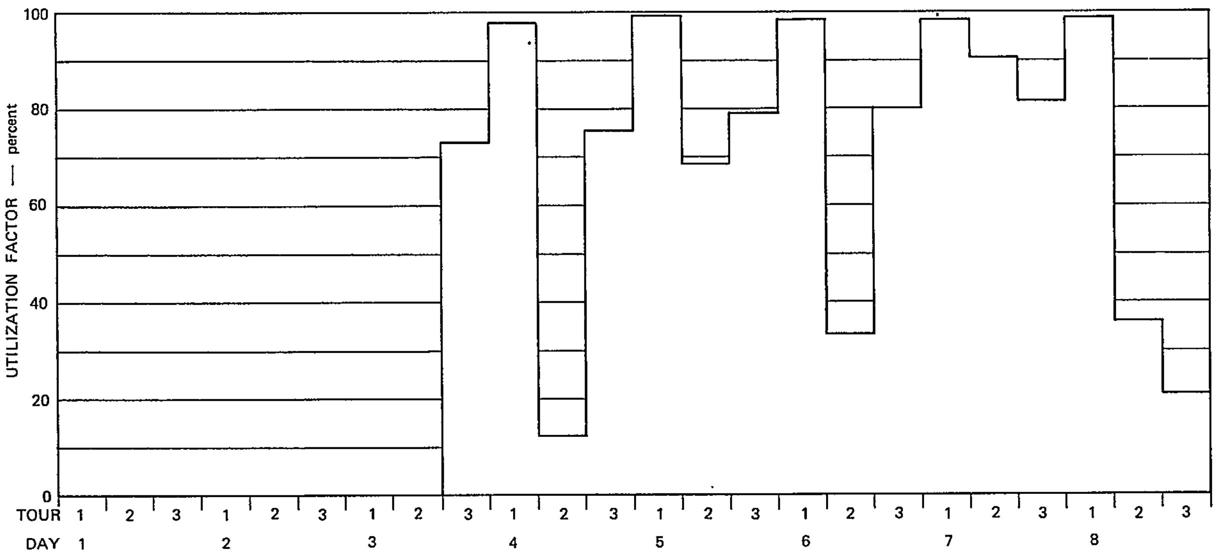


FIG. 6 - (a) UTILIZATION FACTOR OF NTUN1 -- SACK SYSTEM (EASTBOUND)

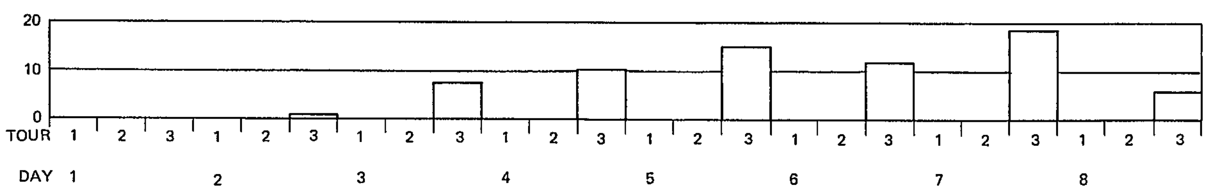


FIG. 6 - (b) UTILIZATION FACTOR OF FDUN1 BY TOUR -- CONTAINER SYSTEM (EASTBOUND)