Abridgment

Analysis of Fare-Collection-System Dependability

DAVID I. HEIMANN

The collection of transit system fares has become more sophisticated in recent years with more flexible fare structures. However, the more complex equipment such fare structures require has often been plagued by reliability problems, which results in significant passenger congestion and delay. Although development efforts are under way to improve reliability, one needs to know by how much the reliability needs to be improved. Attempting either too small or too large an improvement may result in a waste of transit funds and/or no relief from the congestion and delay problems. In order to determine the amount of improvement necessary, a method is needed to determine the dependability of a fare-collection system, i.e., the passenger congestion and delay in the system, given its demand, capacity, reliability, maintainability, etc. This paper discusses how a dependability analysis can be used to obtain reliability and other specifications and presents models to carry out such an analysis. Various types of dependability analyses are described (evaluation, sensitivity analysis, specification determination, and trade-off analysis), and purposes for which transit systems can use such analyses are discussed. Simulation and analytical models to evaluate fare-collection-system dependability are presented, as well as the data requirements for the models. A sample fare-collection dependability analysis that uses data based on an actual transit system is described, and the results and conclusions are discussed.

The collection of transit system fares has become more sophisticated in recent years as transit authorities turn to more flexible fare structures. As the use of extra personnel is often too costly, transit systems have turned to more sophisticated fare-collection machinery, which uses data processing and electronics in order to carry out the more involved fare-collection procedures that arise from such structures (1-4).

However, the newer and more complex a piece of equipment, the more likely it is to have frequent failures. High failure rates have indeed occurred, which leads to significant passenger delay, lower throughput capacity, and general frustration (5,6). Efforts are under way to increase the reliability of fare-collection equipment (5-8). The question that arises, however, is by just how much should the reliability be improved. Under some circumstances, the reliability improvement and its related monetary expenses may be ineffective.

For example, the improvement may be in the wrong service area. Either the main delay does not occur in the service area being improved or the improvement merely causes the delay to shift to a service area further downstream, with no decrease in overall delay.

Another possibility is that the reliability may be improved too much. When the reliability improvement is large enough, failures no longer happen often enough for further improvement to significantly affect system operation.

Measures other than reliability improvement may be more effective. Faster recovery times (i.e., maintainability) or having more units available for service (i.e., redundancy) may improve system performance as well as or better than reliability improvements and may be less expensive.

Finally, system failure may not be the main problem. Large surges of simultaneously arriving passengers, such as those coming from a major feeder bus line, may cause large delays.

In order to properly answer the question, By how much should reliability be improved?, one needs some way to find out the passenger delay in a fare-collection system, given information on its reliability, maintainability, number of machine units (redundancy), nominal processing rate, and passenger demand. In this manner, one can derive the proper mix and extent of improvements necessary.

Described in this paper are models that have been developed to examine the interrelation among reliability, maintainability, number of machine units, and passenger delay by analyzing the flow of passengers through the fare-collection system. These models treat the system as a network of queues, with the passengers moving from one service area to the next (a service area is a specific set of machine units, such as coin and bill changers, ticket vendors, gates, etc.). Superimposed on this network is the failure-recovery process by which units fail at a rate according to their reliability and are repaired according to their maintainability.

ANALYSIS PROCESS

The models calculate the congestion (queue length) and passenger delay in the fare-collection system, given the system configuration and passenger demand, for each of the service areas (i.e., ticket vendors, gates, etc.) in the system, as well as the delay for the overall system.

The models make possible at least four kinds of analyses: evaluation, sensitivity analysis, specification determination, and trade-off analysis. In evaluation, a given fare-collection system is examined, with the required information about the system collected and entered into the model as input data. Sensitivity analysis measures the sensitivity of congestion and delay to changes in input parameters (especially useful if one wishes to make changes or if some of the input parameters are questionable). Specification determination assesses the values of selected input parameters necessary to achieve a desired goal for congestion and delay (this is the reverse of sensitivity analysis, in that it measures the sensitivity of the selected input parameters to the congestion and delay goals). Trade-off analysis examines how two input parameters can be changed, with one being raised in quality while the other is lowered in quality, while keeping the overall performance constant (as compared with sensitivity analysis, which examines the interaction between an input and an output parameter).

The results produced by the fare-collection model are useful to transit properties for a number of different purposes, such as

1. Determination of required number of machine units,
2. Reliability and maintainability specifications,
3. Impact of changes in passenger demand,
4. Effect of maintenance policy changes, and
5. Effect of changes in fare-collection method.

TECHNICAL APPROACH

The basic approach to the model is to investigate the operation of the fare-collection system as a multiple-server queue, with passengers as customers, machine units as servers, and a first-come, first-served service discipline. In addition to the normal queue features, the number of servers (machine units) changes as the machine units fail and are repaired.

Simulation Model

Several machine units, each of which can serve one passenger (provided they are not otherwise busy or out of service because of failure), make up a service area. The simulation model has up to three service areas in sequence. Arriving passengers are assigned to initial service areas according to a given passenger-flow division. After completing service at a service area, a passenger continues to the next area until the service is completed at the final area (usually the gates), at which point the passenger departs the system. The model is an event-oriented simulation in which the next event to be processed is the earliest-occurring of the five basic events: passenger arrivals, passenger departures, equipment breakdown, equipment repair, and passenger continuation to the next service area. Events are processed until the time of the prospective next event is no longer within the time period being simulated.

On arrival, a passenger immediately begins service if a machine unit is available, in which case a departure time is calculated and put into the departure stack (which is an array of departure times of passengers in service, sorted in chronological order, used to determine the time of the next departure). If no machine is available, however, the passenger enters a first-in, first-out queue.

Passengers who find an available unit on arrival and thus immediately enter service have delay times of zero. Passengers who enter service from the queue have their (nonzero) delay times calculated at the time they enter service. The delay time is the interval between the arrival time and the time of entry into service. On departure, the passenger record is removed from the departure stack.

The reliability of the equipment being modeled is given in terms of the mean cycles between failures (MCBF). On the departure of a passenger from service, a random draw is made with probability 1/MCBF that the machine unit just used breaks down. If it does break down, its repair time is calculated and the unit is placed in the repair stack (which is an array of return-to-service times of failed units arranged in chronological order). In addition, if a breakdown occurs, the number of machine units available is decreased by 1.

The time necessary to repair a failed unit is assumed to be an independent random variable with an exponential distribution. (Note, the interarrival times between successive passengers and the passenger processing times are also assumed to be independent and exponentially distributed.) Repair time (maintainability) includes the time to report the failure and dispatch repair personnel as well as the time to actually do the repair. On repair, the unit is removed from the repair stack and the number of units available is increased by 1. If a queue exists when a unit returns to service, the first passenger in the queue enters service.

A passenger continuation is a departure by a passenger from an upstream service area to the next one in sequence. A continuation is treated as a simultaneous departure at one area and arrival at the next one in sequence.

Analytical Model

A simulation model does have some drawbacks. The randomly obtained congestion and delay probability distributions are subject to a number of statistical sensitivities; therefore, the simulation must be run several times for each situation. Furthermore, a simulation will require a large amount of computer time if many passengers must be processed, as would be the case at an important station during the peak period, which is the type of station one would most likely wish to investigate. Therefore, the simulation will require a large amount of computer time to carry out its analysis.

Therefore, in addition to the simulation model, an analytical model was developed to examine a fare-collection system. The analytical model directly solves the equations for the queue length probabilities from which the mean (and variance of) congestion and delay are obtained by using a modification of the Neuts and Lucantoni model (9) for the multiple-server exponential queue with a randomly varying number of servers. The analytical model needs to be run only once, as it avoids the statistical sensitivities that affect the simulation model, thereby reducing the amount of computer time required for the analysis.

DATA REQUIREMENTS

Two kinds of data are required for the model: hardware and passenger flow. The hardware data include reliability and maintainability data as well as the passenger processing rate per machine unit and the number of units provided for each of the service areas in the station. The passenger-flow data include the passenger arrival rate, group size, and division of passenger flow to the various service areas. Specifically, the data requirements are as follows:

1. Passenger arrival rate (one parameter for the entire system): The hourly rate at which passengers arrive at the fare-collection system during the peak period.
2. Group size (one parameter for the entire system): The size of a group of arriving passengers.
3. Passenger processing rate (one parameter for each service area): The hourly rate at which a machine unit in the service area can process passengers, and hence the unit's capacity to handle passenger flow. (Note, the actual rate in the field will be significantly less than the machine design capacity because of various types of passenger-induced delays. A special collection effort may need to be made to obtain this rate.)
4. Failure rates or reliability (one parameter for each service area): The rate at which failures occur to a machine unit, which makes it unable to process passengers. Because the basic measure of exposure to failure is the use of the unit by an individual passenger, the measure of failure rate is given as MCBF.
5. Repair times or maintainability (one parameter for each service area): The elapsed time (in hours) between the failure of a machine unit and its return to service. This is the sum of the times necessary to detect the failure, dispatch repair personnel, and perform the actual repair.
6. Number of machine units (one parameter for each service area): The number of machine units nominally available for passenger use in the absence of failures.
7. Division of passenger flow to service areas (one parameter for each service area): The proportion of arriving passengers who begin their use of the fare-collection system in that particular service area.

SAMPLE ANALYSIS

In order to demonstrate the use of the model for fare-collection-system analysis, a sample run was made based on preliminary data obtained from the Miami Dade County Transit Authority (the analysis is

Table 1. Effects of changes in number of gates and their reliability and maintainability.

Case	Gate Arrival Rate (per hour)	No. of Gate Units	Reliability (MCBF)	Maintainability[a] (h)	Gate Processing Rate (per hour)
1	5400	5	60 000	0.8	1350
2	5400	5	10 000[b]	0.8	1350
3	5400	5	3 000[b]	0.8	1350
4	5400	5	1 000[b]	0.8	1350
5	5400	6[b]	1 000[b]	0.8	1350
6	5400	5	1 000[b]	0.6[b]	1350
7	5400	5	1 000[b]	0.3[b]	1350
8	5400	5	1 000[b]	0.2[b]	1350
9	5400	5	1 000[b]	0.1[b]	1350
10	5400	4[b]	3 000[b]	0.8	1350

[a]Mean total downtime.
[b]Changes from base case (case 1).

of course based on preliminary configurations and does not necessarily reflect the final configuration of the actual Miami system).

The station analyzed in the model run is derived from the Dadeland North station during peak-hour operation. This station was selected because it is a relatively important station that has enough passenger demand to result in significant congestion and delay if enough machine units fail. The estimated peak-hour passenger flow at the sample station is 5400 passengers/h. Passengers are assumed to arrive singly, not in groups.

There are five gates at the station. Each gate has a physical capacity to process 1800 passengers/h. A rough rule-of-thumb for field processing capacity of 75 percent of the physical processing capacity is assumed for this analysis. Therefore, the gate processing rate used in the model runs is 0.75 x 1800, or 1350 passengers/h.

The reliability is 60 000 MCBF. The mean total downtime due to a failure (MTTR) is 0.8 h (48 min).

The analysis investigates the effects of changes in the number of gates and their reliability and maintainability. Ten cases are examined, as given in Table 1. The results for the gates are as follows [note, * = infinity (queue length exceeds 500)]:

Case	Mean Queue Length (congestion)	Mean Passenger Delays Excluding Processing Time (s)
1	3.2	0.5
2	3.2	0.5
3	9.2	6.6
4	40.2	34.1
5	3.3	2.6
6	28.0	24.2
7	20.4	16.7
8	7.3	4.6
9	5.3	2.6
10	*	*

A number of conclusions can be drawn from these results:

1. Evaluation of the given situation (case 1): no serious delay problems are expected from the fare-collection system as specified.

2. Sensitivity analysis of gate reliability (cases 1-4): The specification for gate reliability can be significantly reduced from its original level of 60 000 MCBF without seriously affecting delay. In fact, the reliability can decrease by almost an order of magnitude without serious impact. Delays start becoming significant when the MCBF reaches 3000 and become a problem when the MCBF reaches 1000.

3. Sensitivity analysis of increased number of gates under conditions of low reliability (cases 4 and 5): Adding one additional gate, which makes six units in all, when the gate reliability is low (1000 MCBF) is equivalent to improving the reliability to 10 000 MCBF.

4. Sensitivity analysis of maintainability under conditions of low reliability (cases 4 and 6-9): A delay problem due to low reliability can be solved for this system by improving maintenance response, but the improvement must be considerable (even an improvement from 0.8 to 0.1 h does not completely restore the performance of the base case).

5. Sensitivity analysis of decreased number of gates under conditions of marginal reliability (cases 3 and 10): The system cannot operate with fewer than five gates. If failures occur under a four-gate operation, the system will sustain catastrophic congestion and delay.

6. Trade-off analysis of reliability versus maintainability under conditions of low reliability (cases 3, 4, and 8): An increase in the reliability (of case 4) from 1000 to 3000 is approximately equivalent in delay impact to an improvement in the maintainability from 0.8 to 0.2 h.

SPECIAL NOTE

To supplement the above models, a cost module has been developed that computes the annual costs relevant to fare-collection dependability (i.e., equipment acquisition costs, spares costs, equipment operating costs, and scheduled and corrective maintenance costs). The module makes possible such analyses as cost/performance evaluations, sensitivity analyses of costs to changes in specifications, trade-offs between costs and performance, and trade-offs between different types of costs. A full-length report that describes in detail the models discussed in this paper, as well as the cost module, is available on request from the author.

REFERENCES

1. L. Rubenstein, J. Land, G. Desphande, and B. Harrow; Jet Propulsion Laboratory. Overview of Rail Transit Fare Collection. U.S. Department of Transportation, Rept. UMTA-CA-06-0116-80-1 (JPL Publ. 80-89), Aug. 1980.
2. G. Desphande, J. Cucchissi, and R. Heft; Jet Propulsion Laboratory. Rail Transit Fare-Collection Study; Volume 1: Policy and Technology Assessment. Transportation Systems Center, U.S. Department of Transportation, Cambridge, MA, Aug. 1981.
3. A Study of Transit Fare Policies, Fare Structures, and Fare-Collection Methods. Urban Transportation Research Branch, Department of Transport, Ottawa, Ontario, Canada, April 1978.
4. G. Surabian. Automated Magnetic Card Fare-Collection Systems: An Overview with Emphasis on the Washington Metro Experience. Urban Transportation Department, Northeastern Univ., Boston, May 1980.
5. J. Heisler and R. Stevens; Input Output Computer Services, Inc. Assessment of WMATA's Automatic Fare-Collection Equipment Performance. U.S. Department of Transportation, Rept. UMTA-MA-06-0080-81-1, Jan. 1981.
6. Automated Services, Inc. Automatic Fare-Collection Equipment Reliability and Maintainability Assessment Plan for Urban Rail Properties. U.S. Department of Transportation, Rept. UMTA-MA-06-0025-81-1, March 1981.
7. J. Koziol. Summary Report of Fare-Collection Workshop, March 11 and 12, 1980. Transportation

Systems Center, U.S. Department of Transportation, Cambridge, MA, March 1980.
8. Input Output Computer Services, Inc. An Assessment of ICG's Automatic Fare-Collection Equipment. Transportation Systems Center, U.S. Department of Transportation, Cambridge, MA, June 1981.
9. M. Neuts and D. Lucantoni. A Markovian Queue with N Servers Subject to Breakdowns and Repairs. Management Science, Vol. 25, No. 9, Sept. 1979, pp. 849-861.

Publication of this paper sponsored by Committee on Intermodal Transfer Facilities.

Bus Terminal Planning and Operation at the 1982 World's Fair

DAVID R. MILLER AND M. JANET REID

The design and operation of charter and tour bus and shuttle bus terminals at the 1982 World's Fair in Knoxville, Tennessee, are described. Constraints governing the design principles are discussed and operation policies are defined. Each terminal required a different type of layout and operating concept because of land availability and differences in the loading and unloading requirements of users of the types of services offered. Operating labor requirements, other factors influencing cost, and flow rates actually achieved at each terminal are discussed.

The 1982 World's Fair in Knoxville, Tennessee, was planned to attract 11 million visitors during its six-month duration. A modal split of 30 percent by public transit was predicted for the designed day volume of 80 000 persons. The public transit component of the Fair's planned transportation system included provisions for charter and tour buses; shuttle buses from local hotels, motels, and nearby communities; shuttle buses from parking lots included in the official World's Fair parking system; and the local bus service provided by the Knoxville Transit Authority through its operating arm, K-TRANS. Early estimates were that, on peak days, 700-800 charter buses might arrive, carrying some 30 000 Fair visitors. The local hotel and motel shuttles were predicted to carry a maximum of 5000 visitors/day, and the official parking lot shuttles an additional 10 000 visitors on peak days. (The parking lot shuttles, both official and unofficial, were counted as part of the automobile modal split and thus were not included in the 30 percent forecast.)

The World's Fair site (see Figure 1), which is bounded by the Knoxville central business district (CBD) to the east, the Tennessee River to the south, the University of Tennessee campus to the west, and an Interstate highway and local arterial streets on the north, posed many challenges to the transportation planners. The overall goal was to get visitors to and from the Fair as efficiently as possible while imposing a minimum of added congestion on the Knoxville street and highway system. Planners for the Fair's transportation system had to work within the following constraints:

1. Land adjacent to the Fair was scarce and costly,
2. The Fair management wished to invest the minimum amount possible in transportation facilities consistent with the goals stated above,
3. Charter and tour buses required parking for the day as well as terminal facilities for loading and unloading,
4. Terminal plans had to be compatible with existing or achievable highway capacity on the adjacent streets, and
5. The terminal system and traffic-flow rates had to mesh with the Fair's entrance gate designs and capacities.

The solution adopted was to assign the different types of bus traffic to terminals at the various Fair gates, thereby distributing the volumes and enabling the most appropriate type of facility to be designed for each kind of service. The design and operation of each of the three bus terminals are described in the remaining sections of this paper. Information on operating labor requirements and flow rates achieved is included for each type of terminal.

CHARTER AND TOUR BUS TERMINAL

A charter or tour bus was defined as a bus that transported a group to the Fair, dropped off the passengers, and then picked them up at a designated time. The buses used were typically standard intercity coaches or school buses with one front door for loading and unloading.

The area designated for the charter terminal was a triangular piece of land immediately adjacent to the Fair's north gate. It was selected because of its proximity and ease of access to the Interstate highway system that serves Knoxville, which made it possible to keep most of the long-haul bus traffic off the downtown Knoxville streets en route to and from the Fair. A policy decision was made that charter and tour buses would unload at a Fair gate, but that no attempt would be made to provide all-day parking for the buses in the immediate vicinity of the Fair due to lack of land. Hence, buses would have to deadhead to a parking area immediately on unloading and return to the bus terminal only to pick up their passengers and depart for the next destination. Anticipating the need for fueling, dumping station, and cleaning services, as well as minor maintenance, Fair management entered into an agreement with a local entrepreneur to provide parking for a minimum of 175 buses, with room for an additional 250 to be provided if demand warranted. Servicing and minor maintenance were to be available at the same location, which was approximately 4.5 miles from the north gate. Proposals were solicited from existing bus facilities to provide the layover area based on services available and acreage.