Interactive Network Model for Analysis and Cartographic Display of Railroad Traffic Flow

Jerome M. Lutin, Yehonathan Hazony, and Alain L. Kornhauser,
Princeton University, Princeton, New Jersey

To implement provisions of Section 503 of the Railroad Revitalization and Regulatory Reform Act of 1976, the Federal Railroad Administration sponsored the development at Princeton University of an interactive computer graphics model of the U.S. rail system. The model has been implemented and applied to a number of railroad problems, including the flow of freight traffic, accident locations, flows of hazardous materials, traffic diversions, and Amtrak passenger volumes. Some of the problems encountered in the development of the network model and the data base are described, and examples of applications of the model are offered and illustrated with copies of the graphic displays produced for each case study. The interactive system used for model development and implementation is discussed. Finally, some general observations are made on the value of interactive computer graphics in rail system planning.

In recent years, the U. S. rail industry has undergone a number of major crises and significant erosion in terms of physical plant, traffic, and revenue. Because of the poor health of the industry and its vital role in the nation's economy, the U. S. Department of Transportation (DOT) has become increasingly involved in planning and systems analysis of railroad traffic. Through the Railroad Revitalization and Regulatory Reform (RRRR) Act of 1976, the Federal Railroad Administration (FRA) received a mandate from Congress to begin planning for restructuring of the U. S. rail industry (1). FRA was to consider line abandonments, corridor consolidation, and rehabilitation priorities and in general define a national system that would serve U. S. rail transportation needs.

Under Section 503 of the RRRR Act, FRA was required to identify, classify, and designate class 1 railroad lines according to priority. Each segment of the rail system was to be evaluated as to its importance as a main or branch line within the network. To facilitate this analysis, FRA turned to interactive computer graphics as a technique for communicating information about the rail network.

To disaggregate the rail system into segments of comprehensible size, 132 geographic sectors were created (see Figure 1) (2). This geographic subdivision required the development of 132 maps for the final report, each map containing several hundred kilometers of railroad line. Only through use of computer-generated maps could the Section 503 report be produced within the limited time available.

To develop the necessary computer software and data base needed to produce the maps, the Office of Rail Systems Analysis and Information of FRA awarded a research contract to the Interactive Computer Graphics Laboratory and Transportation Program of Princeton University. Princeton had previously developed a capability to interactively assign traffic to networks and display results in real time by using Tektronix graphics computer terminals.

This interactive graphic traffic-assignment capability had been tested in an October 1975 study of the financial assumptions underlying the U. S. Railroad Association (USRA) Final System Plan for Conrail (3). Under contract to the Subcommittee on Transportation and Commerce of the U. S. House of Representatives, the Princeton team had digitized the Conrail network and flowed expected traffic to indicate line traffic densities. These estimates of traffic density were used to assess the impact of seasonal variations in traffic demand.

As a result of the development of the interactive graphic version of the FRA railroad network model, many projects have been undertaken in which interactive graphic display has been the primary technique for analysis and planning for rail traffic. This paper discusses some of the steps required to set up the model, shows examples of graphic output, and describes some of the projects in which the interactive model has been used. The interactive system used by the Princeton University Interactive Computer Graphics Laboratory is discussed. The paper concludes with some observations on the value of interactive computer graphics in rail systems planning.

SETTING UP THE INTERACTIVE FRA NETWORK MODEL

The original version of the railroad network model was digitized elsewhere and provided by FRA in machine-readable form (4). When the network was displayed on the screen of the interactive terminal, it became apparent that it would be necessary to correct for obvious errors, including a New York-Bermuda railroad connection and a node in the middle of Lake Michigan. Errors such as these are to be expected in any data base as large as the U. S. rail network model, which contains approximately 18 000 links and 16 000 nodes. Such errors are corrected from the terminal by using a microscopic graphical editor—a specialized processor written in APL and designed to address localized errors in the geocoded data base. The analyst uses thumbwheels mounted on the Tektronix 4015 keyboard to control a graphical cross-hair cursor that appears over the map display. The analyst moves the cursor to the nodes to be corrected and enters appropriate commands to delete the errant point or to relocate it to the proper position on the map by moving the cross-hair cursor across the screen.

The original digitized model did not include a geographic representation of the United States in terms of state and county political boundaries or the major water boundaries. Any proposed solution to the problems of the U. S. railroads includes severe geopolitical implications. The most fundamental analytical tool required, therefore, was the capability to correctly overlay the railroad network on a political map of the United States with reasonable accuracy to the county level, to show the geographical consequences of any analysis or proposed modifications. Thus, the first task encountered was the merger of the geographic component of the FRA data base with a Federal Highway Administration (FHWA) geopolitical data base that contained digitized state and county boundaries (5).

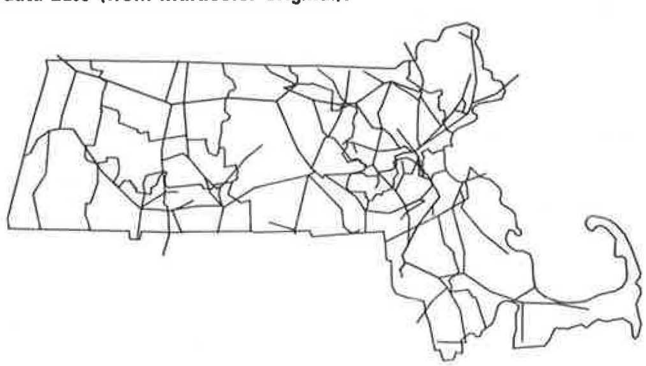
It turned out that the two data bases were produced under different cartographic projections, and the information for the projection used in the creation of the railroad model was not available. To overcome this problem, the geographic representation of the railroad system was converted to one that corresponded to the projection used for the geopolitical data base. The results of this conversion presented the first challenge to the interactive system since the two maps did not match even after the appropriate transformation had been applied. This called for the development of a macroscopic graphical editor, which is written in APL and is capable of addressing global errors in the data base (6).

A careful inspection of the two overlaid maps revealed the simple nature of the distortion and provided the basis for a very simple but accurate solution. The distortion, being unique, required a unique solution: The part of the railroad network east of Chicago had to be horizontally stretched and the part west of Chicago horizontally contracted so that the East Coast and West Coast ends of the

Figure 1. FRA map of U.S. rail planning sectors.

Figure 2. Federal railroad system.

Figure 3. Massachusetts railroad network overlaid on geopolitical data base (from multicolor original).



network matched the corresponding U. S. coastlines. This was done by means of a combination of two linear transformations in which the stretching and contraction were proportional to the horizontal distance of all points from Chicago. The resulting merged map is shown in Figure 2. This transformation survived close inspection along rivers and state boundaries and became the geographic data base for the project.

Figure 3 shows a black-and-white representation of a multicolor original in which the subnetwork of a single state is overlaid by the corresponding part of the geopolitical data base. Note that the solution to this crucial problem did not come from an available software package but rather through flexibilities of interactive graphic display and ease of programming for improvised one-shot solutions. The macroscopic editor did an excellent matching job on the Massachusetts case shown in Figure 3. However, there remain local coding errors (some links extend into the ocean), correction of which will require the use of the microscopic editor. Some of the more subtle errors will emerge only after more sophisticated analysis.

The next stage in the processing of the geographic component of the data base was the search for faults in the connectivity of the network, such as missing segments and "tunnel-link" connections. Since the network contains almost 20 000 links, it was difficult to verify connectivity by visual inspection of many expanded displays. However, inspection of the graphical flow of minimum-path trees from key nodes very quickly revealed faults of connectivity caused either by the omission of major parts of the network or by evolution through unrealistic paths. Similarly, tunnel-link effects have been discovered in the disappearance of paths off the screen and their reappearance in other parts of the country. These tunnel-link effects have been traced to inconsistencies in either the station code number level or the station abbreviation code level of the data base. Although these inconsistencies occurred on a nongraphical level of the data base, graphical methods proved to be very powerful in tracking them down.

In a study of the seasonality of traffic flow on the national network, it was necessary to flow each freight transaction in the 1 percent waybill data base available from FRA and summarize the flow over the links (7). The computer used the minimum-path algorithm to route the traffic from Detroit to New York State along the Canadian side of Lake Erie (actually, the major part of the traffic flows on the southern shore of the lake). Close inspection of the data revealed that the error occurred at the commodity level and that the Canadian link included an old tunnel that is too narrow to accommodate triple-decker automobile-carrying cars. In other words, one encountered situations in which the connectivity of the network depended on attributes other than the physical fact of a railroad track connecting two geographic points.

It took a graphical analysis of seasonal freight flows to identify this particular feature of the railroad system. This was accomplished by plotting freight volumes as rectangles on each link by direction. The width of the rectangles was proportional to the traffic volume in carloads, and direction was indicated by drawing rectangles on each side of the link to represent respective directional flows. The combined value of pictorial analysis and large-memory computers becomes apparent when one realizes that the 1 percent waybill data base used in the analysis consisted of about 600 000 records of 80 columns each. This is equivalent to 600 000 computer cards on which all 80 columns are punched.

A more complicated connectivity problem arose in another study in which the 1 percent waybill data were flowed over the network and a record was kept of track ownership along the way. After considerable analysis had been performed, execution was halted when the minimum-path algorithm flagged that it had encountered a disconnected network. This was surprising since the model was well "beaten" at this point and errors in network connectivity were not expected to occur.

In this instance, a technique known as "zoom" was applied in which the analyst selected a small portion of the display with the cross-hair cursor and commanded the machine to enlarge that portion to fill the screen. By using the zoom capability, it is possible, in several stages, to achieve any degree of enlargement desired. By using graphical zooming, the particular transaction at which execution halted was traced and the geographic location identified.

It turned out that the connectivity fault occurred on the track-ownership level of the data base. In other words, the physical network was connected through a link that represented tracks of different ownership, which was inconsistent with the record of the 1 percent waybill data. Inspection of the network in the location revealed further inconsistencies related to records of dual ownership of links and nodes. The network was edited interactively, and execution was resumed without repeating the entire costly analysis. Only a small portion of the data had to be rerun.

In a study of a hypothetical merger situation, a competitive line was added to an existing system (8). An inspection of freight flow on the combined network revealed an unacceptable pattern in which the freight flow did not take full advantage of the newly added line (see Figure 4). Graphical zooming on the display revealed a missing link that was too short to be noticed in any other way. The expanded window shown in Figure 5 clearly shows the missing link.

Figure 4. Unexpected freight flow caused by missing link (from multicolor original).

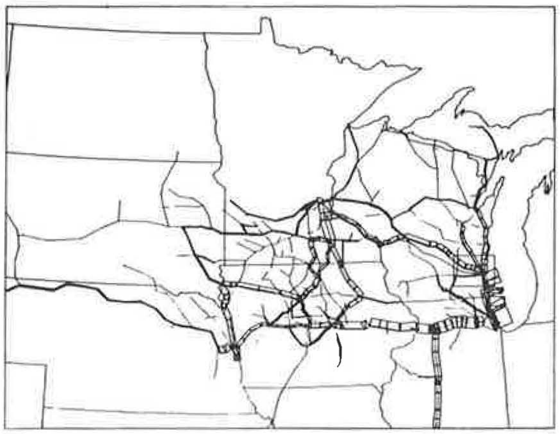


Figure 5. Missing link found.

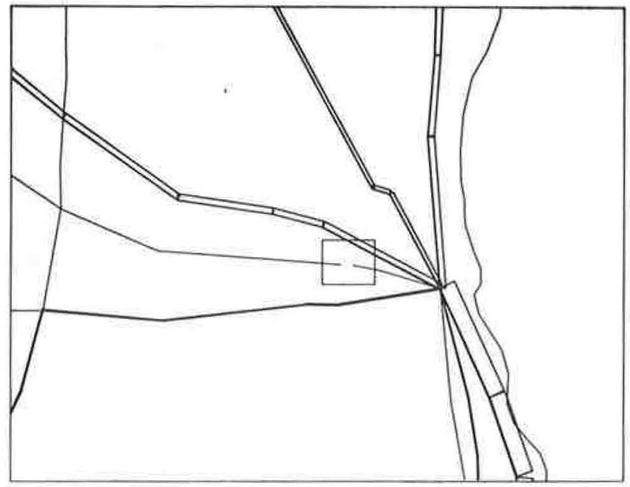


Figure 6. Geocoding of rail accidents: stage 1.

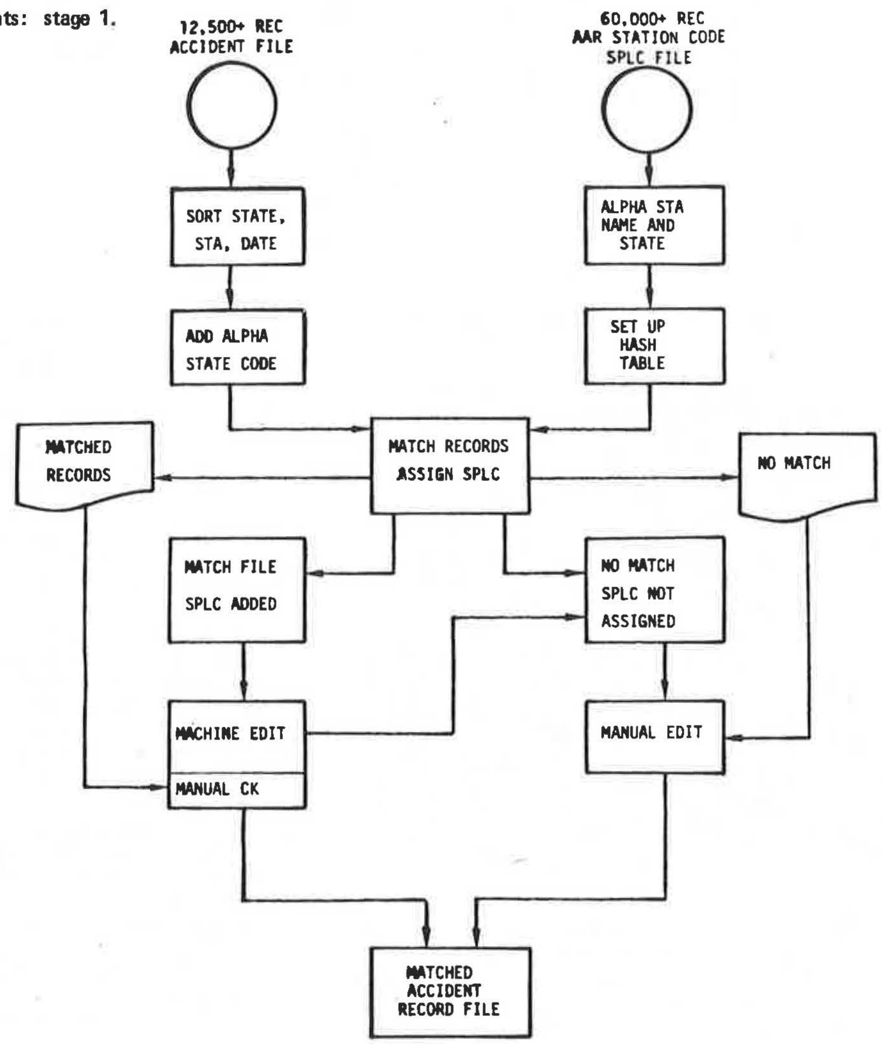


Figure 7. Geocoding of rail accidents: stage 2.

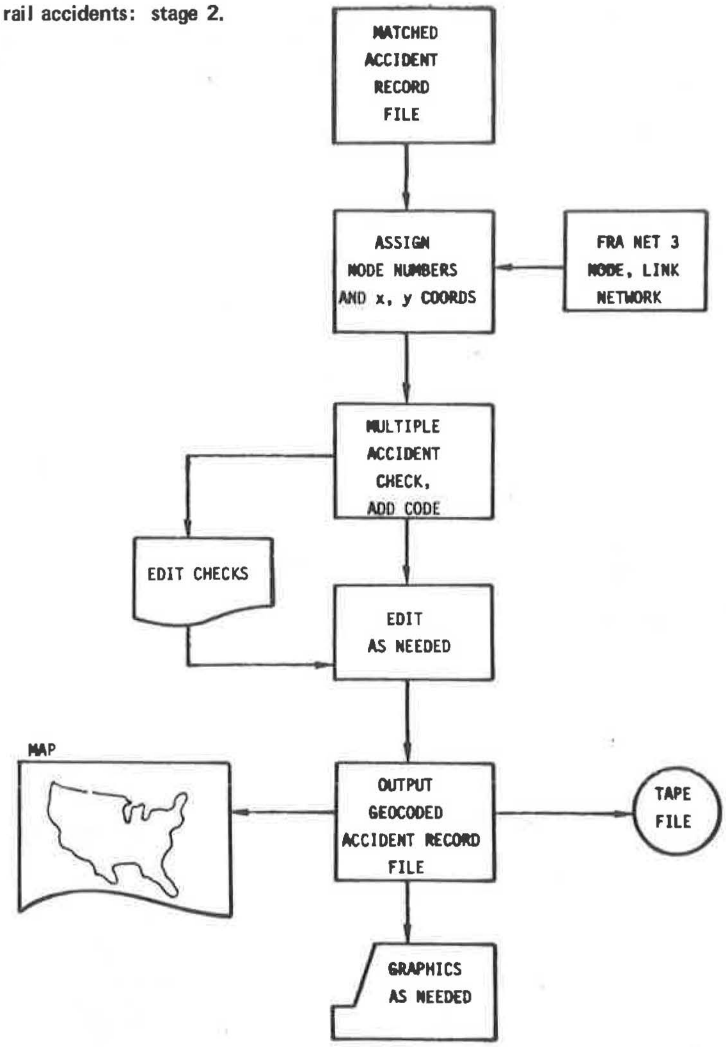
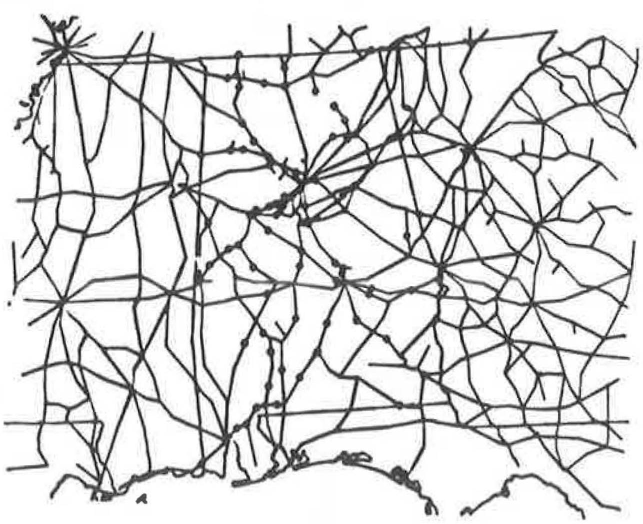


Figure 8. Alabama accident sites superimposed on FRA network links.



The situations and solutions described above concentrate on the pictorial explorations and interactive handling of problems of data integrity in geographically based, large-scale management information systems. However, the flexibility demonstrated above is useful not only for chasing after subtle errors in the data base but also for responding to challenges that were not originally planned for when the data base was designed and created.

MAPPING ACCIDENT LOCATIONS

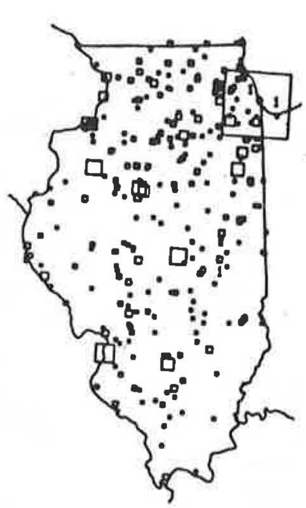
Since 1966, all class 1 and 2 railroads have been required to file monthly reports of all accidents that involve casualties and equipment or property damage above a given monetary threshold (readjusted periodically). These data are placed on computer tape and used to produce annual tabulations and summaries (9). In a program initiated by FRA to prepare more detailed analyses of these accident reports, one of the highest priorities was the identification of areas that have a higher incidence of accidents, especially accidents involving hazardous materials. To obtain a quick impression of the geographic distribution of accidents, the Reports and Analysis Division of the Office of Standards and Procedures of FRA turned to the use of interactive computer graphics to prepare maps of geocoded rail accidents. A pilot project involved geocoding and plotting reported accidents for 1976 (10).

A computer tape containing 12 555 records that represent rail equipment accident-incident reports (FRA form F6180-54) for 1976 was received from FRA. Each record contained the name of the freight station nearest the accident site. But, to locate each accident on the FRA network model, which contained the graphical coordinates needed for plotting, each accident had to be assigned a location code that corresponded with FRA identification codes for

Figure 9. U.S. railroad accidents in 1976.

Figure 10. U.S. railroad accidents in 1976: spills of hazardous materials.

Figure 11. Illinois railroad accidents in 1976.



network nodes. Fortunately, a correspondence table existed between FRA network node numbers and the six-digit standard point location codes (SPLCs) used by the American Association of Railroads (AAR) to reference freight stations (11).

A tape that contained approximately 60 000 station names and SPLCs was obtained from AAR and, by matching the station name on each accident record with those on the AAR tape, a numeric SPLC was obtained for each accident record that contained a valid station name. The SPLC was added to each record, which permitted each accident to be assigned to a point on the FRA network model. The accident records were then amended to include alphabetic state code, SPLC, FRA network node number, location coordinates, and counts of accidents and spills of hazardous materials that occurred at each network node. The entire process is shown in Figures 6 and 7.

A variety of graphical output was produced for this project. Figure 8 shows an example produced for an interim presentation. In the figure, accident sites (denoted by large dots) have been superimposed on the FRA network model links and state boundaries for Alabama. A full-size 92-cm (36-in) Calcomp plot produced for the project was done in three colors. The state boundaries were drawn with a blue technical pen. Each accident node was shown symbolically as a square drawn in black technical pen, and the side of the square was proportional to the number of accidents reported at that node. Hazardous spills were denoted similarly by a red circle, the diameter of the circle being proportional to the number of accidents involving spills.

Figure 9 shows a Tektronix photocopy of the data file plotted for the entire continental United States. A plot of hazardous accidents is shown in Figure 10. Figure 11 shows a plot of the state of Illinois, one of several state maps requested by FRA. In these plots, the area of the square is proportional to the number of accidents. A constant scale is used for all accident symbols; map scale varies to permit the largest possible image to be drawn. A circle is used to denote spills of hazardous materials, and the number of these accidents is shown to the right of the circle.

In a subsequent study for the Transportation Systems Center of DOT, the pressure for immediate solutions to the hazardous-materials accident problem required the speedy creation and analysis of a data system that correlates the flow of hazardous materials and train accidents (12). An interactive traffic assignment was run for hazardous freight to determine the number of carloads flowing on each link of the FRA network. In this assignment, traffic records from the FRA 1 percent sample of carload waybills for 1976 were selected based on the commodity codes for hazardous materials. These shipments were assigned to network links and displayed as rectangles, the width of the rectangles being proportional to the number of carloads. The direction of flow was shown by drawing the rectangle to the right side of each link. The traffic-flow data were superimposed on a plot of accident data, as shown in Figure 12. Here, accidents were disaggregated by type, and different symbols—octagons and triangles—were used to distinguish main-line from other accidents. These plots were prepared in four colors: green to indicate flow volumes, black for rail links and accident symbols, blue for state boundaries, and red for hazardous-materials accidents.

As the data were being processed for plotting, accident tabulations were calculated and printed in the legend. Variable scaling was used, and maximum symbol sizes were calculated and inserted in the legend. Thirty-two class 1 railroads were plotted, and color Xerox copies were prepared for inclusion in copies of the final report.

ANALYZING AMTRAK PASSENGER FLOWS

Another example of interactive graphic analysis of rail traffic is shown in Figure 13. In conjunction with a DOT study of the future of Amtrak, it was necessary to examine present and expected passenger volumes on each link of the Amtrak network (7). Machine-readable records of ticket sales were used to obtain station-to-station volumes on the network. A correspondence file was created to match each Amtrak station with the closest node on the FRA model. Then an "Amlink" network was created by extracting the appropriate subset of nodes from the FRA network. Station-to-station flows were assigned to the network and plotted.

Figure 12. Flow of hazardous materials superimposed on plot of accident occurrence.

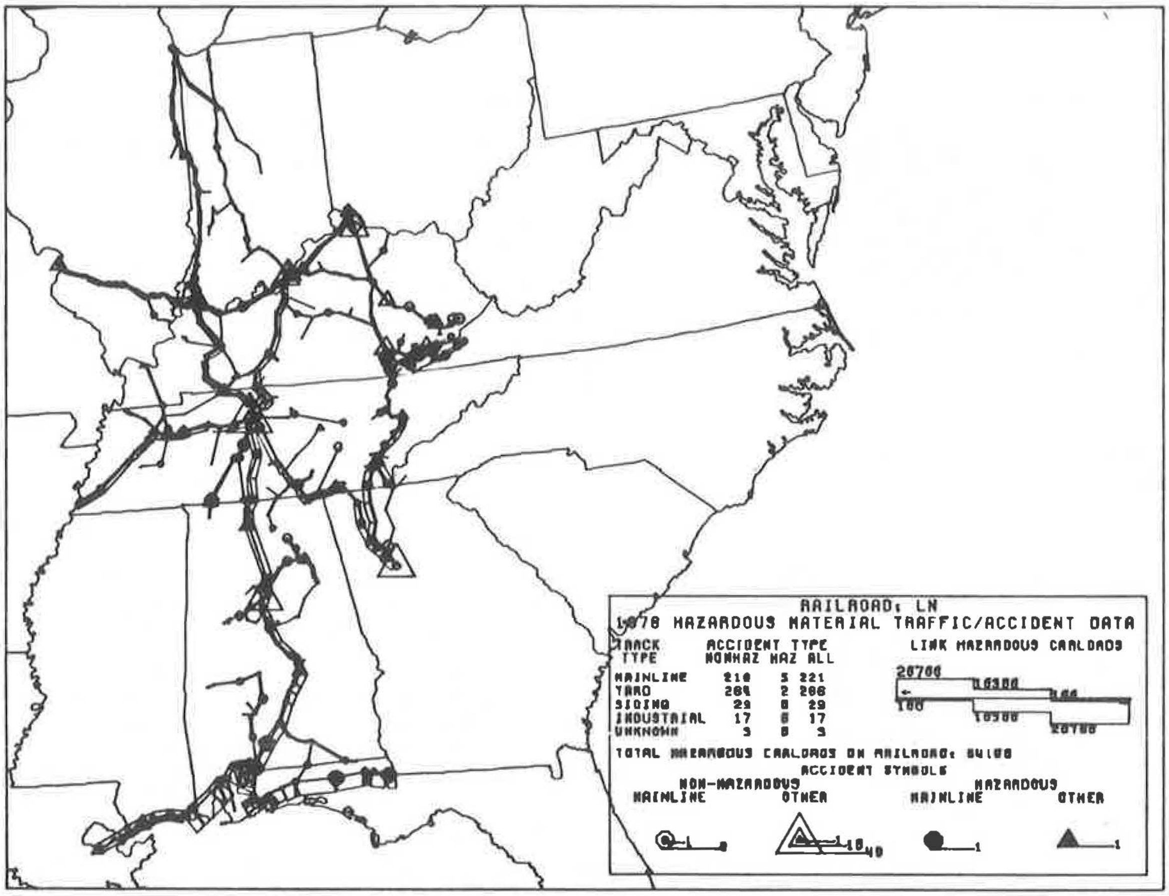


Figure 13. Passenger flow on Amtrak subsystem.

SYSTEM CONFIGURATION

The railroad analyses reported above have all been conducted by the Interactive Computer Graphics Laboratory of Princeton University. An IBM 370/158 computer, running under VM370/CMS, is used for time-sharing applications. This machine is supported by an IBM 360-91, which is used for batch operations running under the OS operating system. Both systems are interconnected so that jobs can be entered interactively and processed in batch mode if desired. The system supports approximately 20 Tektronix 4013/4015 graphic terminals and a similar number of nongraphic terminals. Two Tektronix 77x102-cm (30x40-in) digitizing tablets are available as well. Graphical output is obtained from a Calcomp 936 multicolor 92-cm (36-in) plotter, two smaller Calcomp plotters, and three Tektronix photo-processing hard-copy units. In several projects, the system was accessed remotely via data phone from a remote job-entry station at FRA in Washington, D. C. Programmers at Princeton developed software and map displays that were transmitted to Washington, where they were plotted by FRA personnel on a Calcomp 1036 plotter. A test run was also conducted from the Transportation Systems Center in Cambridge, Massachusetts, by using a Tektronix plotter.

The core of the system is VSAPL, enhanced by a graphical processor and an auxiliary processor that interfaces APL with lower-level languages like ASSEMBLER, FORTRAN, and PL1. The virtual machine environment permits the authorized user to define a virtual computer of a size of up to 16 million bytes. The main graphical processor has been implemented under APL as an auxiliary processor written in ASSEMBLER but highly transparent to the APL user. It has been designed for the efficient display of a high density of information on Tektronix 4013 and 4015 terminals in a time-sharing environment.

The mathematical power of the APL language is available for interactive data analysis. Arithmetic and logical manipulations of arrays and tools of Boolean algebra obviate many of the tasks required for the creation and manipulation of large-scale geocoded management information systems. Abstraction and aggregation of data by geographic and nongeographic attributes are easily achieved, as are analysis and display of statistical data.

CONCLUSIONS

Planning for the future of the nation's railroads is one of the most difficult and complex tasks facing the U. S. Department of Transportation. The problem is particularly difficult in that it demands a great deal of data and the results of any action will affect specific regions of the country and have obvious political repercussions. Consequently, federal rail system planners must, in a short time, test many alternatives to arrive at a feasible solution. In this planning context, interactive graphic analysis has great potential for application. The speed of the large computer coupled with the facility of APL makes it possible for the analyst to generate many solutions in a short time, and the graphic display permits him or her to grasp, almost at a glance, the geographic implications of the analysis. Finally, the power and clarity of the graphical display permit the analyst to communicate results immediately to policymakers and to the public.

REFERENCES

1. Office of the Secretary of Transportation. Preliminary Standards, Classification and Designation of Lines of Class I Railroads in the United States. Federal Railroad Administration, U. S. Department of Transportation, Aug. 3, 1976.
2. Office of the Secretary of Transportation. Final Standards, Classification and Designation of Lines of Class I Railroads in the United States. Federal Railroad Administration, U. S. Department of Transportation, June 30, 1977.
3. Princeton University Transportation Program. Evaluation of the Financial Assumptions Contained in the USRA Final System Plan. Subcommittee on Transportation and Commerce of the Committee on Interstate and Foreign Commerce, U. S. House of Representatives, Committee Print 13, U. S. Government Printing Office, Oct. 1975.
4. Federal Railroad Administration Network Model User's Manual. Federal Systems Division, International Business Machines Corp., Gaithersburg, MD, May 1975.
5. Urban Transportation Planning: General Information and Introduction to System 360. Federal Highway Administration, U. S. Department of Transportation, March 1972.
6. Y. Hazony. Interactive Cartography: An Analytical Tool for Management Information Needs. Transportation Program, Princeton Univ., Princeton, NJ, Transportation Rept. 78-TR-47, 1978.
7. A. L. Kornhauser. Final Report: Development of an Interactive-Graphic Computer Model for the Nationwide Assignment of Railroad Traffic. Transportation Program, Princeton Univ., Princeton, NJ, Transportation Rept. 77-TR-17, Sept. 1977.
8. R. Antush and A. L. Kornhauser. Railroad Traffic Diversion: Elementary Analysis of Originated Traffic. Transportation Program, Princeton Univ., Princeton, NJ, Transportation Rept. 78-TR-43, 1978.
9. FRA Guide for Preparing Accident/Incident Reports. Federal Railroad Administration, U. S. Department of Transportation, 1976.
10. J. M. Lutin; Gellman Research Associates, Inc. Railroad Accident Network Data for 1976. Federal Railroad Administration, U. S. Department of Transportation, Final Rept., 1978.
11. Freight Station Accounting Code Directory. Economics and Finance Department, Assn. of American Railroads, Washington, DC, 1976.
12. J. M. Lutin. Hazardous Materials Commodity Flows and Railroad Accidents for 1976. Gellman Research Associates, Inc., Jenkintown, PA, 1978.

Computer Graphic Animation of UTCS-1/NETSIM Traffic Flows

Amir Eiger, Shih-Miao Chin, and David Woodin, Department of Civil Engineering, Rensselaer Polytechnic Institute, Troy, New York

The development of a program for computer graphic animation of network traffic flows to increase the potential effectiveness of the UTCS-1/NETSIM network simulation package by visually displaying vehicle movements through street networks is described. The program is a tool that can aid traffic engineers in generating and evaluating alternative control strategies by enabling them to view network traffic in real time. It thereby provides them with additional information, such as lane-specific delays and queue buildups, that would not be readily available through the standard statistical output of UTCS-1.

As traffic flows through street networks, it experiences periods of congestion that may result from inadequate geometric design or signalization or simply excessive demand. Traffic simulation techniques are becoming important tools for the traffic engineer in investigating the impacts of various traffic control strategies. These simulation experiments can yield an enormous amount of data that could not be obtained in real life for economic or other reasons.

Among the network traffic simulation models, the Urban Traffic Control System (UTCS-1) model produced for the Federal Highway Administration (FHWA) has been the most popular. UTCS-1 has been extensively validated and is generally considered to yield reasonable results. It is excellent for evaluating signal control schemes and the effect of buses on traffic movement. The UTCS-1/NETSIM model is a fully microscopic model that collects and updates statistical network data at 1-s intervals. At the end of the simulation period, link-specific and system-wide data are printed out for use by the traffic engineer in evaluating the operational characteristics of the network. The aggregated nature of the data makes it useful in defining the existence of potential traffic problems. However, the user would be likely to get a more intuitive understanding of the conditions that gave rise to these potential problems and be more capable of isolating effective control strategies if vehicle flows were displayed visually.

The use of computer graphic techniques and computer-generated films for visual displays of traffic flow information is a relatively recent development. Parakh (1) proposed to improve the analysis of large-scale flow problems by visually displaying the output of computer simulation.