

Bureau of Transportation Statistics

Technical Report

June 2007

How Freight Moves: Estimating Mileage and Routes Using an Innovative GIS Tool

by **Stephen M. Lewis** and
Felix Ammah-Tagoe, Ph.D.

The Bureau of Transportation Statistics (BTS) has developed an innovative software tool, called GeoMiler, that is helping researchers better estimate freight travel. GeoMiler is being used to compute mileages along likely routes for the nearly 6 million freight shipments expected to be reported in the 2007 Commodity Flow Survey (CFS), the nation's largest survey of freight movements. These computations are used in estimating modal ton-miles of freight—a key measure for understanding the use and performance of our nation's freight transportation system.

BTS, part of the Research and Innovative Technology Administration (RITA) of the U.S. Department of Transportation (USDOT), developed GeoMiler using current Geographic Information System (GIS) technology to assign routes and calculate mileage from the true origin to the true destination of each freight shipment even when more than one freight mode is used. While developed for use in the CFS, the tool's integration

of core GIS technology and its modeling approach can be used for any multimodal freight movement at all geographic levels.

Background

The Commodity Flow Survey, a nationwide survey of U.S. businesses that generate freight shipments, is the primary source of national data on the flow of goods—from where to where, how far, and by what means. The 2007 survey, sponsored by BTS and conducted by the Census Bureau, is currently underway; it covers 100,000 U.S. business establishments primarily engaged in manufacturing, mining, and wholesale activities. Visit www.bts.gov/cfs for additional information about the survey.

For the prior CFSs, conducted in 1993, 1997, and 2002, shipment distances were estimated by using a set of routing modules that are now dated and do not fit within the current USDOT Enterprise Architecture. There were individual modules for air cargo, surface modes, and exports. As a result, the CFS records had to be processed multiple times. Because GIS technology was not used, the

models relied on flat ASCII-file representations of the transportation networks. These networks were cumbersome to maintain. Without the GIS component, there was no way to visualize on a map the routes that were generated.

For the 2007 CFS, BTS wanted a tool that met CFS' operational requirements while being easy to maintain and upgrade. This effort was part of a major performance push to improve overall efficiency, methods, and quality of mileage information in the 2007 CFS.

Why the Need for Mileage Estimation?

CFS shipment mileage is estimated by BTS after the survey data have been collected from respondents. The respondents were shippers and not transportation providers (carriers). Shippers know the actual origin and destination of their shipments but may know little about the specific modes of transportation used by carriers, the ultimate routes taken, and the distance traveled. The survey asks for origin and destination ZIP Code and the mode sequence (e.g., highway-rail-highway) used for each shipment.

From this information, BTS must calculate the mileage traveled by single mode (highway, rail, water, air, parcel, and pipeline) and by many multimodal combinations.

Few, if any, off-the-shelf national-level multimodal transportation networks are available for use in a GIS. In addition, there are almost no commercial routing modules capable of seamlessly handling intermodal transfers in the specific modal sequence provided by shippers. For example, routing a grain shipment by truck, rail, inland waterway, and truck is not a simple shortest path problem. It requires routing logic that allows the shipment to transfer at commodity-appropriate intermodal terminals and water docks. This routing logic is built into GeoMiler.

The absence of a readily available commercial solution for such multimodal routing led BTS to develop the GeoMiler network routing tool, using current GIS technology and robust spatial network data. It integrates map visualization features with route solvers to handle many alternative multimodal combinations from the following list of modes represented in the CFS:

- parcel delivery, courier, U.S. postal post;
- private truck;
- for-hire truck;
- railroad;
- shallow-draft (Inland Water) Vessel;
- deep-draft (Ocean) Vessel;
- great Lakes vessel (generated by routing models);
- air;
- pipeline;
- unknown (don't know or missing);
- other; and
- multimode (any combination of the above).

GIS Platform and System Architecture

The operational requirements for accurately estimating mileages for the nearly 6 million shipments, of which over one-tenth involved some multimodal transfer, required a multitier logical and physical

Figure 1. GeoMiler System Flow

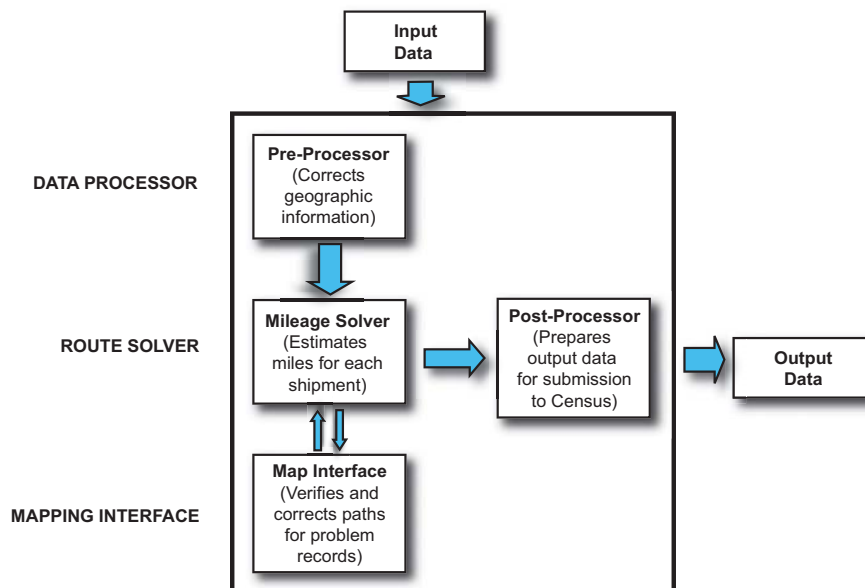


Figure 1 shows the three GeoMiler components – data processor, route solver, and map interface. The data pre-processor hosts all the input survey data received from the Census Bureau, and the post-processor returns the output data to the Census Bureau. The route solver performs the mileage calculations for all the freight modes. The mapping interface allows visualization of routes for any shipment records that have problematic input data and allows shipment paths to be checked, verified, and corrected.

SOURCE: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics.

design (figure 1). This design allowed data pre-processing separate from the route solver, modeling logic, and mapping interface. This system design and the development platform meet USDOT's Enterprise Architecture requirements.

- In this multitier configuration, the entire tool, including the data pre-processor and post-processor, was developed with a software tool that uses the BASIC programming language.
- With the exception of pipelines, the route solver module for all modes – highway, rail, water, air, parcel, and multimodal shipments – was based on the analytical logic and route solvers from off the shelf GIS software. Due to restrictions on use of the national pipeline network, GeoMiler calculates great circle distance (GCD) for all pipeline shipments.
- The mapping visualization interface was also based on off the shelf GIS software, which was customized to allow the assigned paths to be checked and corrected.

The use of GIS in GeoMiler allows the most likely paths taken by shipments from ZIP Code to ZIP Code to be displayed on a computer screen. GIS was used to create a series of programming and modeling routines that allowed for analysis of the individual freight transportation networks for roads, rail, waterways, and air cargo. More importantly, it allowed for analysis across all of these modes. Specifically, GeoMiler creates connectivity rules among the different networks, models the impedance factors based on cost descriptors and network hierarchies, and defines the routing logic that uses these impedances. Each link in the multimodal network is assigned an impedance value, which represents the likelihood of flow on each link. The higher the impedance value, the less likely the link will be selected for the path. By leveraging and customizing GIS technology, GeoMiler allows estimation of the optimal paths for all modes and allows each path to be displayed for analysis or decision making.

Developing Multimodal Spatial Network Data

GeoMiler uses a series of spatially referenced individual transportation networks (highway, rail, waterway, and airway) linked at points (airports, seaports,

and intermodal transfer facilities) with straight line access and egress connections that we call “spatial joins.”¹ This approach makes it unnecessary to merge the individual mode-specific networks into one large integrated multimodal network. Keeping the modal networks separate and connecting them with spatial joins allow distance estimation for both single-mode shipments and any combination of multimodal shipments from any ZIP Code to any ZIP Code.

It also allows calculations for export shipments by any mode to Canada and Mexico as well as by air freight and maritime vessel to all U.S. overseas trading partners.

This spatial network database includes ZIP Code centroids, the U.S. highway network, Canadian and Mexican highway and rail networks, national rail network, national waterway and docks network (including inland, coastal, and Great Lakes), global seaway network, intermodal truck-rail-ports terminals, and North American border crossing locations. BTS constructed this geospatial multimodal network from the following sources:

- **Highways and Roads** – Highway route data from a commercial source provided coverage of roads in the United States. Additional sources were used for Canada and Mexico.
- **Railroads** – Federal Railroad Administration (FRA) maintains the North American rail network used in the Rail Waybill data (enhanced with Waybill ownership, trackage rights, and rail densities).
- **Airports and Airways** – Federal Aviation Administration sources provided location information for large, medium, and small hubs that handle freight. For Alaska, it included all hub and non-hub airports. A domestic and international airway routes network was created from official air cargo data from the BTS Office of Airline Information (OAI).

¹ Transportation facilities (e.g., airports, highway/rail transfer yards) can be extensive and contain many miles of roads or rail. Rather than navigating these labyrinths of rail and asphalt when connecting two or more different freight transportation modes, which is cumbersome and time consuming, GeoMiler represents each of these facilities as a single point. The “shortest path solvers” function in GeoMiler then uses straight-line links, called “spatial joins,” to connect these points to the appropriate transportation mode(s).

- **Waterways** – Routable network maintained by the U.S. Army Corps of Engineers (USACE) and distributed by BTS.
- **Water Ports** – USACE sources, including dock-to-dock shipment location and commodity information for all docks on the inland waterways and seaports.
- **Intermodal Facilities** – database created and maintained by BTS.
- **Zip Code Centroids** – United States Postal Service and a commercial source.

Complementary Input Data

Working together through interagency cooperation, BTS received official freight data, including information not available for public use, from the FRA and the USACE to support the modeling, validation, and selection of the optimal paths. The major primary complementary data sources included:

- FRA Waybill Sample Data – used to determine the link densities, interlining, and rail station use;
- USACE waterborne traffic data – used to determine the most likely docks by commodity for multimodal shipments involving the U.S. waterway system;
- Census Bureau U.S.-international trade data – used to select and validate the most likely air and sea routes for all trading partners, including oceanborne shipments to landlocked countries;
- BTS OAI air cargo data – used to control cargo traffic among airport pairs for both domestic traffic and exports to all countries;
- BTS TransBorder and Border Crossing data – used to select the most likely crossing points for exports into Canada and Mexico by truck and rail.

Modeling Simulated Pathfinder Routes

GeoMiler's basic modeling routines are straight forward. With data on specific origins and destinations and mode sequence, the tool determines the set of spatial links and nodes that make up the "best possible" or optimal route between the two locations (figure 2). This selected route simulates a typical likely

path within the multimodal network that optimizes a set of pre-defined conditions that are mode and commodity specific.

For the highway (both for-hire and private) and water (inland shallow draft, deep draft, and Great Lakes) components, the tool finds the least-impedance path from origin point to destination point over the highway and waterway networks and then sums the lengths of individual links on these paths.

For rail and rail-inclusive intermodal, the tool estimates a least-impedance path based on railroad specific operating paths and selects the most likely route. Rail shipments are routed over the latest FRA rail network updated with the latest ownership changes, abandonments, trackage rights, interlining, and track density.

For air shipments, GeoMiler finds the least-impedance truck routes to the origin airport and from the destination airport. It then uses airport-to-airport traffic availability to estimate the most likely route, giving priority to direct flights between major hub airports, single carrier routes (circuitry constrained), and maximum air cargo lift from airports, while controlling for the ratio between the truck mileage and the air mileage.

Because air traffic does not physically cross the U.S. border at designated gateways, as do the other modes such as rail, GeoMiler measures the U.S. portion of an international flight by extracting that portion of the flight that extends from the U.S. airport of exit to the point where the flight path intersects the U.S. border.

GeoMiler Deployed

As deployed, the GIS-based GeoMiler uses current generation programming languages with standard commercial GIS software to provide a fast and efficient data processing module with the capability to estimate and verify shipment mileages. It offers a dynamic modal routing tool able to process data and generate shipment distances for single modes and multimodal shipments.

It also offers a very user-friendly interface with a consistent diagnostic tool for reconciling records with problematic origin and destination state and ZIP

Code information. New route viewing capability, including multimodal routes and information on choice of intermodal facilities, allows user-analysts to make informed choices when the input modal information needs correction.

While GeoMiler was developed for the CFS, the approach used in creating the geospatial databases and modeling the multimodal movements could be leveraged for other applications.

Figure 2. Illustration of an Intermodal Freight Shipment Path

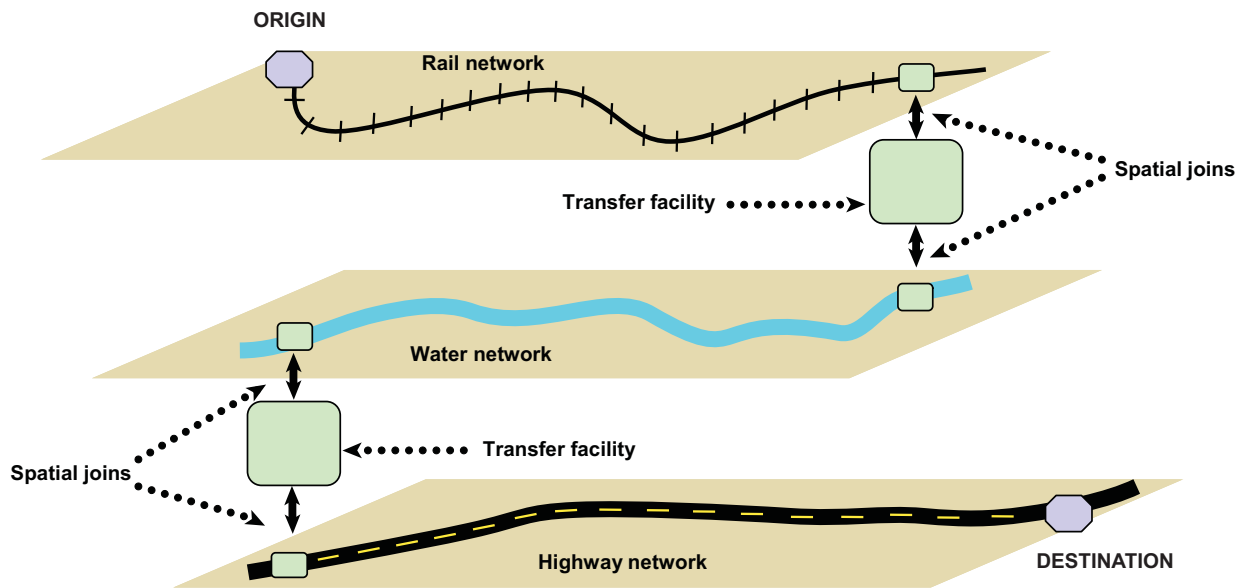


Figure 2 illustrates the routing of a multimodal shipment from origin to destination. In this example, the shipment leaves the true origin by rail, travels to a transfer facility where the cargo is transferred to a vessel, then travels by water to another transfer facility close to the destination, from where it is transferred to truck and travels on the highway system to the shipment's true destination.

SOURCE: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics.

Figure 3. Example of a GeoMiler Path Determination

Great Lakes -> Inland Water -> Truck: Chicago to Haines City, FL

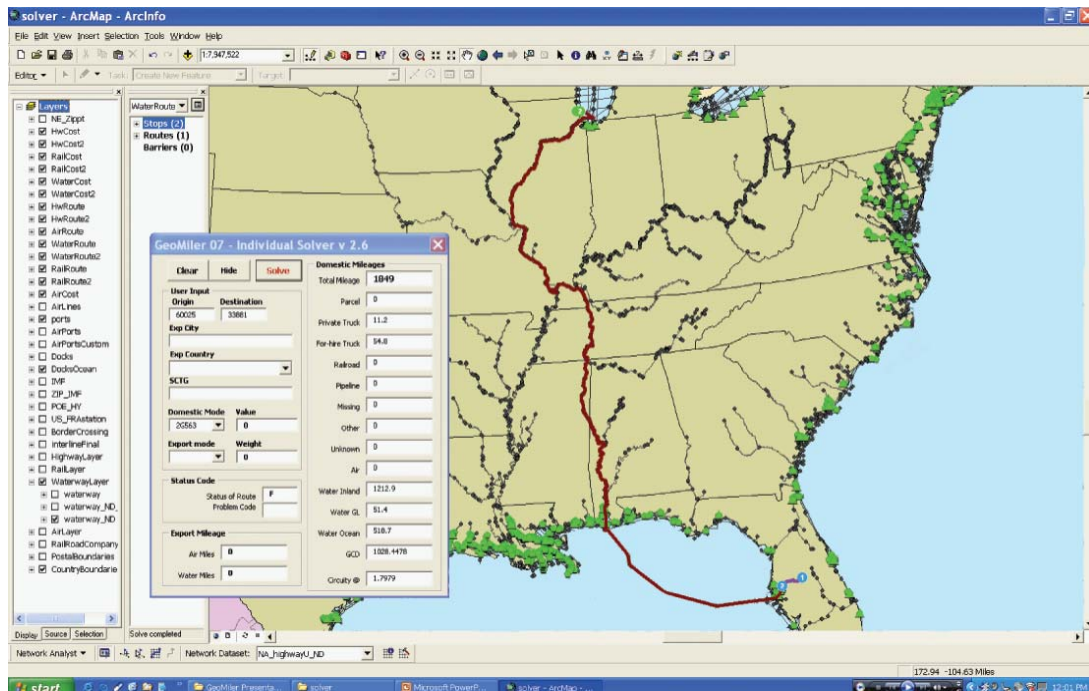


Figure 3 is a snapshot view from the mapping interface. In this case, the input origin is Chicago, Illinois and the destination is Haines City, Florida. The input mode sequence is water followed by highway. Using this information, GeoMiler calculates a path in which the shipment leaves Chicago, travels by inland waterways through the port of Mobile, Alabama, across the Gulf of Mexico to a port in Florida, and then continues by truck to its final destination in Haines City, Florida, traveling a distance of 1,849 miles.

SOURCE: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics.

Acknowledgments

A BTS contract team under the management of Stephen Lewis and Michael Margreta developed GeoMiler. The software developing, programming, and data expertise of Adhi Dipo, Fahim Mohamed, Raquel Wright, and Derald Dudley were critical. The freight modeling and CFS insights were provided by Felix Ammah-Tagoe. The Federal Railroad Administration and the Waterborne Commerce Statistics Center of the Army Corps of Engineers provided critical complementary data and technical support that made this tool possible.

• • • •

The Bureau of Transportation Statistics is a component of USDOT's Research and Innovative Technology Administration.

For questions about this or other BTS reports, call 1-800-853-1351, email answers@bts.gov or visit www.bts.gov.

For related data and reports visit www.bts.gov

Data -

- **NTAD**—National Transportation Atlas Database
- **Commodity Flow Survey**—value, weight, and ton-miles by commodity, mode, and origins and destinations
- **Transborder Freight Data**—monthly release of U.S.-Canada and U.S.-Mexico trade data
- **Border Crossing/Entry Data**—monthly data on incoming vehicle crossings from Canada and Mexico

Reports -

- *America's Container Ports: Delivering the Goods*
- *North American Freight Transportation 2006*
- *Freight in America*
- *America's Freight Transportation Gateways 2004*