Intermodal Freight GIS Network Risk Assessment – Final Report

Introduction

Previous Center for Intermodal Freight Transportation Studies (CIFTS) projects funded the development of an intermodal freight transportation geographic information system (GIS) network. This network contains the highway, rail and waterway modes of transportation, originating from the Bureau of Transportation Statistics’ (BTS) National Transportation Atlas Databases (NTAD). The modes are connected at major intermodal container terminals, verified using satellite imagery and aerial photography from Google Earth and Microsoft Virtual Earth, respectively. The network contains rules and adjustable transfer penalties that govern permissible modal combinations and transfer times at terminals, respectively. Container flows may be loaded onto the network (multiple origins to multiple destinations) and the resulting shortest path or least travel time paths will be displayed.

In January 2009, a new project was undertaken, “Intermodal Freight GIS Network Risk Assessment,” under the Center for Advanced Intermodal Technologies (CAIT). The objective of this project was to add risk attributes to the network so that minimum-risk routing results could be compared with baseline traffic volumes and shortest time results. The risk attributes included population within 5 miles, safety performance (i.e., accident rates for every link) and proximity to emergency services of each network link. This document describes the effort and the results of the research project.

Data

As mentioned in the introduction, the network consists of BTS NTAD networks merged at verified intermodal container terminals. This network contains over 346,000 links and covers the entire United States (intermodal connections are only made in the lower 48 states).

The incident databases originated from several different sources. Each mode of transportation is listed below along with its corresponding data source(s).

Highway – two truck crash data sources were investigated during the project, the Motor Carrier Management Information System (MCMIS) maintained by the Federal Motor Carrier Safety Administration (FMCSA)\(^1\) and the Fatality Accident Reporting System (FARS) maintained by the National

\(^1\) [http://mcmiscatalog.fmcsa.dot.gov/](http://mcmiscatalog.fmcsa.dot.gov/)
Highway Traffic Safety Administration (NHTSA). MCMIS contains accidents involving trucks and buses in the United States and cost $209 to order from FMCSA. There were more than 1.8 million accidents contained in the database from 1993 through April 2009. FARS data contains all accidents where a fatality occurred and using the vehicle configuration field, those not involving trucks can be removed from the database.

Rail – The Federal Rail Administration’s (FRA) Office of Safety Analysis maintains rail network accident data. The dataset analyzed for inclusion in the intermodal network was the equipment accident database, which contains accidents occurring between 1975 through 2008. The accidents in this database include derailments, collisions, fires, explosions, and other events involving the operation of on-track rail equipment.

Waterway – The U.S. Coast Guard is responsible for collecting and maintaining waterborne vessel casualty data. The U.S. Army Corps of Engineers Waterborne Commerce Statistics Center (WCSC) disseminates this data. However, since the focus of the network is on intermodal containerized shipments, waterway accidents were not analyzed during the research. There is not a significant amount of containerized intermodal activity on inland waterways, and deepwater accidents occur before intermodal containers enter the network. Dobbins and Abkowitz have performed an intensive review of USCG waterway accident data using GIS in the Journal of Transportation Safety and Security (Dobbins and Abkowitz, accepted and awaiting publication: 2010).

Emergency Response – The National Fire Department Census served as the database of emergency responders. This nationwide database contains basic information such as address, department type, manning levels at each department and number of stations. The fire station database was divided into two layers based on station type (whether the station was mostly/all volunteer firefighters or mostly/all career firefighters). This was done since the two types typically have different response times, training, manning levels, equipment and hours. The distance of each network link to the closest responder type within 100 miles was computed for all network links.

---

2 http://www-fars.nhtsa.dot.gov/
3 http://safetydata.fra.dot.gov/officeofsafety/
4 http://www.iwr.usace.army.mil/ndc/data/datauscg.htm
5 http://www.usfa.dhs.gov/applications/census/
Population - The population data was developed by the U.S. Census Bureau and included census tracts and census block groups, which are polygon features. GIS routines were used to create centroid layers for tracts and block groups, which were then used to determine the average population density and total population within 5 miles of every network link (“sealanes” and “open water” waterway links were excluded from the population analysis).

Methodology

A fundamental aim of this research was to compute accident rates for every segment of the network. Had all modes’ accident data been easily (and accurately) geocoded and the number of trips consistently reported or ascertainable in the national network datasets, this would have been a straightforward task. Ideally, each accident record would have a geographic coordinate pair (latitude and longitude) that placed it near the appropriate network link. The accident could be tagged with the ID number of the nearest network link, and the number of accidents occurring on or near each network link could be computed. This number of accidents divided by the number of trips, trains, ton-miles or other modal metrics would constitute the accident rates. Unfortunately, this methodology could not be completed due to limitations with the data.

An additional objective of the project was to determine the population and population density within 5 miles of every network link. By buffering each network link by 5 miles, it was hoped that GIS routines could determine the resulting polygon intersection percentages (between Census block groups) and estimate the population within each network link buffer. However, this methodology proved too computationally intense for the GIS to handle on a national, or even a regional, scale. Instead, centroids of the Census tract polygon features were created and the resulting “point in polygon” test (i.e., add up the population of all centroids within each network link buffer) served as the population estimate for the network.

As mentioned previously, the distance from each network link to the nearest responder within 100 miles was computed. There are four fields that contain this information: nearcareerID (ID number of the nearest career fire station), nearcareerdist (distance in miles to the nearest career fire station), nearvolID (ID number of the nearest volunteer fire station), and nearvoldist (distance to the nearest volunteer fire station). This computation was easily computed using spatial join routines that related the nearest features’ straightline distances. For map presentation purposes, all of the career fire stations

http://factfinder.census.gov/servlet/DownloadDatasetServlet?_lang=en
were buffered by 30 miles to see which areas of the U.S. (and intermodal network) that do not have an emergency response team within 30 miles.

**Results**

Population – this analysis was designed to allow the network to determine minimal population exposure routing solutions. By adding population characteristics within 5 miles as attributes, corridor risk assessments may be performed and shortest path solutions may be compared with minimal-exposure solutions. Five miles was chosen as the analysis distance due to most airborne toxic chemicals’ conservative isolation distances being within this range (as reported in the Emergency Response Guidebook 2008). 58,000+ miles of the network (all modes), or 7% of the total network mileage, is located in the top 20% of population densities (3,000 people per square mile). These areas are illustrated in the following figure.

*Figure 1. Areas with population density greater than 3,000 persons/square mile (preliminary Census data, 2007)*
For every network feature, 3 attributes were populated: pop_dens_07 (average population density within 5 miles), pop_estimate_07 (total population estimate within 5 miles) and pop_per_mi_07 (population per mile, or pop_estimate_07/length).

Rail – the accident records downloaded from the FRA’s Office of Safety Analysis website\(^7\) did not have populated latitude and longitude attributes. However, the railroad and milepost attributes contained in the accident dataset were populated. It was hoped that the latitude and longitude from records in the FRA milepost database (obtained from FRA) could be joined to the accident table and mapped. Unfortunately, the milepost layer is not unique in each county. For example, there are thirteen mile markers in Cook County (Chicago) with a railroad value of “Union Pacific” and mile marker value of 9.0. For this reason, it was only possible to compute accident figures on a county-wide basis.

Each county contains a count of rail accidents occurring between 1975 and 2008. In addition, the number of miles of class 1 and total miles of railroad track were computed. Not surprisingly, Cook County in Illinois (Chicago) leads all counties with 4,047 rail equipment incidents. Each rail network link contains a density range of rail traffic, expressed in million gross ton-miles per mile. Translation of these ranges into trips (trains per day) for use in an accident rate calculation was not attempted since the resulting rates would vary dramatically.

![Figure 2. County rail incidents per mile per year (minimum 25 incidents 1986-2008)](image)

---

Highway – Two databases were analyzed in performing the highway mode risk assessments: MCMIS and FARS. Again, the MCMIS database contains all incidents involving trucks, not just fatality accidents like FARS (which contains all accidents, not just those involving trucks). The MCMIS data could not be geocoded since there were no latitude and longitude fields, nor was there a linear reference (route and milepost). Another difficulty was the location and route fields. The names of highways and interstates were not standard (e.g., Interstate 40 might be denoted as “I40,” “I-40,” “I-40E,” “I-40W,” or “I 40”). A significant amount of time was spent standardizing these route names. Since the MCMIS route data was so inconsistent, crash rates were not calculated using this dataset. Almost 10% of the crash records had no route information whatsoever and a significant yet unknown percentage had the incorrect route information judging from the varying codes and impossible Interstate route/state combinations. For example, in addition to the existing Tennessee Interstates 24, 26, 40, 55, 65, 75, 81, 140, 155, 181, 240, 275, 440, and 640, twenty-four (24) non-existent Interstate numbers were also listed in the MCMIS database in Tennessee. Given this caveat, it was possible to determine for each county and Interstate name the number of incidents, as shown in the map below. One of the problems with the MCMIS dataset is that each state reports their own accidents in a non-standard format. Note the state of Texas and how counts were not able to be displayed for anything but San Antonio and a small stretch between Austin and Dallas. Texas’ use of a non-standard code for routes outside urban areas caused this anomaly.

![Figure 3. Map of Interstates with MCMIS counts of Truck-related Crashes (Jan 93 – Apr 09)](image)

It was much easier to use the FARS dataset in a GIS. Latitude and longitude are available for nearly all accidents from 2001 to the present. The geographic coordinates are very accurate as well. Assuming
constant truck VMT, the trucking fatality rate (fatal accidents per ten million truck miles) is presented below on a segment by segment basis in Tennessee (for segments with more than 1,000 trucks per day).

Figure 4. Fatality rates involving trucks in Tennessee (fatality accidents per 10 million truck miles traveled).

The most remarkable analysis occurred with the FARS data from 2001-2008 using the linear referencing system capabilities of the GIS. Linear referencing refers to the ability of a GIS to map objects (points or lines) according to route and milepost data. The GIS is also capable of determining for a point object (already mapped by latitude and longitude, for example) its route and milepost. Every FARS fatal accident was assigned a route and milepost using the Intermodal network (highway links). Then, the number of fatal accidents (involving trucks between 2001 and 2008) were counted for every 1-mile segment in the U.S. The segment with the most frequent number of fatal crashes involving trucks is Highway 23 (mile marker 14) in Lawrence County, Kentucky with 6 separate fatal crashes involving trucks since 2001. The following maps show the 1-mile segments and the counts of fatality accidents involving trucks between 2001 and 2008 for the entire U.S. and Tennessee.
The FARS dataset goes back to 1979, however, latitude and longitude fields were not populated significantly until 2001. For years 1995-2000, a count of fatal crashes was performed by county (similar to the rail accident analysis) and these counts were added to the 2001-2008 counts. This map is presented below.
Emergency Response – The Fire Census Dataset was geocoded using street address data and 2 layers were created (one for career/mostly career fire stations and another for volunteer/mostly volunteer fire stations). A buffer of 30 miles was created around each career station (red areas) and one can see areas where the network has only volunteer fire department coverage.
Of particular interest in Tennessee, there is no career fire department coverage (within 30 miles) for more than 17 miles of Interstate 40, including the bridge over the Tennessee River. There are 4 volunteer fire departments (red squares) in the vicinity of this important bridge, however, it is not apparent what kind of equipment and training these 64 volunteer fire fighters have if a major closure of Interstate 40 were to occur near or on the bridge. A future research topic that could be undertaken by IFTI is an analysis of vulnerable infrastructure such as this bridge in terms of responder coverage and detour lengths.
Future Directions

While the full objective of this project (accident rates, population exposure and emergency responder coverage) was not completely met, several contributions were made and some additional research topics could be explored.

This research presents a screening-level assessment of modal accident and operational databases for use in a national network environment. A host of issues with the data, specifically those precluding the use of data inside a GIS, have been identified with rail and highway datasets. The FARS database is the model which other modes should emulate. This database has accurate geographic coordinates, and as much detail as any user would desire. For example, the fatal crashes analysis in FARS could have been restricted to truck makes/models, truck body type (e.g., single unit trucks, 3 axles, 4 axles, etc.), time of day, even the height and weight of the truck driver, among others. For the other datasets used in this
effort, county accident counts and rates are interesting, but are not suitable for inclusion in a network model.

Some future research topics have been identified through the conduct of this research. Freight bottlenecks have been identified in numerous studies. These locations could be analyzed to determine the contribution of truck-related incidents to congestion and whether other sources contribute to delays (e.g., urban revitalization, etc.).

As identified in this project, network vulnerability and resiliency is an important issue, which GIS tools are particularly adept at addressing. Emergency responder databases are of sufficient quality to embark on a project analyzing coverage of key infrastructure elements of our nation’s freight network. An improvement on this research would be to route the responders to each network link through the local street network. This would constitute a significant improvement over the “as the bird flies” distances computed in this project. Such a vulnerability analysis would also include detour lengths, alternate routes and the time tradeoff associated with a significant disabling of a bridge, terminal or port area. However, as this research has pointed out, it would be best to restrict the analysis area to a region or even a specific port. Operational and safety data issues can hinder the efforts at developing a methodology. For example, individual MCMIS reports and police accident reports are used extensively by researchers to examine truck accidents.

This research has demonstrated the issues involved with each mode’s primary accident data source. Simple database routines can be performed to trap some of these errors. The “non-existent” Interstate route numbers in a given state is one such easily resolved issue. Database routines can make use of a table that lists Interstate routes for a given state. An error report (or message) could be automatically generated if a user attempts to record an invalid Interstate number (e.g., “I-4” in Tennessee, as an example).

One interesting result of the research was that on a square mile basis, as computed with the three-mile nationwide grid, Manhattan, New York had the 2 highest numbers of fatal crashes involving trucks (using FARS data). It would be interesting to look at trends across U.S. urban areas to determine if these accidents involve a high proportion of single unit (delivery) trucks and their interactions with pedestrians.