Incorporating Disaggregate Land-Use Characteristics in Trip and Tour-Generation Models

Sivaramakrishnan Srinivasan
Russell Provost
Ruth Steiner
University of Florida, Gainesville, FL

This study develops household-level, trip- and tour-generation models incorporating land-use characteristics around the residences as explanatory factors (in addition to other socio-economic-mobility variables of interest). It is useful to note here that we describe residential-location characteristics at three different spatial levels: (1) the characteristics of the land-parcel (example, land area and building square footage in the parcel), (2) the characteristics of the "neighborhood" in which the parcel is located (% residential, commercial square footage, etc.), and (3) the location of the neighborhood within the region (such as the distance to regional activity centers).

Travel-survey data from the 1999 Southeast Florida Regional Travel Characteristics Study was supplemented with parcel-level land-use characteristics obtained from the Florida Department of Revenue (FDOR) and roadway-network data from Florida Department of Transportation (FDOT) and used in the development of the models. In this rest of this document, we present an overview of the data-assembly procedure, focusing on the land-use data. The model estimations (ordered-probit models) will be undertaken shortly.

In the 1999 Southeast Florida Regional Travel Characteristics Study, one-day travel information was collected from about 5,000 households in Miami-Dade, Broward, and Palm Beach Counties. The following information was collected for each trip undertaken by each respondent: trip timing (start- and end- times), mode (including occupancy for auto mode), purpose, and trip-end locations (addresses). Subsequently, the trip-end addresses were geocoded to determine the latitudes and longitudes of the trip ends. The total number of person-trips and person-tours generated by each household was obtained by suitable aggregation. A tour is defined as a sequence of trips such that home represents the origin of the first trip and the destination of the last trip with the destination of all intermediate trips being non-home locations. As the trip-end locations were geo-coded, the residential location of each household is known at a fine spatial resolution.

Transportation system characteristics were obtained from a detailed statewide roadway network file entitled "Dynamap Streets" made available on the FSUTMS GIS web portal (http://www.fsutmsonline.net/index.php?/gisonline/). The data file is highly detailed and includes roadways under the purview of all jurisdictions (federal, state, and local). The acquisition of a detailed statewide dataset allowed for the seamless roadway network analysis across county lines. It is useful to note that the roadway data represents year 2005. The travel-survey data are, however, from the year 1999. As information on the year in which each roadway segment was built was not provided in the attribute table, it was not possible to adjust the roadway network to reflect 1999 conditions.

As already mentioned, the primary source of land-use information is the Florida Department of Revenue (FDOR). Historically, parcel-level land use data have been recorded in paper format individually by county property appraisers for tax record-keeping purposes. In recent years this information has become available in data formats compatible with GIS. The FDOR provides access to parcel data from a majority of the counties within Florida primarily in GIS formats via a public FTP site. These parcel data files contain valuable information for our analysis including the landuse type, area of buildings, and the number of residential units located on each parcel.

The parcel data obtained reflect the land-use characteristics in the year 2008. However, the travel survey data are from 1999. In order to ensure consistency, developments built after 2000 were removed from the data. For this purpose, the research team obtained aggregated parcel data that detailed the decade in which individual parcels' structures were built. This file was overlaid on the detailed parcel datasets obtained from FDOR to identify and exclude parcels built after 2000 from the analysis. It is useful to note that FDOR parcel datasets contain an attribute that lists the year in which the parcel became "effective". However, using this information alone for determining the year a development was built can be misleading. This is because the effective date can also reflect modifications to an existing structure. For example, if a residential unit originally built in 1950 underwent a modification in 2002 that required a permit, 2002 would potentially be recorded as the "effective date".

Once the parcel data from 2008 were adjusted to reflect 1999 conditions, they were subject to further cleaning and updating. A few of these efforts are briefly outlined here.

- The data on the number of residential units were not available in the parcel files for mobile-home parks. Typically lots within mobile home parks are not owned by the individual living in the structure itself, but instead by a person or entity holding ownership over the entire park. Researchers obtained an additional dataset (www.MyFlorida.com) that details the location of registered mobile home parks across the state. The locations were geocoded and converted in a dataset compatible with GIS. Within the dataset the number of lots within each mobile home park is recorded. Parcels categorized as mobile homes were overlaid onto these data points to obtain the number of lots within each parcel. The number of residential units (previously zero) was recalculated to reflect the data.
- A cross tabulation of the parcel's land use against the presence of residential units revealed that several parcels with non residential land uses contained "residential units". Although it is conceivable that some parcels may contain residential units as a secondary use, land uses such as storage unit facilities and gas stations were shown to have at least one residential unit. Upon closer examination it was found that certain counties utilized the residential unit attribute to record all types of "units" located on the parcel. These units can represent a number of items depending on the particular landuse. For example, a storage unit parcel may have been recorded as having 48 "residential units" when in fact this represents the number of storage units located on the parcel. Parcels illogically containing residential units were edited to indicate the presence of zero residential units.
- The digitization of parcels from a paper format to a digital format can incur human error. Erroneous polygons can occur when errors in digitization create slight overlaps in the boundaries of parcels. This overlap creates a smaller polygon that often receives the attributes of the parent parcel. This can cause duplication in the parcel dataset and cause

erroneous results and findings. To remove the erroneous polygons, the GIS analyst was used to dissolve parcels based on their unique ID thereby merging these smaller polygons with their parent polygons.

In summary, substantial processing was undertaken to ensure that the data are clean and internally consistent. The cleaned parcel-level file created contains the following descriptive for each parcel in the three-county region: (1) a parcel identifier, (2) parcel area, (3) aggregate and disaggregate land use classifications, (4) number of residential units for residential parcels, and (5) building square footage. The disaggregate land-use classification scheme provided by the FDOR has 99 categories. These were aggregated to create a more manageable and practically useful land-use classification scheme. The major land-use categories in the aggregate scheme are residential (single-family, multi-family, mobile homes), retail-commercial (large retail, regular retail, convenience stores, drive-through), office (professional and non-professional services buildings), institutional (hospitals, colleges, etc.), industrial (light, heavy, warehousing), and other.

In the next step, “neighborhoods” were created and these were characterized by aggregating the data from the cleaned parcel-level files and the transportation-network files. In this study, neighborhoods are defined as grid-cells of size four square miles. To generate these neighborhoods, a four-square-mile grid was arbitrarily imposed on the study region (three-county region). Each cell in this grid was assigned a unique “neighborhood identifier” value. The entire grid was then shifted horizontally by one mile. The new grid-cells were assigned “identifier” values. In the third step, the grid from the previous step was shifted vertically by one mile and the new grid-cells were assigned new identifiers. Effectively, this procedure creates a set of overlapping neighborhoods of size four-square miles across the entire region with the centroids of these neighborhoods lying on a one-square mile grid (See Figure 1).

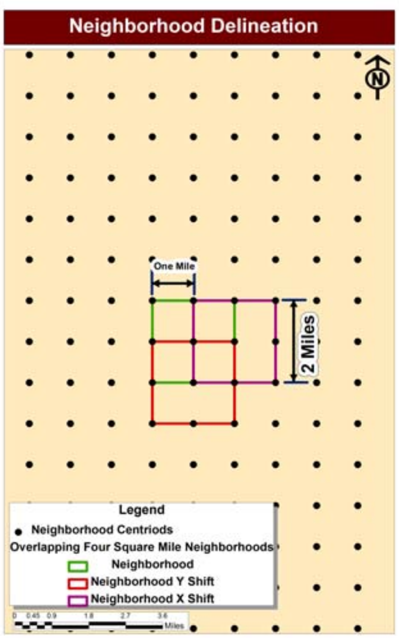

Figure 1 Neighborhood Delineation

The characteristics of the neighborhood were determined by aggregating the parcel-level data. Specifically, each parcel was assigned to one or more neighborhoods if the parcel centroid fell within the corresponding boundaries (note that each parcel can be assigned to multiple neighborhoods as the neighborhood areas themselves overlap). Once the land-parcels were mapped to the neighborhoods, the land-use descriptors of the neighborhood (such as land area under different land uses, building square footage under different uses, etc) could be obtained by spatial aggregation. Similarly, the roadway network data were also aggregated to determine measures such as linear road miles, number of intersections, and number of cul-de-sacs within each neighborhood.

As discussed in the section on methodology, the intent of this study is to capture land-use at three spatial levels: parcel, neighborhood, and regional. The developed of parcel- and neighborhood characteristics have been discussed thus far. Next we present the approach to describe the location of each neighborhood within the overall region. Specifically, three measures are used for this purpose.

- The first measure determines whether each of the neighborhoods is located within the boundaries of one of the four major cities in the region (Miami, Ft. Lauderdale, Boca Raton, and Palm Beach). This was determined by a spatial overlay of the neighborhood centroids with the city boundaries.
- The second measure determines the network distance of each neighborhood to each of four regional activity centers. The activity centers (one in each of the four major cities) were defined as neighborhoods with the highest commercial square footage (includes, retail, office, and entertainment). The distances were determined between the neighborhood centroids along the roadway network.
- The third measure determines the network distance of each neighborhood to each of ten regional residential centers. Residential centers were defined as neighborhoods having over 80 percent of its land use dedicated to residential. Among these, ten neighborhoods (geographically distributed across the study region) that had the largest number of residential units were selected as the regional residential centers. The distances were determined between the neighborhood centroids along the roadway network.

Once all the land-use variables at the parcel- neighborhood- and regional- levels were created, these were suitably linked to the residential locations from the household-travel survey data file. Each location was matched to a specific land parcel based on the proximity to the parcel centroid (ensuring that the land-use type of the parcel is “residential”). Each residential location was also matched to a neighborhood based on the proximity of the location to the centroid.

The efforts thus far have resulted in the development of several descriptives of residential land-use. Subsequent research will focus on the development of ordered-probit models for trip- and tour-generation incorporating these and other variables.