Running head: PSRC URBANSIM TRANSPORTATION INDICATORS

Transportation Indicators

Emad Dawwas
Yegor Malinovskiy
Susan Radke-Sproull
University of Washington

In partial fulfillment of the requirements for URBDP 598E / CEE 599M / CSE 481E

Alan Borning
Paul Waddell

June 5, 2007

FINAL SUBMISSION

Abstract

The purpose of this paper is to report on a project whose goal was to propose and implement a set of new transportation indicators for use in UrbanSim that are derived from transportation network link-level data generated out of the trip assignment step of the Puget Sound Regional Council's (PSRC) travel demand modeling system (EMME/2). By associating land parcels with nearby nodes on the transportation network, data such as traffic volumes and congested travel speed are used to develop indicators of the utility of a particular parcel of land that can inform UrbanSim's Land Price and Location Choice Models.

Table of Contents

List of Figures

Transportation Indicators

"The Puget Sound Regional Council (PSRC) is an association of cities, towns, counties, ports, and state agencies that serves as a forum for developing policies and making decisions about regional growth and transportation issues in the four-county central Puget Sound region." [PSRC, 2007a] Among other functions, PSRC is responsible for land use and travel modeling for the Puget Sound region, as the basis for preparing long-range transportation plans and coordinating regional and local transportation strategies.

"UrbanSim is a software-based simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. It is intended for use by Metropolitan Planning Organizations and others needing to interface existing travel models with new land use forecasting and analysis capabilities." [PSRC, 2007c]

PSRC is actively integrating UrbanSim as the Land Use Forecasting component of their transportation modeling. In turn, UrbanSim's models – in particular, the Land Price and Location Choice Models – are informed by outputs of the transportation modeling process: The transportation network geography and transportation volumes may be useful for deriving indicators that are in turn evaluated as variables in these models.

The project goal was to propose and implement a set of new transportation indicators for use in UrbanSim that are derived from transportation network link-level data generated out of the trip assignment step of PSRC's travel demand modeling system (EMME/2).

Project objectives

In support of this goal, the following objectives were set forth:

- Understand…
    - EMME/2 & UrbanSim basics and interaction
    - Existing transportation indicators & their limitations
    - Available transportation network link-level data
- Learn to extract and manipulate…
    - Link-level data from EMME/2
- Propose…
    - And implement new accessibility indicators that can be implemented now using link-level data
    - And describe additional indicators requiring new data or further analysis

Guiding questions

To guide our work and the specification of new indicators, we identified the following sorts of questions as representative of the *type* of analysis our indicators should support:

- Is there a grocery store within a 15-minute drive of this parcel, and does this fact impact utility (as reflected in land price and/or the probability that a household would choose to live here)?
- Is it feasible to live at this location without a car, and does this fact impact utility?
- Does the volume of vehicle traffic adjacent to this parcel impact utility, and is this different for residential versus commercial land?

Possible uses of the indicators were identified to include:

- Visualizing patterns or differences based on different modes of travel or on different parcel characteristics across the Puget Sound region
- Inform the UrbanSim Land Price Model or Location Choice Model
- Compare results across different regional scenarios, e.g., comparing a 4-lane versus a 6-lane 520 bridge replacement

Structure and scope of this paper

The three project objectives are used as the basis for structuring this paper:

The first section provides a brief background of our understanding of EMME/2, UrbanSim, and their interaction. It discusses the data output from the assignment step of the travel demand modeling process and the use of these data to develop indicators, including limitations inherent in the data.

The second section describes the software architecture used to extract link-level data from EMME/2 and implement the indicators, including visualization. This section discusses aspects such as system design, test, and deployment.

The third section provides a formal definition of the implemented indicators, including documentation of their usage and limitations. This section also offers some ideas for future new indicators as well as refinements to those already implemented.

Finally, a conclusions section considers this project, and specifically the proposed indicators, in the context of the urban planning environment in which UrbanSim is used: social, political, and economic. Issues of professional and ethical responsibility are considered.

Understanding the Systems and the Data

This section sets the context for the implementation of indicators derived from transportation network link-level data generated from the trip assignment step of PSRC's travel demand model. It is not intended to be a complete description of the systems and data, but rather just enough to enable the reader to understand the indicators and why they were defined the way they were. After all, indicators are only as useful as the data used to construct them. To help the reader focus on the key points important to understanding our work, sentences and paragraphs are highlighted with a box. Of particular note are the highlighted statements in the section on Limitations in the Data.

PSRC's travel demand model

EMME/2, the travel demand modeling platform, projects traffic volumes for a target year as a function of projected changes in land use allocation (generated using a system with the acronym DRAM/EMPAL) that are, in turn, responses to economic forecasts (generated using a system called STEP). The following diagram provides a simplified overview of the systems involved. For more detail, refer to [Cambridge Systematics, 2006], as PSRC's travel demand model is well-documented there.

Of particular note relative to our project is the fact that each step in the process gets progressively finer-grained in its projections in terms of geography.

Economic forecasts are generated for each county in the 4-county region. Land use forecasts are generated for each of 219 Forecast Analysis Zones (FAZ) jointly agreed upon region-wide. Travel demand forecasts are finer-grained yet: land use forecasts are mapped from the 219 FAZ to 938 Traffic Analysis Zones (TAZ).

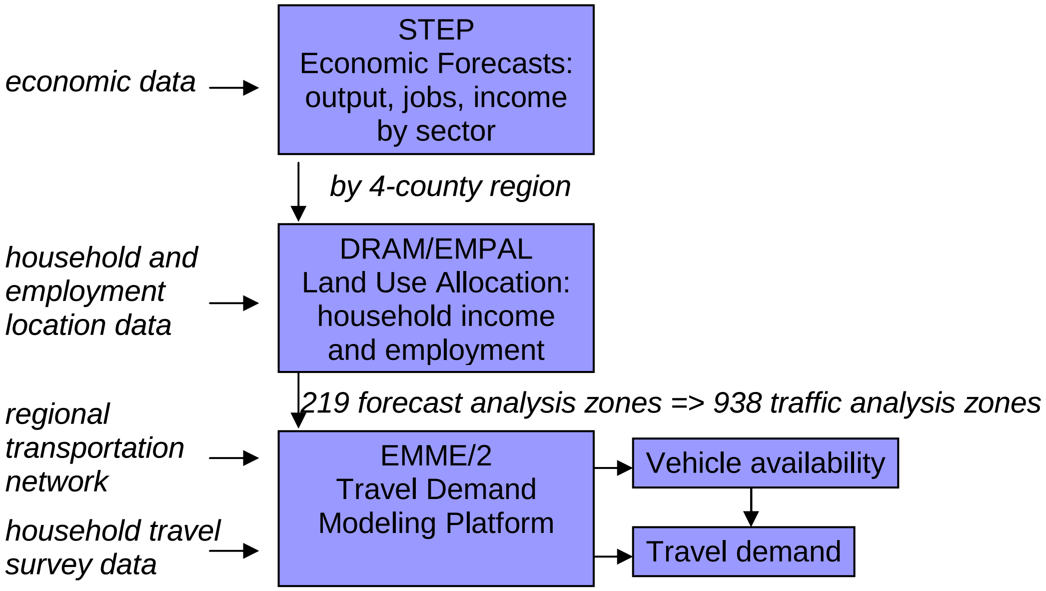


Figure 1: PSRC Travel Model Overview

The TAZ is the geographical unit of interest to EMME/2. Through a progression known in the transportation community as “the 4-step process,” EMME/2: 1) projects, for each identified trip purpose[1], the number of trips expected to be generated by (based on household characteristics) and attracted to each TAZ; 2) projects how those trips will be distributed (using the gravity model and other techniques) between a TAZ and one or more of the other TAZ; 3) projects what mode (based on household surveys, characteristics of the trip purpose, and other patterns) will be selected for each trip and (not formally a part of the traditional 4-step model) what time of day each trip will occur; and 4) after factoring in projected truck volumes (also not formally a part of the 4-step model but certainly a significant contributor to traffic congestion), assigns each of those trips, using a given mode at a given time of day, onto the individual links comprising our regional transportation network: highways/freeways, transit routes, bike routes, etc. This progression is shown in the figure below.

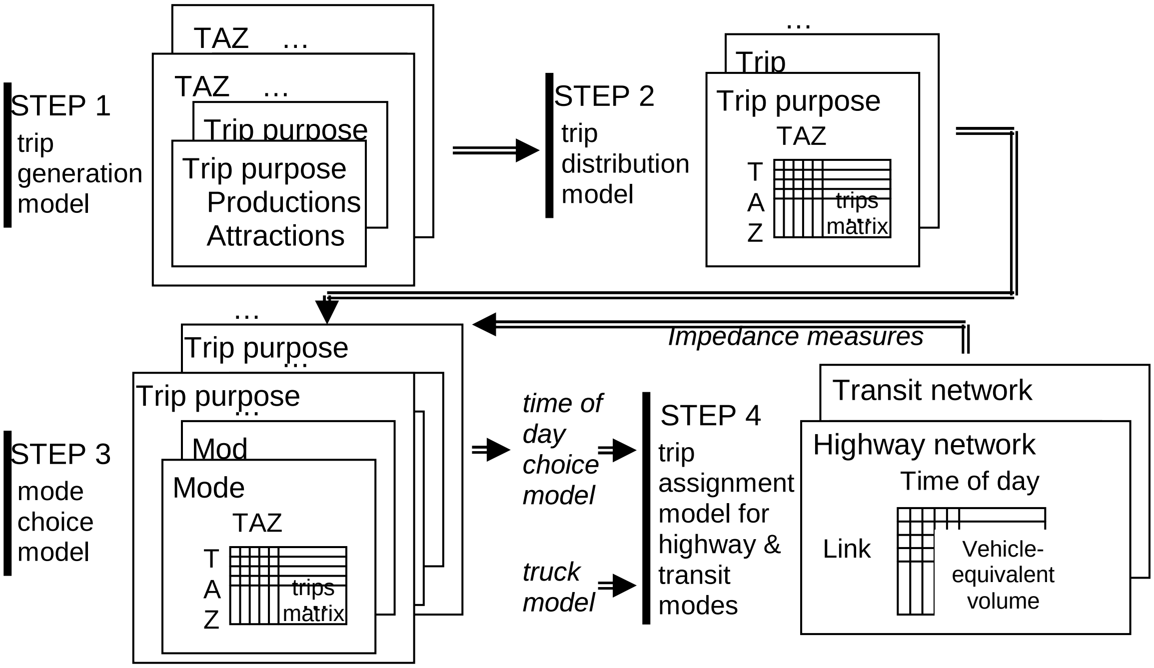


Figure 2: Overview of EMME/2 4-Step Model

It is the output of Step 4 in which we are particularly interested for this project.

The output of Step 4 includes projected traffic volumes, by mode, on each link of the region’s transportation network, by time of day. Once the trip assignments are generated, PSRC calculates several performance indicators for each link in the transportation network, such as congested travel time. The data we used for our project are discussed in more detail later.

---

[1] Each trip purpose has a single origin and destination, such as home-to-work or home-to-school. It does not capture the notion that a trip could be made up of a sequence of such origins and destinations. Thus, it is not possible to analyze the single key purpose in a sequence of purposes that might drive a mode choice decision. PSRC is working on a linked-trip model.

PSRC can compare different regional scenarios by changing elements of the transportation network that is input to EMME/2. For example, scenarios might explore the impact of an Alaska Way Viaduct replacement, or a 4-lane versus a 6-lane 520 bridge replacement.

EMME/2 and UrbanSim

UrbanSim is being integrated into the travel demand model as a replacement for DRAM/EMPAL: UrbanSim will be used to generate the land use allocation projections – households and employment – used by EMME/2 to estimate the number of trips the region is likely to generate given those projections.

One challenge (among many) is that the unit of geography modeled by UrbanSim is the "grid cell", an area of 150m x 150m… much finer grained than the TAZ used in EMME/2. From a data modeling perspective, a TAZ will typically encompass many grid cells, although a grid cell may lie partially in one TAZ and partially in another. At least initially, UrbanSim aggregates data (e.g., jobs, households), using a percentage allocation if necessary, from the grid cell to the TAZ to pass on to EMME/2.

UrbanSim is moving from the grid cell to modeling individual parcels of land… enabling simulation at the level at which decisions are actually made.

With respect to interfacing with EMME/2, the same aggregation process used to map grid cells to TAZ will apply to the parcel-to-TAZ mapping.

In addition to generating input to EMME/2, UrbanSim can also benefit from data produced by EMME/2 to inform its models. For example, land prices and location choices are often influenced by feasibility of various modes of travel, levels of congestion, and side-effects such as noise and emissions. As such, travel demand projections can be used to develop indicators that can be used as variables in UrbanSim's Land Price and Location Choice Models.

In fact, several indicators have already been developed in UrbanSim using the trip assignment data from EMME/2. However, these indicators are implemented using data that have been aggregated up to the TAZ. For example, rather than using the data available for each link in the transportation network, such as number of vehicles or congested travel time, the indicators use vehicle counts that have been aggregated up to the TAZ and congested travel times computed for travel between TAZ centroids. Of course, since UrbanSim actually works with grid cells, not TAZ, the data need to be disaggregated to grid cells, based on which TAZ the grid cell centroid falls within.

As might be obvious from this description, there are benefits to developing indicators, instead, directly from the link-level transportation data, without the interim steps of aggregation and dis-aggregation. By associating parcels with nearby nodes on the transportation network, interesting data such as traffic volumes and congested travel speeds can be understood as they impact land parcels and travel from one parcel to another. This opportunity is the motivation for this project.

Geography and the transportation network

In summary, the following diagram illustrates the various geographic semantics we are working with:

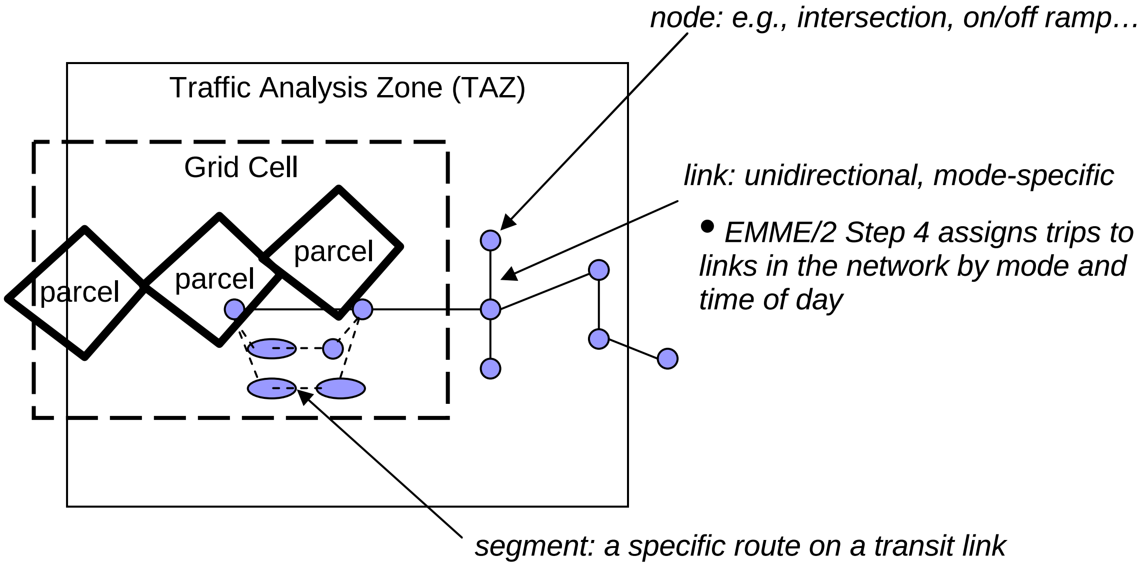


Figure 3: Units of Geography

- A model of the transportation network is an input to EMME/2. It is represented as nodes and links, each of which has a set of attributes that describe it.
  - Nodes: geographically-located points on the transportation network where trips may join or leave the network or shift to another part of the network, or where attributes of the network change (e.g., a shift from 2 lanes to 1)
  - Links: contiguous pieces of the transportation network that connect two nodes uni-directionally (a “from node” and a “to node”) and support one or more modes of travel; attributes include length, number of lanes, and an uncongested travel speed (typically, the posted speed limit)
  - Segments: when transit operates on the transportation network, there are additional attributes for each type of bus, ferry, or train that runs over a link and these are modeled as individual segments associated with a link
- EMME/2 models travel demand by and between Traffic Analysis Zones (TAZ). As the last step in the travel model, EMME/2 assigns trips generated to/from TAZ to the transportation network, as if trips originate and culminate at TAZ centroids. This assignment process generates additional information on each link, such as a volume of modal units (e.g., single-occupancy vehicles). Once the trip assignments are generated, PSRC calculates several performance indicators for each link in the transportation network, such as congested travel time.

- UrbanSim models land use at the level of grid cells, which are finer grained than TAZ: Typically, there are multiple grid cells in a TAZ, but a grid cell may span multiple TAZ.
- UrbanSim is moving to parcel-level modeling, simulating decisions of individual households and businesses. The transportation network passes by land parcels (although the transportation network itself is made up of land parcels, these constrain rather than participate in land use dynamics).

The association of a parcel with the transportation network enables us to ask questions of accessibility and traffic impact.

Relevant data and scenarios

A complete description of data used in the travel demand model and its encoding is available in [Cambridge Systematics, 2006] and [PSRC, 2007b]: We do not attempt to rival the completeness of these documents but rather seek to point out the key attributes relevant to our project.

For reference, Appendix A lists the network link attributes: a) the data available to EMME/2 about the transportation network, including possible values for the mode and the facility type attributes; b) the additional attributes generated as a result of the EMME/2 assignment step. Appendix B lists the transit segment attributes: a) the data available to EMME/2 about the transit routes, including the possible values for vehicle type, operating agency, and transit time function; b) the additional attributes generated as a result of the EMME/2 assignment step.

The transportation network assignment data produced by EMME/2 is organized by time period. A snapshot of the loaded transportation network is produced for each of:

- AM peak
- Mid day
- PM peak
- Evening
- Night

Thus, indicators based on transportation network data will be generated based on a selected snapshot and reflect conditions at a selected time of day.

Typically, indicators will also be mode-specific.

It will rarely make sense to ask about travel volumes that are an aggregation of counts of cars, trucks, bikes, and pedestrians, or shortest routes that combine congested travel speeds on general-purpose lanes and HOV lanes if you are interested in SOV travel. As such, indicators will generally specify a mode, such as (refer to Appendix A):

- All links valid for SOV travel (mode = s)
- All links valid for HOV3 travel (mode = a, where a is the encoding for “all road vehicles”)

- All links valid for HOV2 travel (mode = a and not mode = i, where i is the encoding for HOV3)

In the introductory paragraphs, we outlined a number of guiding questions for our project. In the context of link-level data, these questions compel us to consider mode, congested travel time – from a parcel to selected other parcels – and traffic volumes. Thus, in addition to time of day and mode, we focused on the following assignment and performance data available for each link (refer to Appendix A):

- Total vehicle volume on link (@tveh)
- Congested travel time (timeau)

From these, the scenarios we implemented, discussed in more detail in the remainder of this report, are:

- Path algorithm based on congested travel time (timeau) at a selected time of day
    - Shortest path, based on time using a specified mode, from a parcel to selected other parcels. Scenario: How long does it take as a SOV driver on the shortest path to a grocery store or CBD?
    - From a given parcel, how many of a particular type of parcel are within a selected number of minutes using a specified mode? Scenario: Are there grocery stores within 15 minutes for a SOV driver?

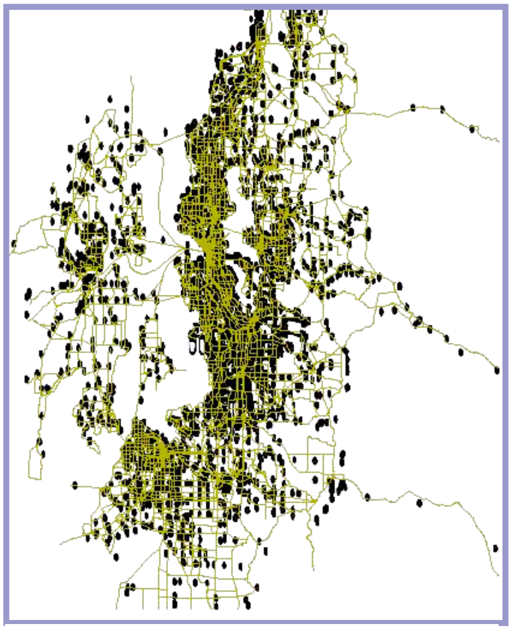


Figure 4: Transportation Network and Nodes

- Traffic impact algorithm based on vehicle volume (@tveh) at a selected time of day
    - Weighted impact (number of vehicles divided by the distance from the parcel centroid to its nearest node) of a particular mode of traffic approaching the nearest node to a selected parcel. Scenario: How impacted by rush-hour traffic is a particular residential parcel? a particular commercial parcel?

As mentioned earlier, these indicators may be incorporated in UrbanSim's Land Price or Location Choice Models as variables that may impact land price and/or location choice, respectively.

Limitations in the data

The previous sections glossed over details about the transportation network and link-level data that are important when defining indicators that are useful interpretations of accessibility, of the impact of congestion, and of the availability of alternative modes of transportation.

The most important limitations we uncovered for our work are summarized here.

- Facility types
  - In EMME/2, the transportation network *for vehicle assignment* reflects freeways and arterials only, not the local road network.

    Thus, when doing parcel-level queries about volumes of traffic passing through nearby links or about congested travel time, local traffic and local routes are not considered and, depending on how far a parcel is from the modeled network, may have a large local impact that will not be reflected in the indicators.
- Trip assignment
  - In EMME/2, trips occur between TAZ centroids and are assigned, using some distribution algorithm, to the transportation network.

    Thus, doing parcel-level queries about volumes of traffic passing through nearby links or about congested travel time between parcels may imply a false sense of precision in the data.
  - In EMME/2, intra-TAZ trips are not modeled at all.

    Thus, when doing parcel-level queries about volumes of traffic passing through nearby links or about congested travel time, there may be in reality additional traffic that is not included in the numbers used.
  - In EMME/2, buses are not modeled as vehicular trips, and are therefore not included in the trip assignments.

    Thus, traffic volumes and congested travel time do not take into account the addition of buses into the mix.

    PSRC believes that for freeways and arterials, the impact of a few additional buses is minimal, particularly because bus stops on these facility types tend to involve dedicated lanes and off/on ramps.

    It is possible that we could estimate the number of buses to add to the vehicle counts, e.g., by dividing the number of transit riders on a given segment (@voltr, see Appendix A) by the capacity of the type of bus for each segment on a link and computing some vehicle equivalent (e.g., a bus = 2) to add into the vehicle counts on the link. In fact, we practiced with aggregating segment-level data to the link level (see Appendix C). However, since the impact of these additional vehicles was not reflected in the calculated congested travel time on the link (timeau, see Appendix A), we believe this needs further thought.

- Transit
    - In addition to the earlier observations about transit not being reflected in link-level vehicular data, we also discovered that congested travel time by bus is not calculated and stored for each link, or even for each transit segment that runs over the link.

    Thus, the indicators we implemented are essentially not valid for bus transit mode (we did not investigate the semantics of the data for other forms of transit, such as rail or ferry). Or, put another way, bus transit must be assumed to have the same travel time as other traffic on a link, ignoring stops.

    One thing that makes this complex is that travel time may be different for each type of bus on a link, e.g., local versus express, so it's difficult to ask generalized questions about bus travel time.

    It might be possible to explore whether we could use the transit time functions (see Appendix B) to perform our own calculations. PSRC has said that congested travel time for transit is already calculated but it is stored in the underlying assignment matrices that are used to make assignments to the transportation network, and we did not explore those matrices.

- Alternative transportation
    - Although links are encoded for bicycle and walking modes, volume and travel time data are not available.

    Thus, the indicators we implemented are essentially not valid for bicycle or walk modes.

    We did explore the data to some level. For example, there appear to be no bike trails in the data for 2001 that we were looking at, and it appears that bike lanes are assumed for no freeways and for all arterials. If bicycles are assigned to the transportation network by EMME/2 at all (and it appears they might not be), it might be possible to calculate a travel time that simply multiplies link length by some assumed average cycling time.

These limitations have the following implications for our work:

- Querying data at a parcel level of granularity could convey a false sense of accuracy.

    It might be possible to create yet another unit of geography (refer to earlier discussion) that is finer grained than a TAZ or even grid cell and yet coarser grained than a parcel, thus not only speeding up calculations (see discussion of implementation details) but also make it clearer that traffic volumes and travel times may be more useful for clusters of parcels than for an individual parcel. We discussed briefly the idea of clustering parcels into "blocks" or into "street fronts," with pros and cons to each approach. This warrants further investigation.

- Discrete numbers are less interesting than relative numbers.

That is, rather than hanging one's hat on an AM congested travel time by SOV of 26.37 minutes, the indicators should be used instead to observe, for example, that travel from parcels in one area takes twice as long to reach the CBD during this time than travel from parcels in another area.

- Indicators may be most relevant when doing comparisons across scenarios than for analyzing the details of any particular scenario.

For example, comparing travel times from one parcel or parcel cluster to another may actually inform decision-making when the scenarios under consideration are, e.g., a 6-lane versus a 4-lane 520 bridge replacement.

Finally, given the increasing interest in alternative forms of transportation, and what it might take to live in an area without a car, it would be quite interesting to spend some time teasing out indicators for non-auto-based-travel.

It is understandable that the model outputs are currently auto/truck focused, because this part of the travel demand model was focused on analyzing congestion. However, we would like to see support for scenarios involving alternative transportation options.

Software Architecture and Implementation

In order to implement the scenarios mentioned in the Relevant Data and Scenarios section, a small and versatile system was written in Python 2.4, under the same conditions as UrbanSim. The NetworkX graph library package was used for graph building and searching. The Boost network library was also explored, but could not be implemented due to some limitations involving property maps. It should be noted however that the NetworkX library provides fast results relative to other procedures, such as the transportation node to parcel mapping described below.

Implementation Details

To begin to answer our motivational questions, several pieces of data had to be obtained and connected. The EMME/2 model output would contain the appropriate node links and the transportation attributes we needed, but no geospatial information. The geospatial information would be found within ArcGIS shapefiles which cannot be parsed directly in Python. Finally, the parcel geospatial and zoning type could be obtained in text form from MySQL queries from UrbanSim's database.

To answer shortest-path questions, the data has to be represented in graph format, with geospatial coordinates for parcel-to-network mapping. To do so, the graph nodes are first built using a .dxf file conversion of the transportation network which contains the coordinates of every single node in the network in Washington State Plane North coordinate system. (In fact, it contains all the points of the polylines used in the shapefile, but only the first point of the polyline is taken.) This conversion is done in ArcGIS, but the built system contains the network that is complete for all possible scenarios and thus this step does not need to be repeated for different scenario runs. Parcels are created in a similar fashion, although their coordinates are extracted from a text file.

1)

SAMPLE CALL : `G = street_graph_nodes("NetExt.txt")`

Where G is a graph and NetExt.txt contains the EMME/2 model output.

EXECUTION TIME : ~ 2 mins

2)

SAMPLE CALL : `G = centroids_graph("parcels.txt")`

Where G is a graph and parcels.txt contains parcel id and coordinate info.

EXECUTION TIME : ~ 10 mins

Next, edges are put into the built network nodes. Edges are built based on the model output attribute values and a-b pairs. For each node pair output by the model, a directed link is created from node a to node b with the weight of the attribute desired. The current implementation creates two graphs – one based on the "timeau" attribute and one based on the "@tveh" attribute. This step *does* need to be rerun for every different EMME/2 model scenario. Therefore, if a comparison between SOV and HOV travel times is desired, a graph of the SOV links has to be made and a separate graph of the HOV links has to be made. Links are assumed to have positive weights, so if an edge is encountered that has a negative, or zero weight, it is increased to .001mins.

3)

SAMPLE CALL : `H = street_graph_traveltime(G, "NetExt.txt")`

Where H is a directed graph, G is the graph obtained in step 1 and NetExt.txt contains the EMME/2 model output.

EXECUTION TIME : ~ 10 mins

The most expensive step now follows. In order to support parcel-to-parcel queries, each parcel must be mapped to its nearest transportation node. This is done by calculating the distances to all the transportation nodes and finding the minimum for every parcel. This distance is then divided by an assumed speed of 50 ft/s, which results in a direct travel time from the parcel centroid to its nearest node. Note that this travel time is an underestimate, as the local street network is unlikely to link directly to the nearest node in a straight line without any turns and or stops. Unfortunately, this step must also be redone every time parcel data changes.

4)

SAMPLE CALL: `nearest(D,A,"nearest.txt","parcelsType.txt")`

Where D is the graph obtained in step 3, A is the graph obtained in step 2, nearest.txt will contain the nearest mappings (sink) and parcelsType.txt has information about the type of every parcel in A.

EXECUTION TIME: ~ 12 hours

Now that the parcels are matched with their respective transportation nodes, the basics of behind our motivational question are in place. The usefulness of the implemented system is now dependent on the type of information that is used to build and search these graphs, as well as the type of queries that are made.

Testing

Because the system implemented largely makes use of available data and library packages, testing was done on an application rather than unit level. Testing the individual method calls would be redundant; as such testing is already being done within the NetworkX package. Thus, the tests were performed by visually inspecting the attained graphs to check for input data inconsistency and creating sample indicators that displayed the properties to be tested.

Once the graphs are constructed properly, the errors obtained can only be due to the input data, which comes from EMME/2 or UrbanSim’s SQL database. Thus far, the data has been fairly clean, with exception of some negative edge weights reported by EMME/2, which were corrected, as mentioned above.

Below are the sample calls to three indicators that were created for proof-of-concept and testing purposes:

1) Shortest path queries

SAMPLE CALL : `shortestPathFromTo('FROM', 'TO',D, "nearest.txt")`

Where FROM is the string id of the origin parcel, TO is the string id of the destination parcel, D is the completed street network (with edges) for the scenario (HOV and SOV are different) and nearest.txt contains the nearest transportation node–parcel mappings.

EXECUTION TIME: ~ 30 secs

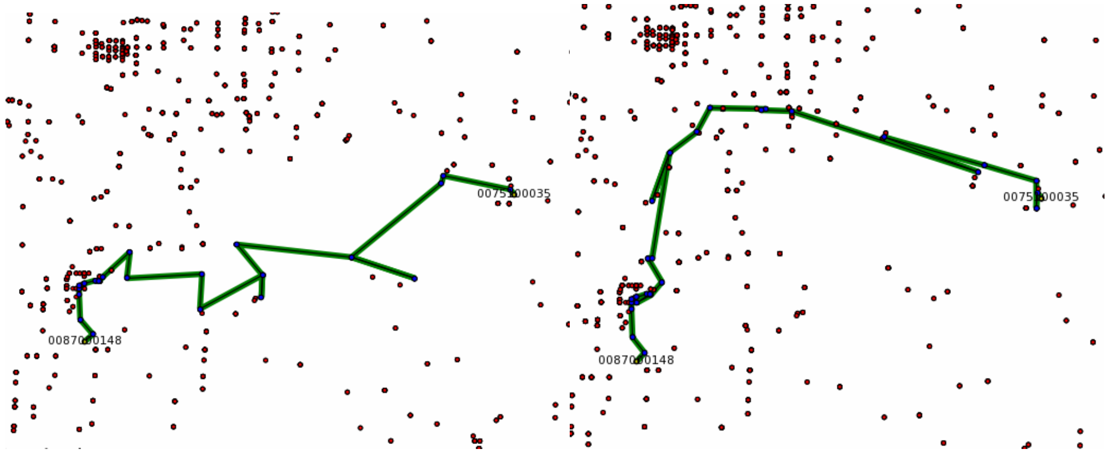


Figure 5: Shortest path queries output

Shortest path queries output between parcels *'0087000148'* and *'0075100035'*.
Comparison of SOV path selection with 25min travel time (L) and HOV3+ path selection with 20min travel time (R) in 2001

2) Passing vehicle impact

SAMPLE CALL : `H = street_graph_volume(G, "NetExt.txt")`

Where H is a directed graph, G is the graph obtained in step 1 and NetExt.txt contains the EMME/2 model output – this time we are interested in volume.

EXECUTION TIME : ~ 10 mins

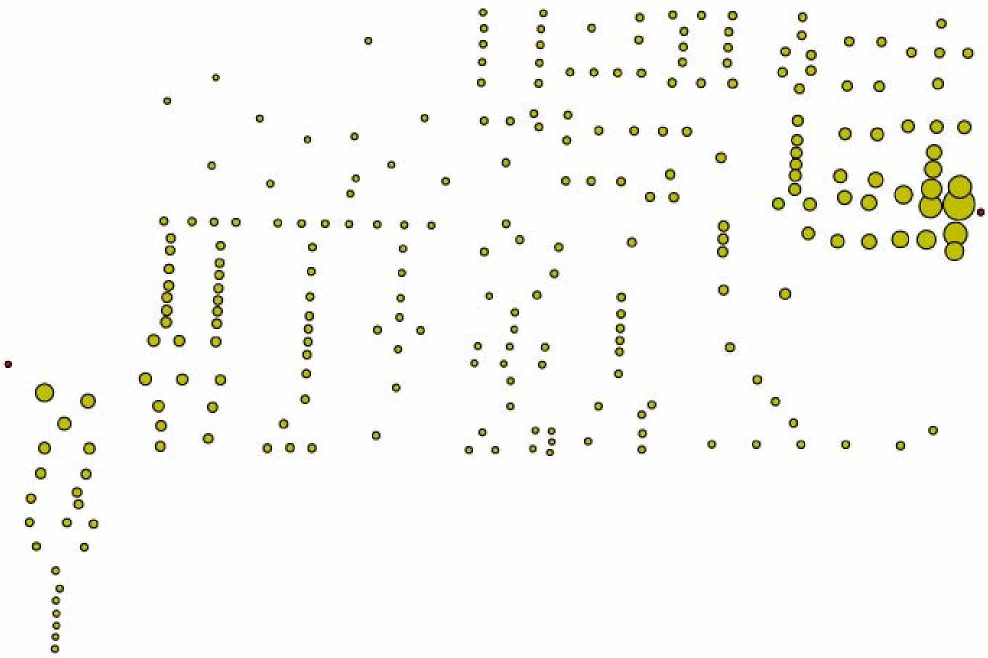


Figure 6: Passing vehicle impact output

Transportation nodes are shown in red (there are one at the far left and one at the far right of the plot), parcels are shown in yellow, proportional to their vehicle impact - the larger the dot, the bigger effect passing traffic has.

3) Types of parcels reachable within a time period

SAMPLE CALL :

`DDD_within_DDD_minutes_of_parcel_SSS('FROM',D,mins,type, "nearest.txt")`

Where FROM is the string id of the origin parcel, D is the completed street network (with edges) for the scenario, mins is the desired time period, type is the county parcel code for the type of establishment to be reached and nearest.txt contains the nearest transportation node–parcel mappings.

EXECUTION TIME: ~ 30 secs

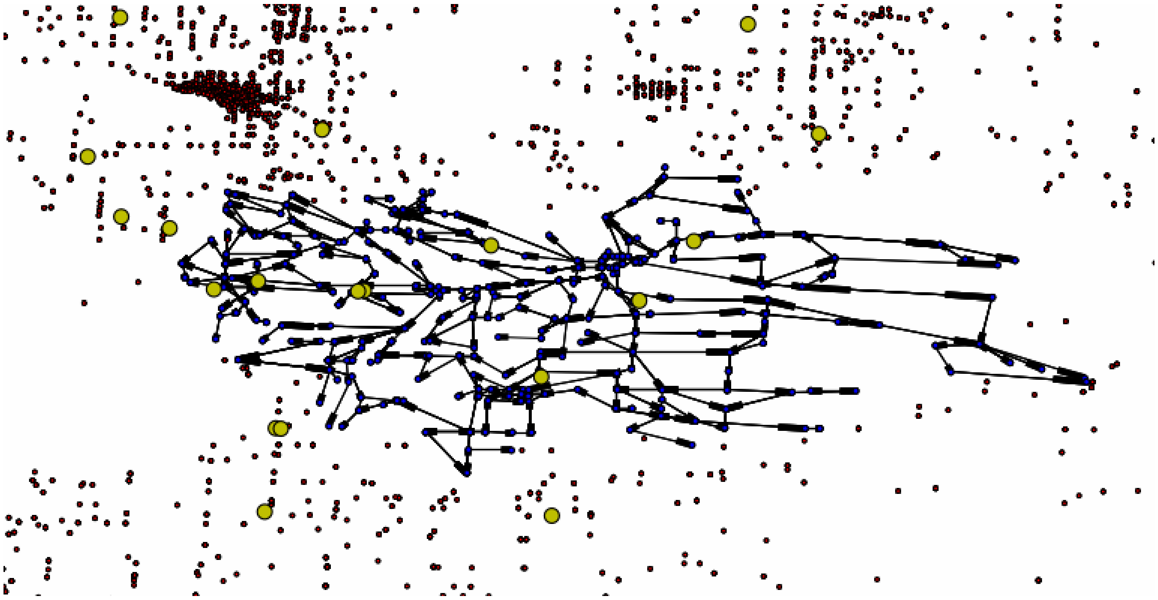


Figure 7: Types of parcels reachable within a time period

Yellow nodes represent desired types of establishments and the FROM parcel (center), smaller red nodes are transportation nodes. Edges are drawn to all nodes that are reachable within 15 mins of parcel ‘*0088000710*’

Specification of Indicators and Proposed Future Work

This section summarizes the implemented indicators that have been generated in this project in addition to potential future indicators.

Indicators Implemented

The project team implemented three indicators: Shortest Path, Land Use Access, and Traffic Impact. For an overview of the scenarios that motivated these specific indicators, as well as the semantics of the underlying data, refer to the sections of this document entitled *Relevant data and scenarios* and *Limitations in the data*.

As discussed in the section entitled *Software Architecture and Implementation*, each of these indicators operates off of a weighted graph of links between transportation nodes, where the weight is calculated from the value of the selected attribute of interest, e.g., congested travel time in our implemented indicators (but a macro could be built to implement other indicators to explore the impacts on parcels of other transportation attributes (refer to appendices for a list of other attributes)). It is also important to understand that the macro used to extract the data from EMME/2 as input to the graph-builder extracts data for a *specific* year and time of day and will likely further filter which links in the transportation network are to be used. In our case, the data for each graph are filtered to extract data for a specified mode, e.g., SOV or HOV3. The following figure captures these concepts.

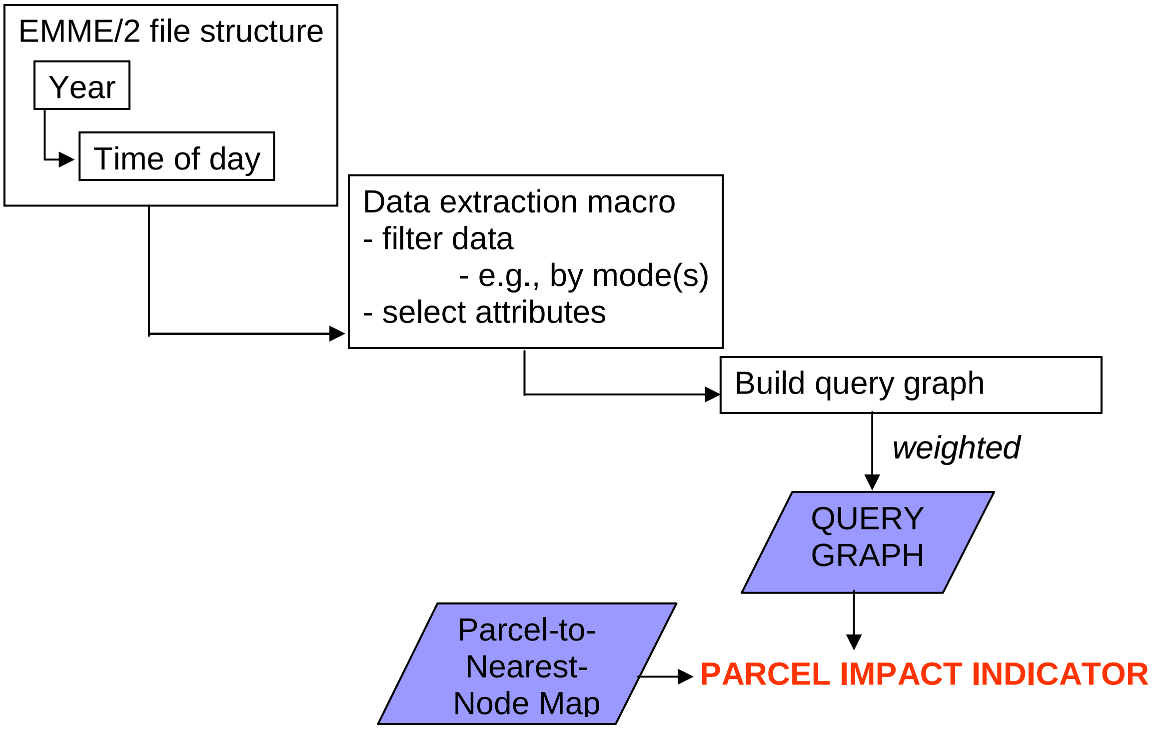


Figure 8: Building data as a pre-step for the indicators

Warning: Since data filtering is done prior to building the graph and calculating the indicator, it will not be obvious from the indicator output how the data were filtered! One suggestion is to name the graph files in such a way that the filtering criteria are a part of the file name, or to supply a string that will be output with the indicator to make it clear what subset of the transportation network data were used.

1- Shortest Path Indicator

Method signature: `shortestPathFromTo('FROM', 'TO',D, "nearest.txt")`

`FROM: string id of the origin parcel`

`TO: string id of the destination parcel`

`D: completed street network graph, filtered for a particular year, time of day, and (likely) mode, and weighted based on a measure of distance (e.g., length or time; our implementation uses congested travel time timeau)`

`“nearest.txt”: parcel-to-nearest-node map`

Indicator purpose/specification: This indicator returns the shortest path, as a sequence of nodes and edges, along the modeled transportation network from one parcel to another. It bases this return value off of a pre-computed graph that contains all of the semantics of the query: year, time of day, weighted attribute (in our case, *timeau*), and other filtering criteria (in our case, we produced one file for SOV travel time and another for HOV3 travel time).

Interpreting results: This indicator could be used to explore how changes in land use location and density or how changes in transportation decisions (e.g., availability of more HOV lanes or a 4-lane versus 6-lane 520 bridge replacement) and the resulting impact on congestion (as estimated out of the 4-step model of EMME/2) might impact the utility of a particular parcel. For example, if the “from” parcel is a residential parcel and the “to” parcel is a CBD work zone, scenarios that imply increased congested travel time might suggest a perceived decrease in the desirability of the residential parcel. If the “to” parcel is a grocery store or a park, it might suggest that certain pockets of residential areas are being positively linked with or negatively cut off from quality of life land uses. If the “from” parcel is in an area of lower economic parcels, it might suggest how subgroups are being affected relative to other subgroups. Time to access services such as grocery stores by various transportation modes might suggest whether it is feasible to live in that area without a car – with the caveat that EMME/2 does not model the local road network, so such analysis, which is really only meaningful for short distances, may not yield useful results given the available data. Of course, it is difficult to draw conclusions from an isolated piece of data, anyway, so the most likely use of this indicator is as a variable in one or more of the UrbanSim models, such as the Land Price or Location Choice Models; its significance can then be understood relative to the significance of other parcel attributes.

Units of measurement and precision: The units of measurement and precision are dependent on the attribute used in the weighted graph. For *timeau*, this is congested time in minutes. Not only is the precision limited by EMME/2's assumptions and calculations for the attribute in question, but due to various limitations in the source data (refer to the section on *Limitations in the data* for a more detailed discussion), any specific parcel-to-parcel value is meaningful only in magnitude and only as compared relative to some other scenario! For example, it might be meaningful to contemplate the shortest travel time from a parcel in one geographic area to a CBD as compared with the shortest travel time from a parcel in another geographic area to that same CBD. It might also be meaningful to compare shortest paths using one set of transportation network assumptions (e.g., a 4-lane 520 bridge replacement) to shortest paths using another (e.g., a 6-lane replacement). It might also be meaningful to contemplate the magnitude of time savings by SOV or by HOV2 or HOV 3. If the network-level data were more accurate for transit or bike, it might be meaningful to do the same for alternative transportation options.

Related indicators: Land Use Access Indicator (see below).

Limitations: (Refer to the section on *Limitations in the data* for a detailed discussion.) 1) Relating a parcel to the transportation network may convey a false sense of precision because EMME/2, which generated the data, based it off of estimated trips between TAZ, not between parcels. 2) We calculate a time to travel from the centroid of the parcel to its nearest node, which is based on a straight-line distance, not the actual local road network, so it may be optimistically short in some but not all cases. 3) The shortest path calculation only considers the portion of the transportation network modeled by EMME/2… which does not include the local road network. 4) Although someone could write a macro to extract data for transit and bicycle modes, the value of the information is only as good as what EMME/2 outputs; at this time, congested travel time for these modes is not available as a link-level calculation, so the output of the macro would not be meaningful. 5) The actual value of congested travel time does not reflect intra-TAZ trips, which are not modeled in EMME/2, so there may be additional traffic on the network whose impact is not accounted for (on the flip side, one might assume that this intra-TAZ impact occurs mostly on local roads, but that might not be true in some geographic regions of the city). 6) Shortest path by congested travel time does not reflect the impact of buses on the transportation network because EMME/2 does not consider buses in trip assignment.

How modeled: The value of this indicator is not completely modeled by Urbansim. Parcels are modeled by UrbanSim, all network path data are obtained from the external travel model, EMME/2.

2- Land Use Access Indicator

Method signature:

```
DDD_within_DDD_minutes_of_parcel_SSS('FROM',D,mins,type,
"nearest.txt")
```

```
FROM: string id of the origin parcel
```

```
D: completed street network graph (with edges), filtered for a
particular year, time of day, and (likely) mode, and weighted
```

```
based on a measure of distance (e.g., length or time; our
implementation uses congested travel time timeau)

mins: upper bound travel duration expressed in minutes

type: county land use code

“nearest.txt”: parcel-to-nearest-node map
```

Indicator purpose/specification: In this indicator, it is possible to find the number parcels with specific land use that can be reached within a specified time from a specified parcel *and using the mode represented by the underlying data*. The implementation is based on the same shortest path algorithm used in the Shortest Path Indicator.

Interpreting results: Warning: Number of parcels returned does not necessarily correlate with a count of businesses, or parks, or whatever land use was specified, because it is often true that a single use spans multiple parcels of land. With this caveat, this indicator could be useful, for example, in analyzing how accessible certain services are from residential parcels in an area of interest. While the Shortest Path Indicator could be used in a similar way, it assumes that one knows the specific parcel of interest, which might be true when considering the CBD or other known destination point but might be less true when any of a number of parcels might fit the desired land use criteria. Like the Shortest Path Indicator, this indicator could be used to assess various aspects of utility related to the origin parcel, or to compare across scenarios to assess how various land use and transportation network decisions impact congestion, which in turn impacts the feasibility of accessing various services via various modes. Rather than a stand-alone indicator, it is expected that it could be used as a variable in one or more of the UrbanSim models, such as the Land Price or Location Choice Models.

Units of measurement and precision: This method returns a set of parcels. The units of measurement and precision used in determining which set of parces to return are dependent on the attribute used in the weighted graph. For *timeau*, this is congested time in minutes. Not only is the precision limited by EMME/2’s assumptions and calculations for the attribute in question, but due to various limitations in the source data (refer to the section on *Limitations in the data* for a more detailed discussion), any specific parcel-to-parcel value is meaningful primarily in comparing relative values across scenarios. Also, as noted above, the number of parcels returned does not necessarily correlate with a count of businesses, or parks, or whatever land use was specified, because it is often true that a single use spans multiple parcels of land.

Related indicators: Shortest Path Indicator (see above).

Limitations: (Refer to the section on *Limitations in the data* for a detailed discussion.) 1) Relating a parcel to the transportation network may convey a false sense of precision because EMME/2, which generated the data, based it off of estimated trips between TAZ, not between parcels. 2) We calculate a time to travel from the centroid of the parcel to its nearest node, which is based on a straight-line distance, not the actual local road network, so it may be optimistically short in some but not all cases. 3) The shortest path calculation only considers the portion of the transportation network modeled by EMME/2… which does not include the local road network. In particular, beware of inputting short durations which might yield no or questionable results if only freeways and arterials are considered as network options. 4) Although someone could write a macro to extract data for transit and bicycle modes, the value of the information is only as good as what EMME/2 outputs; at this time, congested travel time for these modes is not available as a link-level calculation, so the output of the macro would not be meaningful. 5) The actual value of congested travel time does not reflect intra-TAZ trips, which are not modeled in EMME/2, so there may be additional traffic on the network whose impact is not accounted for (on the flip side, one might assume that this intra-TAZ impact occurs mostly on local roads, but that might not be true in some geographic regions of the city). 6) Shortest path by congested travel time does not reflect the impact of buses on the transportation network because EMME/2 does not consider buses in trip assignment.

How modeled: The value of this indicator is not completely modeled by Urbansim. Parcels are modeled by UrbanSim, all network path data are obtained from the external travel model, EMME/2.

3 – Traffic Impact Indicator

Method signature: `H = street_graph_volume(G, "NetExt.txt")`

`H: directed graph`

`G: street node graph containing just the nodes of the transportation network`

`NetExt.txt: the EMME/2 macro output file that will be used to fill in the edge information based on the @tveh attribute of the model run`

Indicator purpose/specification: This indicator expresses the impact of traffic on every parcel in the parcel centroids to nearest node mapping file. The indicator is calculated as a ratio between the traffic volume passing the parcel’s nearest intersection and the distance between the parcel centroid and the nearest intersection:

Weighted impact = #vehicles/ distance from parcel centroid to nearest node

Interpreting results: This indicator could be used to look at the relative impact of traffic volumes on freeways and major arterials relative to the desirability of such traffic to the land use type of specific parcels. For example, residential parcels may be negatively impacted by high volumes of traffic while positively impacted by an easily-accessible major transportation link; commercial parcels may be positively impacted by high volumes of traffic. A huge caveat is that, since the only transportation networks considered by EMME/2 are freeways and major arterials, it is dangerous to use this as an indicator for success of a nearby commercial property since the local road network may not be conducive to accessing the parcel. In general, though, this indicator could be an interesting variable in the UrbanSim Land Price or Location Choice Models.

Units of measurement and precision: This method returns a directed graph with a weighted value for the traffic impact. The same caveats discussed in the earlier indicators apply.

Related indicators: Shortest Path Indicator (see above).

Limitations: (Refer to the section on *Limitations in the data* for a detailed discussion.) 1) The traffic volume used in calculating the indicator represents traffic only on arterials and freeways and thus is not a good indicator of the impact of local traffic. 2) Relating a parcel to the transportation network may convey a false sense of precision because EMME/2, which generated the data, based it off of estimated trips between TAZ, not between parcels. 3) Although someone could write a macro to extract data for transit and bicycle modes, the value of the information is only as good as what EMME/2 outputs; at this time, traffic volumes and congested travel time for these modes are not available as a link-level calculation, so the output of the macro would not be meaningful. 4) The actual values for traffic volumes and congested travel time do not reflect intra-TAZ trips, which are not modeled in EMME/2, so there may be additional traffic on the network whose impact is not accounted for. 5) The data does not reflect the impact of buses on the transportation network because EMME/2 does not consider buses in trip assignment.

How modeled: The value of this indicator is not completely modeled by Urbansim. Parcels are modeled by UrbanSim, all network path and traffic volume data are obtained from the external travel model, EMME/2.

Recommended Areas of Focus for Future Work

There are two recommended areas of focus for future work that can be undertaken in UrbanSim (of course, there are several limitations, discussed in earlier sections, in the underlying EMME/2 data that are larger scope issues):

1) Consider introducing a new unit of geography that is coarser-grained than a parcel but finer-grained than a grid cell. This might be a collection of parcels that share geographic characteristics (like a block) or perhaps further share transportation access characteristics (like a set of parcels sharing a street front). The purpose of such a new unit of geography would be to correct, somewhat, the false sense of precision that is conveyed in relating individual parcels to the transportation network. It would likely have the side benefit of reducing calculation time in building the underlying graphs.

2) Investigate the feasibility of constructing new link-level attributes, using the techniques described in Appendix C, or new indicators that estimate volumes and travel times for alternative transportation, such as transit and bike travel.

Potential New Indicators: Feasibility of Living without a Car

This is likely a collection of indicators addressing the question about the feasibility of living in a parcel, or an area, without a car. The Shortest Path and Land Use Access Indicators are a piece of this puzzle. However in order to live in a parcel without a car comfortably, you need to have alternate modes so that you can access different necessary facilities like offices and grocery stores. The existing indicators should be expanded, or new indicators implemented, to support queries by alternative travel modes such as transit or bike.

With respect to transit, it may be possible to mine the EMME/2 matrices that were used to populate the transit link and segment tables for ways to estimate volume and congested travel time.

1. Bus-Route Availability Indicator

This indicator would look at the existence of an actual connected bus route (not just any random transit link) between two specified parcels. Answering this question would require manipulation of segment-level transit data which was out of scope of this project.

2. Bike Feasibility Indicator:

This indicator would estimate bike commute time over a bike lane or path as a function of path length and a bike speed assumption. In the EMME/2 link-level data, bike lanes are encoded on the links (although it looks as if the encoding simply assumes that no freeways support bikes and all arterials have bike lanes) and separate bike trails networks are available. While our investigation suggested that volume data are not available for bikes, it may be possible to estimate travel time based solely on length of a link and an assumed bicycle travel speed.

Potential New Indicator: Level of Service (LOS)

Level of Service is defined as the level of operating performance (that is, the quality of flow changes with the traffic density) on the highway. As shown in the diagram below, the network could be classified according to the LOS which depends on:

- number of lanes in each direction;
- traffic volume;
- and free flow speed.

This indicator could be used in UrbanSim for:

1. Traffic management and planning:

Building new facilities and expanding the existing facilities depending on the urban growth areas.

2. Commercial choice:

Some kinds of businesses prefer to be away from low LOS like E or F.

3. Housing choice:

Generally, house location is preferred to be in areas with high LOS and people usually avoid living near intersection of LOS E and F.

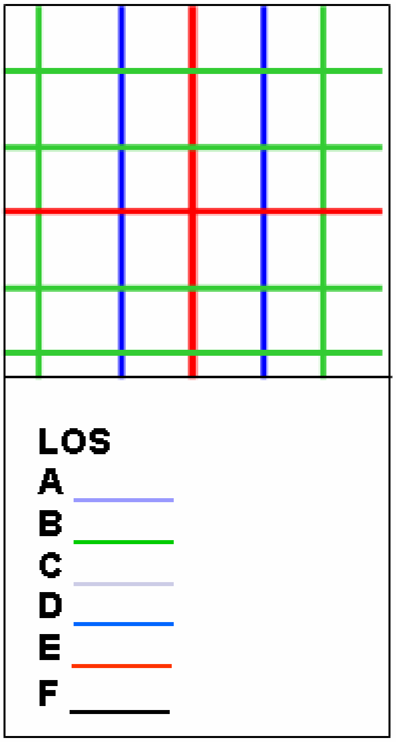


Figure 9: LOS assigned to network

Conclusions

Parcel-level indicators are a feasible alternative to the currently used zone-level indicators. The results obtained are compatible with the notion of "household indicators" as described by a recent work by Davis and provide valuable insight towards more realistic and plebian model results: "…there has been a push for greater realism and detail in reporting Household Indicators, making them even more directly related to citizens' experiences living in the region."

Of course, the interpretation of the indicators must be based on an understanding of the limitations in the underlying data. Several limitations in the underlying data described in detail in this paper were identified as:

- Data at a parcel level of granularity could convey a false sense of accuracy, since the travel data were not generated as parcel-to-parcel trips in the first place, the congestion volumes and travel times do not include buses, intra-zone trips are not modeled, and the local road network is not represented.
- Because of the lack of precision in the underlying data, discrete numbers are less interesting than relative numbers. That is, rather than hanging one's hat on an AM congested travel time by SOV of 26.37 minutes, the indicators should be used instead to observe, for example, that travel from parcels in one area takes twice as long to reach the CBD during this time than travel from parcels in another area.
- Indicators may be most relevant when doing comparisons across scenarios than for analyzing the details of any particular scenario. For example, comparing travel times from one parcel or parcel cluster to another may actually inform decision-making when the scenarios under consideration are, e.g., a 6-lane versus a 4-lane 520 bridge replacement.

With increasing precision in these parcel-based indicators, an individual could increasingly obtain information that is truly meaningful to them. However, more likely, urban planners would use the information to assess improvements or degradation in accessibility across alternate scenarios. For instance, one could compare the impact of building a 6-lane vs. a 4-lane bridge across Lake Washington on the commute times for various geographic areas, or how land prices might be impacted by changes in vehicular traffic passing nearby intersections. This in turn could be connected to the household pricing model to see how the value of the home would change. Many queries can be designed to shed some light on many of the proposed transportation improvements. It is also not unthinkable to be able to produce an individual monetary value for each proposal, provided some more input from the individual. This value could be then compared to the cost of the proposal to the individual and a direct comparison can be made.

An ethical question of social engineering could be posed, as due to the non-transparency of the model and its dependency on PSRC data, it could be conceived that this model could be used to sway voters to a particular conclusion. This type of analysis would also have to place costs on less tangible items such as environmental and aesthetic costs which are dependent on the individual and are highly variable with economic climates. These are all issues that will have to be dealt with in the most transparent way possible, in order to allow citizens to understand the process by which the information is processed and presented. A larger question, no doubt on the mind of anyone that would be using this model will be – can we really rely on a model to represent and predict the impact of our present day decisions? Can a handful of variables, many of which are crudely derived, tell us what’s best for us 30 years on? It’s important for users to keep the results of the simulation in perspective and not base decisions solely on its results, as the results of the simulation are only as accurate and complete as the input data.

Being able to extract parcel-level travel information is a significant step in the development of the “household indicators.” The work presented in this report provides a solid base for implementing these indicators, which in turn will be a valuable tool in the planning process.

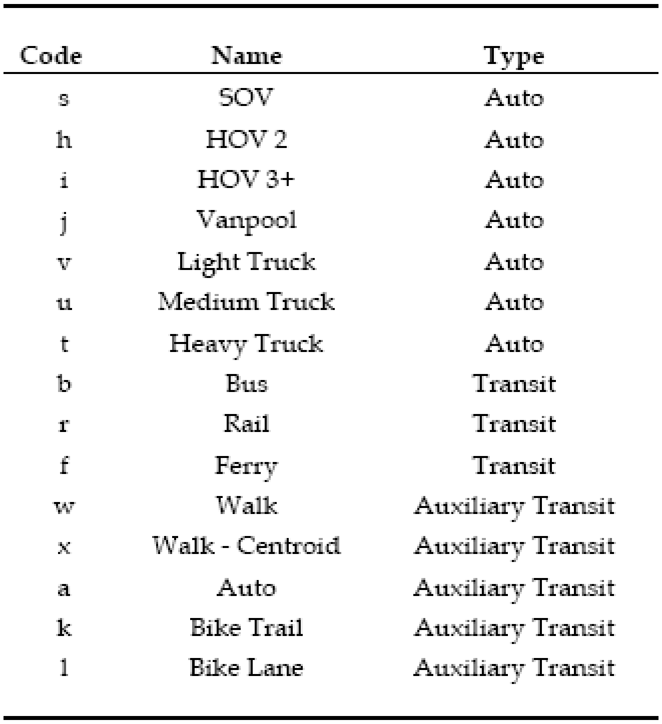
Appendix A: Network Link Attributes

The information in this Appendix is copied from [PSRC, 2007b].

Link descriptors

- a-node;
- b-node;
- length in miles;
- modes allowed to travel on the link;
- type used for screenline number (linktype = 90 are links that are not on any screenline);
- number of lanes or boats per hour on ferry links;
- volume-delay function index;
- user-defined data 1: capacity per lane or vehicles per boat on ferry links;
- user-defined data 2: observed speed limit;
- user-defined data 3: facility type.

Allowed values for mode



Code	Name	Type
s	SOV	Auto
h	HOV 2	Auto
i	HOV 3+	Auto
j	Vanpool	Auto
v	Light Truck	Auto
u	Medium Truck	Auto
t	Heavy Truck	Auto
b	Bus	Transit
r	Rail	Transit
f	Ferry	Transit
w	Walk	Auxiliary Transit
x	Walk - Centroid	Auxiliary Transit
a	Auto	Auxiliary Transit
k	Bike Trail	Auxiliary Transit
l	Bike Lane	Auxiliary Transit

Allowed values for facility type

- freeways (1);
- expressways (2);

- urban 2-way arterials (3);
- urban 1-way arterials (4);
- centroid connectors (5);
- rural 2-way arterials (6);
- ferries (0);
- bike (0);
- walk (0).

EMME/2-generated values after assignment (selected)

volau: link volume in vehicle equivalents[2]
timeau: travel time in minutes resulting from the link having been congested by volau

Mode –specific volumes in equivalent vehicles (trucks: light =1, medium =1.5, heavy=2)
@sov Single Occupant Vehicle (including all four HBW-SOV volumes)
@hbw1 HBW1 SOV Volumes
@hbw2 HBW1 SOV Volumes
@hbw3 HBW1 SOV Volumes
@hbw4 HBW1 SOV Volumes
@hov2 2 person carpool
@hov3 3+ person carpool
@vpool Vanpool (8 person)
@light Light truck
@mediu Medium truck
@heavy Heavy truck

Truck vehicle volumes (in vehicles)
@lveh light-truck volume on link (vehicles)
@mveh medium-truck volume on link (vehicles)
@hveh heavy-truck volume on link (vehicles)

Total vehicle volumes (in vehicles)
@tveh total volume on link (vehicles)

---

[2] Assigned light, medium, and heavy trucks are represented as automobile equivalents, e.g., a light truck = 1, medium = 1.5, heavy = 2.

Appendix B: Transit Segment Attributes

The information in this Appendix is copied from [PSRC, 2007b].

Transit line (segment) attributes

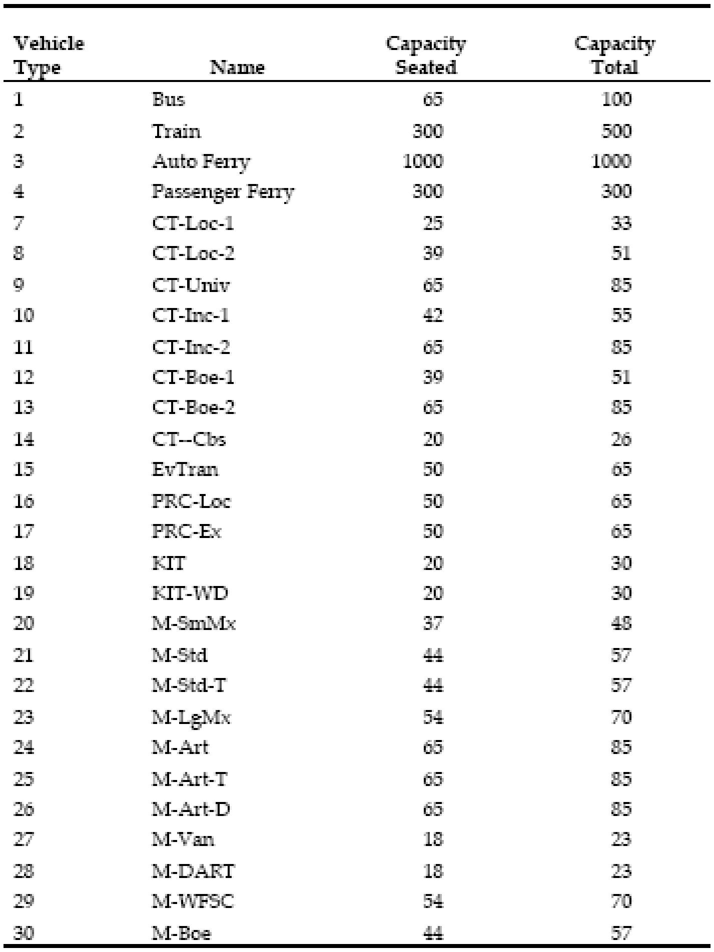
Header:

- line ID and description;
- mode (b or f or r);
- vehicle type;
- headway;
- default speed;
- user defined 1;
- user defined 2;
- user defined 3 (operating agency code).

Itinerary:

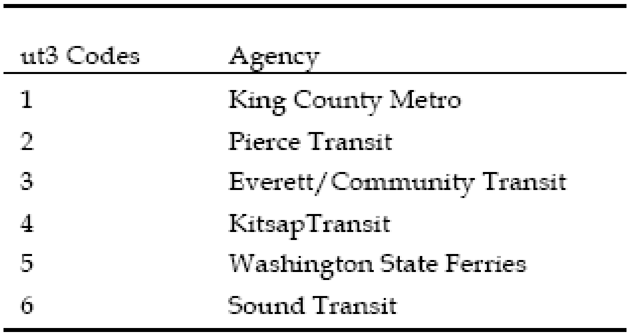
- a-node;
- b-node;
- dwell time (generally 0.25 minute, #.00 indicates that the link does not contain a stop);
- transit time function index;
- user defined data (not used).

Allowed values for vehicle type



Vehicle Type	Name	Capacity Seated	Capacity Total
1	Bus	65	100
2	Train	300	500
3	Auto Ferry	1000	1000
4	Passenger Ferry	300	300
7	CT-Loc-1	25	33
8	CT-Loc-2	39	51
9	CT-Univ	65	85
10	CT-Inc-1	42	55
11	CT-Inc-2	65	85
12	CT-Boe-1	39	51
13	CT-Boe-2	65	85
14	CT--Cbs	20	26
15	EvTran	50	65
16	PRC-Loc	50	65
17	PRC-Ex	50	65
18	KIT	20	30
19	KIT-WD	20	30
20	M-SmMx	37	48
21	M-Std	44	57
22	M-Std-T	44	57
23	M-LgMx	54	70
24	M-Art	65	85
25	M-Art-T	65	85
26	M-Art-D	65	85
27	M-Van	18	23
28	M-DART	18	23
29	M-WFSC	54	70
30	M-Boe	44	57

Allowed values for operating agency



ut3 Codes	Agency
1	King County Metro
2	Pierce Transit
3	Everett/Community Transit
4	KitsapTransit
5	Washington State Ferries
6	Sound Transit

Allowed values for transit time function

The transit time function is the algorithm used to calculate congested travel time for selected transit modes as a function of time of day and link length and/or vehicle congested time (timeau). The resulting value is not stored, although it may be in the EMME/2 matrices from which the link and segment data files are generated.

- ft4 (Train): Station to Station Time;
- ft5 (Passenger Ferry): Ferry Route Sailing Time / 3;
- ft 11 / 15 (Bus - AM Peak): 1.64934 * (Minimum of Auto Congested Time and Length * 12);
- ft 12 (Bus - AM Peak): 2.044616 * (Minimum of Auto Congested Time and Length * 12);
- ft 13 (Bus - AM Peak): 2.013138 * (Minimum of Auto Congested Time and Length * 12);
- ft 14 (Bus - AM Peak): 1.331*(Minimum of Auto Congested Time and Length*12);
- ft 11 / 15 (Bus - Mid-day): 1.72431*(Minimum of Auto Congested Time and Length* 12);
- ft 12 (Bus - Mid-day): 1.94323*(Minimum of Auto Congested Time and Length* 12);
- ft 13 (Bus - Mid-day): 1.913313*(Minimum of Auto Congested Time and Length* 12);
- ft 14 (Bus - Mid-day): 1.21*(Minimum of Auto Congested Time and length*12).

EMME/2-generated values after assignment (selected)

@voltr: number of persons traveling on the transit segment
@ehedwy: effective headway (calculated as a function of headway (one of the transit line attributes, above, and congestion)

Appendix C: A Session with Interactive EMME/2

The macro used to extract link-level data from EMME/2 is essentially just a scripted version of an interactive session. As such, it is often useful to first go through the menu sequences required to obtain needed data, so that it's clearer how to encode the menu responses in the macro. This appendix captures a sample interactive session to:

1. Create a new link-level attribute that aggregates transit segment level data (remember that there is a transit segment record for every bus route that runs on a link, so each link has 0-to-many segments).
2. Assign values to an attribute.

This particular session was performed using the data in D:\baseline_travel_model_TEST\2000_06\bank1.

1. Use Module 2.42 to create a new link-level attribute from transit segment data

```
- Menu choice 2, CREATE attribute
- Menu choice 2, it'll be a LINK attribute
- Enter attribute name, here I chose @sumtr (@ + 5-character name)
- Enter description for attribute: volume of transit riders on the link
- Enter default value, here I chose 0
```

2. Use Module 2.41 to assign values to the attribute.

```
- Menu choice 1, network calculation
- Answer save result, yes
- Enter attribute to save result in: @sumtr
- Answer change description of attribute, no
- Enter expression to calculate @sumtr = voltr (that is, @sumtr will be
derived from the volume-of-transit-riders on that link)
- Enter "enter" to end expression (because can extend over mult lines)
- Menu choice 4, it'll be calculated as the SUM of the segment-level
attribute voltr for each segment on the link
- Enter *, to indicate want all transit lines
- Enter *, to indicate want all links calculated
- Menu choice 4, no output (there are other options, but since we've
saved the values in the db we'll just extract them using our macro)

      I calculated the value and got a summary report sent to the
      screen:
      - minimum value calculated: 0 riders
      - max value calculated: 11,938 (does this seem possible?)
      - sum of all values: 2908147
      - average result: 375.9 (this seems way high)
```

Note that Module 2.41 is the same module used in the link-level data extraction macro. Specifying values for expressions, as we do when selecting specific modes, requires an extra carriage return, just as did the expression to calculate @sumtr, above. That is something that would not be obvious without using the interactive system. (Note: An asterisk * is used to specify that all values are being selected, and this does not require the extra carriage return because the interactive menu does not initiate expression-building.)

Bibliography

Cambridge Systematics, Inc. (2006). *PSRC Travel Model Documentation Draft Final Report, Updated for Congestion Relief Analysis*. Seattle, WA: Washington State Department of Transportation and Puget Sound Regional Council.

Household Indicators: Design to Inform and Engage Citizens. (2006) Davis, Janet

Puget Sound Regional Council. (2007a). About the Puget Sound Regional Council. Retrieved June 2, 2007, from http://www.psrc.org/about/index.htm.

Puget Sound Regional Council. (2007b). *Land Use and Travel Demand Forecasting Models: PSRC Model User's Guide for Version 1.0*. Seattle, WA: Originally prepared by Cambridge Systematics, Inc, 2001.

Puget Sound Regional Council. (2007c). UrbanSim. Retrieved June 2, 2007, from http://www.psrc.org/data/urbansim/index.htm.