Tools to Create Network-Based Accessibility Grids for Land Use Modeling

Abdunaser Arafat – Shimberg Center for Housing Studies, University of Florida

Abstract

The coordination between land-use and transportation has been the focus of several research studies. Transportation impacts land use and land use impacts transportation. Traditionally, the land use-transportation planning cycle is one of the main causes of urban sprawl. The construction of freeways facilitated mobility and urban expansion which in turn generated demand for more freeways. Researchers have identified this problem and recommended planning for accessibility instead of planning for mobility. The Land Use Conflict Identification Strategy (LUCIS) is an automated GIS-based land use modeling system that uses suitability and criteria evaluation matrices (CEMs) for envisioning future growth. However, LUCIS uses Euclidian proximity and point density grids to model accessibility to services. This paper will introduce new tools created by Model Builder, Python and geo-processing objects to create accessibility grids based on network distance/time as well as service density and attraction.

Introduction

Land-use suitability surfaces are used in hierarchal structures that include constructing single utility and multiple utility grid surfaces in which the composition of these utilities are performed using preferences and community values (Carr & Zwick, 2007). In these utility compositions, accessibility is simplified as the proximity to facilities such as proximity to highways and transit stations. This proximity estimation discriminates between destinations based only on distance. However, travel behavior literature indicates other important variables such as the attraction of the destinations
are important. This paper investigates and compares mathematical models that are used to estimate accessibility in the transportation and land use planning literature.

Gravity, opportunity and distance models are used to model accessibility (e.g. see, Waddell, 2002; Handy, 2004; Hanson 2004), however, they are applied on either a small random data sample or on aggregate and zonal level data such as TAZs and/or are based on Euclidean distance estimations. In this paper, the distance estimation component can be currently performed by network or Manhattan distance as alternatives to using Euclidean distance. The network distance is the distance travelled between two locations using the road network. The Manhattan distance is the distance travelled between two locations following a grid network while the Euclidean distance is the straight line distance between two locations. The aforementioned distance estimation methods are used to evaluate the suitability in the proximity component, and in creating the capture area for the opportunity suitability component.

**Literature on Estimating Accessibility**

The measurement of accessibility varies in existing research, ranging from linear distance to network distance, travel time and the number of activities within a distance from an attraction or a certain residential location. Accessibility is defined as the potential to interact. To differentiate between the accessibility and mobility, we can say that mobility is the potential to move. In these terms, accessibility is connected to destinations and the mobility is connected to the networks and vehicles. Accessibility for example, measures the number of jobs in a certain area or the number of destinations in a specified area or the availability of choices between modes, while mobility deals with traffic delay and level of service (Handy, 2004). However, this explains that accessibility could be different between modes, but does not take congestion into
consideration because it is a mobility measure according to the definition. The same approach for accessibility is taken by Hanson (2004). According to Hanson, accessibility is the number of opportunities within a distance or travel time while mobility refers to the ability to move between different sites. Furthermore, Hanson explains that because the distance between activities becomes larger as density decreases, accessibility becomes dependent on mobility (Hanson, 2004). This adds interdependency between density, accessibility and mobility. It is clear that there is a relationship between accessibility and mobility and that this relationship is stronger for regional destinations other than neighborhood destinations particularly when different modes of transportation are taken into consideration (Hanson, 2004; Salomon et al., 1998). In regional destinations, using highways and freeways increases the dependency of accessibility on mobility. Accordingly, when transit is involved accessibility may depend on the level of service and thus on mobility.

Accessibility can be also defined as the ease with which a destination can be reached and it is one of the important factors in location decision choices. This definition clearly connects accessibility as a function of land-use and transportation patterns. Accessibility also can be defined as the ease with which people can participate in activities. Such a definition acknowledges that the destination activities and location properties are important factors in accessibility (Primerano & Tylor, 2004).

Accessibility measurements can also be divided into personal accessibility and place accessibility. Personal accessibility can be measured by counting the number of activities within a certain distance of a person’s home. The measurement can also include the magnitude of the distance for each location in a gravity cumulative
approach. The accessibility for a place investigates the number of activities at a certain
distance from a place. These are simple methods for calculating accessibility. More
advanced methods of time-space analysis are needed to address the effect of time on
accessibility (Hanson, 2004). However doing time-space analysis on a disaggregate
level of data at a dependable accuracy is not always possible in places with poor travel
activity diaries.

To simplify, choosing to use personal accessibility or place accessibility depends
on whether personal characteristics are included in the estimation of accessibility. For
example, it is possible to use personal characteristics to estimate accessibility for
existing urban development but it is more complicated to predict personal
characteristics for new development. Generally, the estimation for accessibility can be
a topological or opportunity measurement or both. The topological estimations are an
estimation of physical proximity from origins to destinations which includes the
measurement of distance such as the distance to the nearest location. The opportunity
models measure a density or attraction of accessible places. Incorporating both gives
the relative accessibility which can be clearly shown in gravity models. This relative
accessibility if accumulated for a large scale will result in a measure of an absolute
accessibility (Levinson & Krizek, 2008).

Land-use change and land-use prediction models use accessibility estimations to
model the change of land-use over time. Topological accessibility using proximity is
used generally in suitability models of locations (Carr & Zwick, 2007). Gravity models,
which can be based on the available opportunities and their travel distances to
locations, are used in the modeling of location using statistical methods (e.g. modeling
the employment opportunities for residential location, Waddell et al, 2003). Parcel-level opportunity and gravity accessibility indicators using either Euclidian or network distances are not yet used in land-use suitability analysis. Usually the simple accessibility estimation, as defined as the proximity by the Euclidian distance measurement is used in LUCIS models (Carr & Zwick, 2005).

Many mathematical forms are used to estimate accessibility. Bhat et al (2002), summarized accessibility measurements into different equations for cumulative opportunity and gravity. However, they are applied on either a small random data sample or on aggregate and zonal level data such as traffic analysis zones (TAZs). These estimations are classified as Gaussian, composite impedance, activity distance and in-vehicle travel time. Table 1 compares the variables used for accessibility estimation while Table 2 compares the merits and limitations in accessibility estimation using different methods.

In the absence of measured travel times, the accuracy of topological accessibility estimations depend on the method used to estimate distance. Generally, distance measurement methods in land use research are one of three methods; Euclidean distance, rectilinear distance (Manhattan) and network distance. Network distance can be obtained from a network property approach by measuring the length of street segments as a percentage of the whole street network (O’nell, 2004), or by measuring the actual distance travelled (Zhao et al, 2003). The use of travel time may be more sophisticated and take additional variables into consideration. However, in measuring network distance barriers can be included to give a more accurate indication of travel distance. Arafat, Steiner and Bejleri (2008) compared network distance to Euclidean
and Manhattan distance in research on school siting. Their research found that the use of network distance gives a better estimation for walking distance than Euclidean or Manhattan distances. Additionally they found that the catchment area for population, which is an accessibility indicator, is exaggerated when using a Euclidean buffer. The network distance is used in transportation research to build accessibility indices in Texas (Bhat et al, 2002), where, the travel distance had been obtained from travel surveys which may not be available on a disaggregate level. An alternative methodology can be used to generate the network distance at a parcel level using ArcGIS network analyst which is software that can calculate distance from origins to destination following the road networks. In this methodology, the shortest network distance can be measured from each origin to each destination (Arafat et al, 2008).

Traditionally, proximity as straight line distance has been used as an accessibility measurement in deterministic land-use models (e.g. see, Zwick & Carr, 2007) while gravity models have been used in statistical and stochastic models (e.g. see, Waddell, 2002). There are also other measurements of accessibility used in statistical models such as opportunity access (e.g. see, Handy, 2004; Hanson 2004). However, using gravity models on a parcel-level scale requires generating huge origin-destination matrices that contain billions of records for a large county as a study area. These origin-destination matrices exceed the capacity of hardware and software used in the analysis. Therefore, a methodology to create smaller representative datasets is used in the paper. The distance estimation component can be performed by network analysis. This distance estimation method can be also used to generate the capture area for opportunity estimation which has historically used Euclidean distances. Arafat, Steiner
and Bejliri (2008) discussed the differences between capture areas based on network and Euclidean distances in terms of the number of students within a walking distance to schools and recommended the use of network distance to generate the capture area.

Table 1. Variables used in accessibility equations.

<table>
<thead>
<tr>
<th>Article</th>
<th>Distance</th>
<th>Network</th>
<th>Opportunity</th>
<th>Gravity</th>
<th>Gravity</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr and Zwick (2007).</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D- Distance</td>
</tr>
<tr>
<td>Levinson and Krizek (2008).</td>
<td>D or TM</td>
<td>O</td>
<td>O, TM, α</td>
<td></td>
<td></td>
<td>TM- Travel Miles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α- Decay Factor</td>
</tr>
<tr>
<td>Arafat, Steiner and Bejliri, (2008).</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ND- Network Distance</td>
</tr>
<tr>
<td>Bhat et al. (2002).</td>
<td>TM, TT</td>
<td>O</td>
<td>O, TM, α</td>
<td>O, TT, α</td>
<td></td>
<td>TT- Travel Time</td>
</tr>
<tr>
<td>Handy (2004).</td>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanson (2004).</td>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waddell et al. (2003).</td>
<td></td>
<td></td>
<td></td>
<td>O, TM, α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grengs (2009).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O, TT, α</td>
<td></td>
</tr>
<tr>
<td>Srour et al (2002).</td>
<td>D</td>
<td></td>
<td></td>
<td>O- (Driving time buffer)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhat and Guo (2004).</td>
<td>TT</td>
<td></td>
<td>O, TM, α</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creating distance a raster is one of the simplest models to measure accessibility.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This is traditionally created using the Euclidean distance raster tool provided by the ArcGIS software. Creating raster grids using other type of distances such as the Manhattan and network distance is not provided by the ArcGIS software. This paper uses Python programming and other ArcGIS tools programmatically to create distance raster grids based on Manhattan and network distance. The Manhattan distance grid is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
created by map algebra operations that are performed on the output of ArcGIS Euclidean rasters. The procedure is performed in Python and produced as a customized tool for Manhattan distance as shown in Figure 1.

![Figure 1 Manhattan Raster Tool](image)

The paper also introduces the network distance raster which provides the shortest network distance to the nearest facility. The network distance raster is also a customized Python tool that estimates distances from all the locations to their nearest destinations and generates a network distance raster instead of a Euclidean distance raster. To reduce the memory used by the tool and to increase the processing speed, the tool offers the use random origins and uses them to estimate the distance from origins to destinations using Network analyst. The random origins are generated by the user before using the tool. The use of the random sample is optional and the user can replace them by the whole population. The output raster is created by interpolating the
results of the distance estimated for each origin point to the closest destinations. Figure 2 shows the interface of the network raster tool.

![Network Distance Raster Tool](image)

**Fig 2 Network Distance Raster Tool**

The opportunity access is also one of the common methods for estimating accessibility and can be defined as the number of opportunities within a defined area. It can also be weighted to include the floor area of services according to the following equation:

\[ A_i = \sum_i O_i C_i \]

Where:
\( O_i \) = the weight or the attraction for a facility

\( C_i = 1 \) for \( D_i < \) buffer distance and zero else where

The opportunity is calculated by accumulating the attraction value of the service, which depends on the type of service. This attraction could be the square feet for retail services, the number of beds for hospitals or any other criteria that can be used to discriminate our preference of one service to another. ArcGIS 10 provides many tools to estimate the opportunities around any location using the point density or the focal statistics tools. These tools usually work on Euclidean based distances. However, a Manhattan distance can be defined as a neighborhood for opportunity estimated by focal statistics. This paper provides a new customized tool to estimate opportunity within a certain travel time or travel distance based on street networks. The result of the tool is represented by a raster grid.

Fig 3 Network Opportunity Tool
The gravity access has also been used to estimate accessibility. According to Bhat et al (2002), the accessibility measure is summarized into different equations for cumulative opportunity and gravity measures which are classified as Gaussian, composite impedance, activity-distance and in-vehicle travel time measures. The equation for calculating access from gravity is:

\[ A_j = \sum_{j} \frac{O_j}{d_{ij}^\alpha} \]

Where,

- \( O_j \) is the weight or the attraction for the facility
- \( d_{ij} \) is the distance from each to each destination
- \( J \) is the number of destinations connected to each origin
- \( \alpha \) is the distance decay factor

The distance as impedance estimation is a gravity estimation. However, the gravity model is a simple model for use on a zonal level such as TAZs. On a parcel level, the model will generate an origin-destination matrix that contains hundreds of millions of trip combinations if applied to any of the large counties in Florida. Applying the gravity mathematical equation at that level is impractical in terms of hardware and software limitations. However, in this paper the gravity approach is used by using a representative random sample of origins to capture the accessibility to destinations and interpolate the results to predict the accessibility from any point. To provide flexibility in the tool, the use of random points is optional and they are added to the set of input data for the new tool.
The use of accessibility is important in coordinating land use and transportation. Therefore, the accessibility grids that are created using the aforementioned tools are used in land use planning models. The Land use Conflict Identification Strategy (LUCIS) is a GIS-based land use model that identifies the conflict between major land uses such as agriculture, conservation and urban. LUCIS is based on suitability and the Analytical Hierarchy Process (AHP) to model land use (Carr & Zwick, 2007). LUCIS models have been recently modified to account for network distance instead of Euclidean distance for accessibility estimation and for more coordination of land use with transportation. The access tools introduced in this paper are used in LUCIS in two levels. The first level is to replace some of the Euclidean proximity and density grids in the suitability models. The second level is in a post processing process for the allocation of future population. LUCIS+ is the new version of LUCIS. Plus stands for the processing of land use
scenarios. LUCIS+ adds flexibility to running land use scenarios in a post processing environment using a Criteria Evaluation Matrix (CEM). The CEM is an enumeration matrix that contains the output of LUCIS conflict analysis as well as other criteria and uses that for the population allocation following a user controlled automated process. The accessibility tools presented in this paper are used extensively in running land use scenarios using the CEM and LUCIS+.

**Results and discussion**

The tools have been used in estimating the access to services in the land use models. Traditionally, proximity has been used to estimate access to services. This is usually performed by estimating the Euclidian distance from an origin to a destination. The Euclidean distance raster tool provided by ESRI estimates the distance between any cell in the raster to the nearest facility or destination feature (ESRI, 2011). Facilities like schools and hospitals are usually favorable locations in terms of proximity while facilities like prisons and noise sources are regarded as a disadvantage in terms of proximity. For example, distances to amenities will generally have decreasing suitability meaning that if the distance away from the amenity increases, the suitability of a location decreases. Conversely, for facilities that are less desirable to be located near, suitability will increase as the distance away from the facility increases.

The reclassification to decreasing and increasing suitability usually follows a sequential process including descriptive statistics. An example of descriptive statistics used in the process is the zonal statistics of the Euclidean distances for all the residential parcels. From the zonal statistics, the mean and standard deviation is calculated and their values are used to assign the suitability values. It should be noted,
however, that proximity does not represent the actual travel distance from an origin to a destination. It also does not discriminate between destinations according to their size or attractiveness.

Proximity can also be projected by estimating the network distance from any cell in a raster to the nearest facility or destination feature. However, the estimation process is not a direct output from ArcGIS tools. Network analyst can be used to locate the nearest network points for each origin and its nearest destination. Using the shortest path method, the network distance is also calculated. Then the values for the distance of the whole trip is summed and assigned as a score for the origin in the raster, allowing the user to create a network distance raster. This procedure is automated using a customized Python tool created for this research. The network distance raster (Figure 5) is created by the tool and shows the distance to the nearest retail locations in Heartland region in central Florida. It should also be noted that the network distance raster may be a better estimation of travel distance to services but it does not discriminate between destinations according to their size or attraction. The raster only estimates the distance to the nearest facility. Furthermore, the method is used for destinations that can be reached by walking or driving using street networks. Therefore, it is not used to estimate the proximity to noise generators, sources of pollution and other point sources where the Euclidean distance maybe a better estimation than the network distances.
As mentioned before, proximity does not discriminate between services that a person may choose as a destination. It only estimates the distance to the nearest facility. However, a person may travel more for shopping if the destinations have more shopping options. The opportunity accessibility measurement captures how many services are within a specified distance. The score for services can be different depending on the type of service such as the area of a retail store or the number of beds for a hospital. The neighborhood or the surrounding distance can be Euclidean.
distance, Manhattan distance or network distance. The Euclidean and Manhattan based distance opportunities can be executed using the density or focal statistics tools provided by ESRI. This paper provides a customized ArcGIS tool to estimate opportunity based on network distance or driving time. The cumulative opportunity score within the specified distance from a parcel is assigned as an opportunity score for that parcel (or cell in the raster). Figure 6 is an example showing retail density within one mile of driving distance.

Figure 6 Network based Opportunity Raster
The distance and opportunity tools are embedded in the LUCIS model used in the land use modeling for the Heartland 2060 region. The Heartland region contains seven rural counties in central Florida. Based on the accessibility and mobility literature, the project team decided to use the tools in their effort to plan for accessibility instead of planning for mobility. The first level of the tool application was in the suitability model. Figure 7 shows the output of the urban suitability model after applying the tools. The differences between the urban suitability model before and after the application of the tools is the use of network based proximity instead of Euclidean distance and the use of the network based opportunity instead of the density functions.

Figure 7 Urban suitability for Heartland 2060 region
The third tool which is the gravity accessibility tool is used for the allocation of population scenarios. The planning for the Heartland region’s future population allocation incorporates running multiple scenarios using the LUCIS$^+$ and the CEM method. These scenarios include transportation, economic and sea level rise population migration scenarios.

In scenario building, gravity accessibility is very important. For example if we take the economic development scenario, accessibility to major activity centers, logistic terminal and ports are very important. As the travel distance increases accessibility becomes more dependent on mobility and transportation networks. Additionally, people’s choice of destination depend on combined attraction, distance and friction are more evident for larger distances. Therefore, the gravity access tool is one of the main tools used to prioritize the allocation of future population using the CEMs. Accessibility is estimated using the gravity model as explained methods section. An example on this accessibility estimation is to generate an accessibility surface for major retail and commercial activities. In this variable estimation, the major activities are identified, which are the major activities in retail and commercial. Then the resultant activities are integrated within walking distances to reduce the number of major activities. However, the value of attractions is accumulated in the integrated points. To reduce the size of the origin-destination matrix, the random points are used as origin points and the major activities are taken as destination points.

The resultant origin-destination matrix that contained the network distances generated by ArcGIS Network Analyst is joined to major activities to calculate the attraction for each trip. The trips of each parcel to all activities are used to estimate the
accessibility of the parcels to major activities using the gravity access mathematical equation explained earlier in the paper.

Figure 8 Network Gravity Access

The three tools were validated for accuracy. The tools depend on two main estimation steps. The first is to estimate the accessibility score for the each point in the population. This can be also represented by the random points for increasing the processing speed. For the network proximity access, the accessibility score is the
network distance to the nearest destination and can be computed by ArcMap Network Analyst for any point.

The second tool is the opportunity tool. The accessibility score in the opportunity tool is the density inside a network driving shed. This also can be generated by Network Analyst and the service are can be used to capture the density of services inside the network buffer.

For the third tool, the accessibility score is applying the gravity equation on network distance and the attraction as well as the decay factor. These are estimation steps that should not have any errors. To make sure the estimation is error-free a sample from the random points were taken and estimated using the ArcGIS interface and Network Analyst and compared with the corresponding list generated as an intermediate step by the tools and the two estimations were identical. However, there is a prediction step used in the tool in its final stages. This step is the Inverse Distance Weighting (IDW). The IDW is an exact interpolator, which means that it does not change the value for the input points. Theoretically, the more points used for the interpolation, the less interpolation errors will be generated. The tools were used to generate an access surface for many locations. The number of random points used varies according to the study area size. For the Heartland region in Florida, 50 thousands random points are used to cover seven counties. Other areas like Hillsborough County in Florida 10,000 were used. The prediction errors in all the cases were very small. However, the number of random point as well as the distribution of them is generated by the user and therefore establishing a standard prediction error may not be an easy task.
Conclusion, recommendations and future research

Accessibility can be estimated using different mathematical equations. The tools created in this paper demonstrate three main types of accessibility estimations which are: proximity, opportunity and gravity. The tools use road network in the estimation of accessibility scores. Therefore, distance or time can be used as a cost field in the analysis. The tools also can be updated to include congestion and barriers in the road network. The choice of the suitable tool for estimating accessibility depends mostly on the study area and why the accessibility score is being estimated. For example, investigating the National Household Travel Survey data in Florida (NHTS 2010), it is not difficult to conclude that frequency of shopping trips is decreasing if the travel distance is increasing. This is true for shopping trips that are more than two miles. This means that if the shopping trip distance is less than two miles, people may choose to shop in a further shopping center within the two-mile shed. This makes the use of the opportunity estimation more suitable for capturing neighborhood access.

The comparison between the three methods of estimation is complicated and may not be practical. In the Heartland region, the three tools where used in the suitability models. For the neighborhood and local accessibility level, a combination of network-based opportunity and network-based proximity is used in the suitability modeling. In suitability modeling the nearer to the shopping center, the more suitable the location will be for residential locations which is a direct application of the network proximity tool. In terms of a commercial location, the higher the density of commercial footage, the more suitable the location for commercial uses. The latter is a direct application of the opportunity tool. For regional destinations, the choice of locations to shop depends on the attraction of the destination as well as the travel time or distance, which is an
application of the gravity model, a direct application of the third tool. However, the traveler may not travel 50 miles for shopping even if the attraction of the destination is very high. That increases the need for decay and friction factors in the equation. For regional destinations also, accessibility is more dependent on mobility (Hanson, 2004) and that makes the need for including capacity, barriers and congestion more important in an accessibility estimation.

The tools created in this paper are parcel-level population tools that should not have any estimation error. The user can use all the parcels in the study area as origins and the tools will create an accessibility score from each origin to destinations and create an accessibility raster from the values of the access score for each parcel. This is a time consuming process and its success depends on the processor speed, the memory of the computer as well as software limitations. The use of a random sample can reduce the estimation time. At the same time it will induce some estimation error due to the interpolation of values using the origin location. To make the tool flexible, the use of population or sample data is optional so that the user can decide on the trade-off between time and accuracy.

The tools are automated tools and simple to use. Because the tools are based on the network distance and travel time, it is recommended to use these tools in land use modeling instead of the proximity tools that are based on Euclidean distance. However, it should be noted that the proximity to noise or pollutants are not dependent on network distance and the use of Euclidean distance should continue for these cases. The tools can be applied on distance and travel time. Including barriers and congestion in the tools will be studied in future research.


