**Transportation and Accessibility**

**Introduction**

Accessibility is a key element to transport geography, and to geography in general, since it is a direct expression of mobility either in terms of people, freight or information.

**Defining Accessibility**

Mobility is a choice made by users and is, therefore, a mean to evaluate the impacts of infrastructure investment and related transport policies on regional development. Well-developed and efficient transportation systems offer high levels of accessibility, while less-developed ones have lower levels of accessibility. Thus, accessibility is [linked with an array of economic and social opportunities](https://transportgeography.org/?page_id=6950), but congestion can also have a negative impact on mobility.

Accessibility is the measure of the capacity of a location to be reached by, or to reach different locations. Therefore, the capacity and the arrangement of transport infrastructure are key elements in the determination of accessibility.

All locations are not equal because some are more accessible than others, which implies inequalities. Thus, accessibility is a proxy for spatial inequalities. The notion of accessibility consequently relies on two core concepts:

* The first is location where the relativity of space is estimated in relation to transport infrastructures since they offer the means to support mobility. Each location has a set of referential attributes, such as its population or level of economic activity.
* The second is distance, which derived from the physical separation between locations. Distance can only exist when there is a possibility to link two locations through transportation. It expresses the friction of distance and the location which has the least friction relative to others is likely to be the most accessible. Commonly, the friction of distance is expressed in units such as in kilometers or in time, but variables such as cost or energy spent can also be used.

There are [two spatial categories applicable to accessibility problems](https://transportgeography.org/?page_id=6957), which are interdependent:

* The first type is known as topological accessibility and is related to measuring accessibility in a system of nodes and paths (a transportation network). It is assumed that accessibility is a measurable attribute significant only to specific elements of a transportation system, such as terminals (airports, ports or subway stations).
* The second type is known as contiguous accessibility and involves measuring accessibility over a surface. Under such conditions, accessibility is a cumulative measure of the attributes of every location over a predefined distance, as space is considered in a contiguous manner. It is also referred to as isochrone accessibility.

Last, accessibility is a good indicator of the [underlying spatial structure](https://transportgeography.org/?page_id=6962) since it takes into consideration location as well as the inequality conferred by distance to other locations.



Relationship between Distance and Opportunities

Accessibility is a determining factor behind the availability of opportunities (jobs, customers, suppliers, etc.) and if they can be realized or not. In a high accessibility setting, an individual will have access to a wider array of goods and services, employment as well as additional social interactions. The same applies to a business with potentially more customers and suppliers. Keeping accessibility constant, density is also a factor impacting on opportunities. In a high density setting, a distance will confer more opportunities than the same distance on a low density setting.



Accessibility can be measured in two different ways:

* Topological. Considers a system of nodes linked by transport infrastructures. In this case, accessibility is calculated at the nodal level and a function of the network structure. For seven nodes (a to g) located at an equal distance of one another, node d is the most accessible because it represents the minimal summation of total distances with all other nodes. Accessibility is measured only for nodes, while the intervening spaces are not considered outside the distance they represent.
* Contiguous. Considers a continuous space, here represented as a grid where each cell was assigned a level of accessibility. In this case, accessibility is a function of the spatial structure. Although accessibility values are here qualitative (ranking from least to most) a quantitative value can also be allocated for each cell.



**Accessibility and Spatial Structure**

Due to different spatial structures, two locations of the same importance can have different accessibility levels. The above figure presents three cases that compares the differences in accessibility of two locations according to the variations in the spatial structure.

* **(A) Uniform distribution**. For a spatial structure where locations are uniformly distributed, locations 1 and 2 have different accessibility levels, with location 1 being the most accessible. As distance (Euclidean) increases, location 1 has access to a larger number of locations than location 2. To access all locations, location 2 would require about double the traveled distance than location 1.
* **(B) Clustering in central area**. In this case, which is reflective of the distribution of urban populations, the number of locations that can be reached by location 1 increase rapidly and then eventually peaks. Location 1 has a clear accessibility advantage over location 2.
* **(C) Clustering in periphery**. Although the number of locations that can be reached from location 2 initially increases faster than for location 1, it catches up and is actually the most accessible, but by a lesser margin.



Space – time convergence is far from being a uniform process as differences in transport infrastructures and basic landscape constraints have a discriminatory effect on accessibility. The above figure represents travel time, from less than 1 hour to 10 days, to the nearest city of more than 50,000 people. It is the outcome of an overlay of several friction of distance factors, including the road and rail networks, navigable rivers, shipping lanes and land cover. It can be considered as a proxy for global accessibility with only 10% of the world’s population being more than 48 hours away from a large city. While it depicts the general ease of accessing urban markets, it does not depict well the effectiveness of global freight flows. For instance, while South Asia appears highly accessible because of the density of large cities, the quality and capacity of inland transport infrastructures is generally poor.

**Connectivity and Total Accessibility**

The most basic measure of accessibility involves **network connectivity** where a network is represented as a [connectivity matrix](https://transportgeography.org/?page_id=7620) (C1), which expresses the connectivity of each node with its adjacent nodes. The number of columns and rows in this matrix is equal to the number of nodes in the network and a value of 1 is given for each cell where this is a connected pair and a value of 0 for each cell where there is an unconnected pair. The summation of this matrix provides a very basic measure of accessibility, also known as the **degree of a node**:



* C1 = degree of a node.
* cij = connectivity between node i and node j (either 1 or 0).
* n = number of nodes.

The connectivity matrix does not consider all the possible indirect paths between nodes. Under such circumstances, two nodes could have the same degree but may [have different accessibilities](https://transportgeography.org/?page_id=6978). To consider this attribute, the [Total accessibility matrix](https://transportgeography.org/?page_id=6983) (T) is used to calculate the total number of paths in a network, which includes direct as well as indirect paths. Its calculation involves the following steps:



Thus, total accessibility would be a more comprehensive accessibility measure than network connectivity.







**The Shimbel Index and the Valued Graph**

The main focus of measuring accessibility does not necessarily involve measuring the total number of paths between locations, but rather what are the shortest paths between them. Even if several paths between two locations exist, the shortest one is likely to be selected. In congested networks, the shortest path may change according to the current traffic level on each segment. Consequently, the Shimbel index calculates the **minimum number of paths** necessary to connect one node with all the nodes in a defined network. The [Shimbel accessibility matrix](https://transportgeography.org/?page_id=6990), also known as the **D-Matrix**, thus includes for each possible node pairs the shortest path.

The Shimbel index and its D-Matrix fail to consider that a topological link between two nodes may involve variable distances. It can thus be expanded to include the notion of distance, where value is attributed to each link in the network. The [valued graph matrix](https://transportgeography.org/?page_id=6996), or **L-Matrix**, represents such an attempt. It has a very strong similarity with the Shimbel accessibility matrix and the only difference lies that instead of showing the minimal path in each cell, it provides the **minimal distance** between each node of the network.



The Shimbel Distance Matrix (or D-Matrix) holds the shortest paths between the nodes of a network, which are always equal or lesser to the diameter. To construct this matrix, C matrices of Nth order are built until the diameter of the network is reached. Each C matrix is converted in a corresponding D matrix. In this case, two C matrices, C1(connectivity matrix) and C2 (two-linkages paths; C1\*C1) are built since the diameter is 2.

* **The first order Shimbel Matrix**(D1) is a simple adaptation of C1, where all the direct links are kept. A value of 0 is assigned for all the cii cells since the shortest path between a node and itself is always 0. Cells that have a value of 0 in the C1 matrix (outside cii cells) remain unfilled on the D1 matrix.
* **The second order Shimbel Matrix**(D2) is built from the first order matrix D1 but only from its unfilled cells. A value of 2 is assigned for each cells on the D2 matrix that have a value greater than 0 on the C2 matrix, but if a value of 1 already exists (D1 matrix), this value is kept. This means that on the D2 matrix of the above figure, only the values of the yellow cells have been changed to 2. Since the diameter of this network is 2, the **D2 matrix is the Shimbel distance matrix**.
* **Nth order Shimbel Matrix** (DN). For a network having a diameter of 3, a D3 matrix would have to be built from a C3 matrix (C1\*C2) because at least 1 cell would have remained empty in the D2 matrix. Repeat the construction of Nth order Shimbel matrices until the diameter is reached.
* **The Shimbel Matrix (D)**. The order of the Shimbel distance matrix that corresponds to the diameter is the D matrix. The summation of rows or columns represents the Shimbel distance for each node. In the D matrix of the above example, node C is having the least summation of shortest paths (4) and is thus the most accessible, followed by node A (5), nodes B and D (6) and node E (7). The total summation of minimal paths is 28.



The construction of the valued graph matrix (L-matrix) follows the following procedure:

* **The distances in the network are transcribed in matrix L1** (direct connectivity distance) for each pair directly connected. An infinite value is given for pairs not directly connected.
* **Calculation of the Nth order L matrix**. The operation is similar to the creation of the Shimbel Matrix. What differs in this case is that we are not working with the minimum number of paths, but with the minimal distance, which could give different results. The shortest path between node A and B is obviously the A-B link. However, there is also a A-C-B link, which summation of distances could be smaller (actually it is not). The calculation of the L2 matrix requires the cross-summation of the L1 matrix where each cell in a column is added with each cell in a row. The B-A cell on matrix L2 is thus calculated by the cross summation of column B and row A. Only the smallest value of the five operations is kept, which is 10 in this case. Since the above network has a diameter of 2, only two steps are necessary and the L2 matrix becomes the L-Matrix. The summation of each row on the L2 matrix represents the minimal distance required to reach all the other nodes in the network. For node B, it is 43.

# Geographic and Potential Accessibility

From the accessibility measure developed so far, it is possible to derive two simple and highly practical measures, defined as geographic and potential accessibility. [Geographic accessibility](https://transportgeography.org/?page_id=7001) considers that the accessibility of a location is the **summation of all distances** between other locations divided by the number of locations. The lower its value, the more a location is accessible.



This measure (A(G)) is an adaptation of the Shimbel Index and the Valued Graph, where the most accessible place has the lowest summation of distances. Locations can be nodes in a network or cells in a spatial matrix.

Although geographic accessibility can be solved using a spreadsheet (or manually for simpler problems), **Geographic Information Systems** have proven to be a very useful and flexible tool to measure accessibility, notably over a surface simplified as a matrix (raster representation). This can be done by generating a distance grid for each place and then summing all the grids to form the total summation of distances (Shimbel) grid. The cell having the lowest value is thus the most accessible location.

Potential accessibility is a more complex measure than geographic accessibility since it includes simultaneously the **concept of distance weighted by the attributes of a location**. All locations are not equal and thus some are more important than others. [Potential accessibility](https://transportgeography.org/?page_id=7007) can be measured as follows:



The potential accessibility matrix is not transposable since locations do not have the same attributes, which brings the underlying notions of emissiveness and attractiveness:

* **Emissiveness** is the capacity to leave a location, the sum of the values of a row in the A(P) matrix.
* **Attractiveness** is the capacity to reach a location, the sum of the values of a column in the A(P) matrix.

Likewise, a Geographic Information System can be used to measure potential accessibility, notably over a surface.



The construction of a geographic accessibility matrix, A(G), is a rather simple undertaking:

* **Build the valued graph matrix** (L). The above L-matrix shows the shortest distance in kilometers between five nodes (Node A to Node E).
* **Build the geographic accessibility matrix** A(G). The A(G) matrix is similar to the L-matrix expect that the summation of rows and columns is divided by the number of locations in the network. The summation values are the same for columns and rows since this is a transposable matrix. The most accessible place is Node C, since it has the **lowest summation of distances**.



By considering the same valued graph matrix (L) than the previous example and the population matrix P, the **potential accessibility matrix**, P(G), can be calculated:

* The value of all corresponding cells (A-A, B-B, etc.) equals the value of their respective attributes (P).
* The value of all non-corresponding cells equals their attribute divided by the corresponding cell in the L-matrix.

The higher the value, the more a location is accessible, node C being the most accessible. The matrix being non-transposable, the summation of rows is different from the summation of columns, underlining their respective attractiveness and emissiveness. Node C has more emissiveness than attractiveness (2525.7 versus 2121.3), while Node B has more attractiveness than emissiveness (1358.7 versus 1266.1).