Urban Networks, Analysis and Planning

Affiliation
Andres Sevtsuk
City Form Lab, Singapore University of Technology and Design, Singapore.
asevtsuk@alum.mit.edu

Synonyms
Spatial networks, urban form, GIS, accessibility, centrality, urban design, city planning.

Glossary
Urban Form: The physical pattern of urban infrastructure and buildings.
Land Use Pattern: The spatial distribution of human activities accommodated within urban form.
Built Environment: A combination of urban form and land-use mix of an area.
Accessibility: Property of a location that describes the ease with which the location can be accessed from surrounding urban form and land-use attractions.
Centrality: refers to metrics that describe how centrally an event is located in a spatial network (see Centrality Measures 00227).

Introduction
Network analysis concepts have been used in the design and planning of cities for several decades. Until recently, however, they were common in only highly specialized applications — disaster planning problems, critical facilities’ location problems, and costly utility and transportation infrastructure design problems. Efforts to apply network analyses to the design of ordinary buildings, public spaces and urban districts go back to the 1960s, but only in the recent decade have the necessary tools and data for their widespread use become available to architects and planners.

Most work in urban network analysis has relied on methods that were originally developed for social networks. There are, however, important adjustments in applying social network analysis methods on urban space. In social networks, for instance, connections between network elements are generally described topologically — as degrees of separation between people in a network, for example — where geography and geometry of relationships have little importance. Scholars of the built environment, on the other hand, are more often interested in precise geographical relationships of a spatial network, where distances, angles and travel times are critical to describing adjacencies and proximities between places. Second, whereas in social networks the weighting of network elements — people — according to personal characteristics has until recently been rare, weighting is often critical in spatial network studies. A street lined with small single-family homes has a different effect on a neighbourhood than a street lined with high-rise office buildings. These particularities have led researchers to customize both the representation of urban networks and the metrics applied thereon.

Historical Background
The spread of graph theory among planners in the latter part of the 20th-century was catalyzed by the appearance of numerous applied graph theory publications after the Second World War [1,2]. Spatial applications of graphs were quickly adopted in transportation research, where the precedent for applying graph measures to large-scale road and rail networks was first established [3,4,5,6]. Architects soon also adopted graph representation for the study of building plans [7,8,9,10], typically representing each room by a node and the availability of a direct circulation connection between two rooms by a link (Figure 1). Representing buildings with graphs opened up new opportunities for distinguishing common layouts of architectural plans using graph indices, some of which are discussed below [11].

Figure 1. Adjacency graph for Frank Lloyd Wright’s Aline Devin House. Source: (March and Steadman 1971: 259-261).

Hillier & Hanson applied the representation and analytic tools of graphs to street networks establishing the now well-known Space Syntax methodology [12,13]. Over the last three decades, their work has argued that the spatial configuration of street networks is related to diverse social phenomena including the rates of pedestrian flow, the geography
of crime, and the distribution of business establishments.

Hillier and Hanson have chosen to represent streets not with centerlines, as in most transportation studies, but rather with axial lines. Axial lines are defined as the fewest and longest lines of sight that can be drawn through the open street spaces of a study area [12]. This approach has led to some criticism, since the specification of axial lines is subjective (there is more than one solution), and poorly applicable to sparsely built-up streets [14]. Unlike typical transportation applications of graph theory, Space Syntax researchers have also adopted a so-called dual graph representation, where streets are represented as nodes and intersections as edges. Since most graph theory indices have been designed to focus on the properties of nodes (e.g. in social networks, nodes can represent people), this inverted form of graph representation allows the Space Syntax analysis to focus on streets (axial lines). Whereas 'degree centrality' in social networks indicates how many direct links (e.g. kinship ties or acquaintances) connect to a node of interest (e.g. a person), an analogous measure in Space Syntax describes the number of neighboring axial lines that intersect with a particular axial line of interest.

Though useful for centering the analysis on streets, the dual representation also introduces a well-known problem to the Space Syntax methodology. If streets are represented as nodes, then both long and short streets alike reduce to dimensionless points, thus effectively eliminating metric distance from the analysis. Space Syntax researchers address this problem by measuring travel from one line to another across the graph in topological terms, using the count of lines traversed (i.e. degrees of separation) as a metric of proximity. This metric, commonly referred to as depth, is central to most Space Syntax analysis. It is used as a kind of distance measure, which represents the minimum number of axial lines needed to go from an origin to any other axial line in the network. The depth measure leads to another central metric in Space Syntax literature: integration [13]. The integration measure is simply a relative description of each axial line's depth with respect to all other axial lines in the graph (Figure 2). It is obtained by repeating the depth measure from each line to all other lines in the system and normalizing the obtained sums for each line by the total number of lines in the graph. In mainstream network analysis terms the integration analysis is analogous to the closeness metric, with the difference that distance is being calculated on the basis of topological turns instead of metric units. If integration is computed with a radius of only one turn (also referred to as one step in Space Syntax literature), then the result simply shows how many axial lines intersect with a given line of interest, analogous to the familiar degree centrality of nodes in graph theory.

![Figure 2. (a) Plan drawing of Gassin, a hill-town in Southern France. (b) Axial lines overlaid on its street network. Source: (Hillier and Hanson, 1984).](image)

Several other approaches to graph analysis of street networks have appeared in the recent years. Among those, Porta and Xie have implemented a number of spatial graph indices that rely on primal representation of spatial networks [15,16]. A number of freely accessible software tools have been developed to operationalize spatial network analyses including the Axwoman toolbox [17]; the SANET toolbox [18,19], the Urban Network Analysis Toolbox [20] and other custom built applications for GIS [21,22]. In the following we will primarily rely on the approach introduced by Sevtsuk and Mekonnen [20], which allows us to describe urban spaces of multiple morphologies and scales using a representational framework that is common in transportation studies.
Representational Framework

Most spatial network studies to date have represented networks using two types of network elements – nodes and edges. In the case of urban street networks, edges typically represent street segments, and nodes the junctions where two or more edges intersect [15]. As already discussed, the Space Syntax approach inverts these elements. The Urban Network Analysis (UNA) framework [20,23] has recently introduced two important modifications to this framework. First, it adds buildings (or other location instances, such as land parcels, transit stations etc.) to the representation, adopting a tripartite system that consists of three basic elements: edges, representing paths along which travelers can navigate; nodes, representing the intersections where two or more edges intersect; and buildings, representing the locations where traffic from streets enters into indoor environments or vice versa. The unit of analysis thus becomes a building, enabling the different graph indexes to be computed separately for each building. Should the analyst wish to compute the graph centrality measures for nodes of the network instead of buildings, then the nodes themselves can be used as inputs instead of buildings. This allows a user to account for both uneven building densities and land use patterns throughout the network, neither of which are addressed in most previous urban network analysis methods. The UNA representation assumes that each building connects to a street segment (edge) that lies closest to it along the shortest perpendicular connection. This network representation framework is illustrated in Figure 3. The left side of the figure presents a fragment of Harvard Square in Cambridge, MA in plan drawing. The same plan drawing is shown in graph form on the right.

If the spatial configuration of the environment under study cannot be represented in a two-dimensional graph – as may be the case if the network contains underpasses, overpasses, or three-dimensional circulation routes inside buildings – then a similar graph can also be represented three dimensionally, using vertical z-axis values on each of the network elements.

The three-element representation of spatial networks is well suited for mapping urban and regional networks of various typologies and scales. At the finest architectural scale, the third network element can be used to represent individual establishments or rooms within buildings (Figure 4). Each establishment, room or floor of a building can be described with an appropriate weight that captures the quantity or quality of activities it houses. At a larger scale, units of analysis can instead represent whole buildings, allowing the user to weigh the analysis by substantive attributes of each building (Figure 3). At an even larger scale the units of analysis can become whole city blocks that contain multiple buildings, or zip codes that include multiple blocks. The third network element thus offers a flexible container for analyzing the urban built environment at various scales with consistent methods.

It is important to aggregate the units of analysis at substantively justified levels. Numerous studies have shown that the choice of aggregation can itself affect analysis results. This issue, which has become known as the Modifiable Areal Unit Problem or MAUP in

![Figure 3. Left: Plan drawing of Harvard Square in Cambridge, MA. Right: graph representation of the same plan. Source: author.](image-url)
literature [24], is defined as “a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomenon resulting in the generation of artificial spatial patterns” [25]. In order to avoid MAUP issues, it is important to favor behaviorally justified choices of aggregation to ad-hoc choices of aggregation. One sensible way of aggregating the units of analysis in urban networks is to choose aggregation levels according to spatial control boundaries [26]. Individual rooms in a building, for instance, have a separate control structure from the building as a whole – changing access to the room does not affect access to the entire building. But changing access to the building does affect access to all rooms within the building. It is often convenient to graphically mark the third network elements at the entrances of spatial control boundaries (e.g. doors) of the units of analysis (e.g. business establishments), as shown in Figure 3. Just like rooms in a building, individual business establishments can have autonomous spatial control boundaries, as can floors in a building, buildings in a block, and blocks in a district. It can be practical to keep track of different aggregation levels by adding an extra digit to the location ID at each successive aggregation level (e.g. building ID = 26; floor id = 26-3; room ID = 26-3-1). Aggregation allows each of these elements to be both a parent and a child to other elements. Using spatial control boundaries as a basis for aggregation avoids the hazard of biasing analysis results to arbitrary data groupings, makes the collected data useful for cross-scalar analysis and the results easy to explain to professionals of different disciplines.

Metrics

We can broadly distinguish two types of network analysis indexes on urban networks. The first type – inter-network indices – capture the properties of a spatial graph or sub-graph as a whole [27,28]. Their results become meaningful if compared to other networks. The second type – intra-network indices – characterize the relative relationship of each network element (e.g. node or building) to other surrounding elements in that network.

Inter-Network Indices

Inter-network indices can be used to analyze the overall properties of an area’s spatial network. These indices are most commonly applied to traditional two-element networks consisting of nodes and edges, where nodes can represent street intersections and edges street segments [11]. But they can be equally
applied to the circulation layouts of buildings or to three-element networks that contain other network instances. In this case, each building or other network instance can be represented as a node that is connected to the street-network via a link to the nearest street segment.

Figure 5. Diamond graph with 6 nodes and 15 edges.

The Gamma Index illustrates the extent to which a spatial network resembles a fully connected diamond graph, where each node is directly connected to every other node in the graph (Figure 5). It is calculated as follows:

\[
\text{Gamma Index} = \frac{e}{(v^2-v)/2}
\]

where \( e \) is the number of edges and \( v \) is the number of vertices in the graph. The higher the index, the higher the internal connectivity of the observed street network and the more directly a traveller can commute between intersections in the street network.

The Cyclomatic Number forms another index, which shows the availability of alternative, rather than unique routes between nodes in the network. A cycle is analogous to a hole in a fishnet and the index is defined as follows:

\[
\text{Cyclomatic Number} = e - v + g
\]

where \( g \) the number of connected components in the network (Figure 5).

Figure 6. An urban grid of four city blocks, where \( e=12; v=9 \). Gamma index = 1/3; Cyclomatic Number = 4; Maximum Cycles = 28; Redundancy Index = 1/7.

In urban street layouts, a grid of four blocks surrounded by streets on all sides produces four cycles. A cul-de-sac network, on the other hand has no cycles at all, and is referred to as a tree graph. Tree graphs always have \( v-1 \) edges and therefore only a single shortest path is available between any pair of nodes in a tree (Figure 7). From an urban design perspective, this means that tree networks, which are commonly seen in suburban settings, require the least asphalt to connect a given set of locations. But they also favour hierarchical patterns of organization and limit choice in travel paths.

Figure 7. A ‘tree’ network, where \( e=21, v=22 \). Gamma index \approx 0.09; Cyclomatic Number = 0; Maximum Cycles = 210; Redundancy Index = 0.

The maximum number of cycles for a given number of vertices is calculated as maximum possible edges in a graph minus edges in a tree graph with the same number of vertices:

\[
\text{Max. Cycles} = \frac{(v^2-v)/2}{(v^2-v)/2 - v + 1}
\]

The Redundancy Index shows the vulnerability of the network to divisions. The index is defined as a ratio between the number of observed cycles and the number of maximally possible cycles:

\[
\text{Redundancy Index} = \frac{(e-v+g)}{[(v^2-v)/2 - v + 1]}
\]

When the Redundancy Index is zero, then the network is one of several trees, when the Redundancy Index is one, then the network is totally connected (e.g. diamond). The index can be used to study how volatile a street network is to ruptures that may be caused by natural disasters, such as floods, mudslides etc. Different indices describe different properties of a network and a careful, hypothesis-driven choice of indices can lead to a complementary set of metrics that provide a holistic description of the urban area under study (see Network Representations of Complex Data 00012).

**Intra-Network Indices**

Intra-network indices describe the relative importance of individual nodes or buildings in a network. These indices allow us to compare how well different elements of the same network are connected to the rest of the network. In the following we briefly outline six indices that are readily applicable to the
analysis of urban infrastructure networks. The open-source Urban Network Analysis Toolbox allows an interested user to implement most of these indices in ArcGIS [20].

The degree centrality of a street intersection indicates the number of street segments that intersect at the given node. The vast majority of intersections in cities have three or four intersecting street segments. A typical intersection at the corner of an avenue and a street in the Manhattan grid, for instance, has four intersecting edges. But the maximum number of intersecting streets can be much higher. The Charles de Gaulle Etoile (a.k.a. Arc de Triomphe) intersection in Paris forms an atypically connected node, where twelve radial streets collide. One of the most admired public spaces of Europe – the Campo di Siena in Italy (Figure 8) – also draws foot-traffic from twelve paths that enter the Campo. But even nodes with five or six streets tend to stand out as exceptional places in urban circulation networks. Such atypical intersections have been found to be memorable in people’s mental maps of a city [29,30].

![Figure 8. Left: Plan of Charles de Gaulle Etoile traffic square in Paris. Right: Plan of the pedestrian Campo di Siena in Italy. The Degree centrality of both squares is 12. Source: Google Maps.](image)

While the degree centrality captures the connectivity of a node to its immediately adjacent street segments, a number of centrality indices have also been designed to describe a location’s connectivity to a set of destinations within a larger access radius. The *Reach* index describes the total number of destinations that can be reached within a given network radius from any street intersection or building. The reach centrality \( \text{Reach}^r[i] \) of a node \( i \) in a graph \( G \) at a search radius \( r \), describes the number of other nodes in \( G \) that are reachable from \( i \) at a shortest path distance of at most \( r \). It is defined as follows:

\[
\text{Reach}^r[i] = \sum_{j \in V(G) - \{i\} : d[i,j] \leq r} W[j]
\]

where \([i,j]\) is the shortest path distance between nodes \( i \) and \( j \) in \( G \), and \( W[j] \) is the weight of a destination node \( j \). The weights can represent any numeric attribute of the destination buildings – their size, the number of employees they contain, the number of residents they accommodate etc. Using weights allows the analyst to compute how many of such attributes (e.g. residents, jobs) can be reached from each building within a given network radius. The choice of input network and search radius \( r \) allow the user to model the index from the perspective of different transportation modes (e.g. walking, biking, driving). The measure may be interpreted as an alternative to areal density measures (e.g. households per acre, or jobs per square kilometer). It accounts for opportunities that are reachable along the actual street network as perceived by a pedestrian, bike or vehicle, producing a unique result for each origin location.

Whereas the *Reach* measure simply counts the number of destinations around each building within a given search radius (optionally weighted by building attributes), the *Gravity* measure additionally factors in...
the spatial impedance required to reach each of the destinations [31]. The gravity index, \( Gravity^r[i] \) of a node \( i \) in graph \( G \) at a radius \( r \) postulates that centrality is inversely proportional to the shortest path distance between \( i \) and each of the other nodes in \( G \) that are reachable from \( i \) within a geodesic distance \( r \). It is defined as follows:

\[
Gravity^r[i] = \sum_{j \in V(G) - \{i\}, d[i,j] \leq r} \frac{W[j]}{e^{\beta d[i,j]}}
\]

where \( \beta \) is the exponent that controls the effect of distance decay on each shortest path between \( i \) and \( j \) and \( W[j] \) is the weight of a particular destination \( j \) that is reachable from \( i \) within the radius threshold \( r \). If the buildings in \( G \) are weighted, then the Gravity measure is proportional to the weight of each of the other buildings that can be reached within the given search radius.

The exponent \( \beta \) in the Gravity Index controls the shape of the distance decay function, that is, how strongly the distance between \( i \) and its neighboring destinations \( j \) affects the result. The specification of \( \beta \) should thus be set according to the mode of travel assumed in the analysis (e.g. walking, cycling, driving), as well as the units of distance measurement. An empirical study of pedestrian trips to convenience stores in Oakland, CA by Handy and Niemeier [32] has suggested that for walking distances, measured in minutes, \( \beta \) is approximately 0.1813. The Gravity index offers a powerful measure that combines the number of destinations, the attractiveness of the destinations, and the travel costs of reaching them into a single value.

The Reach and Gravity indices describe how conveniently each location can be accessed from a set of surrounding locations. For some purposes, however, it may be more important to estimate the ease with which a location can be accessed en route while travelling between other locations. Newspaper kiosks, for instance, might find it less desirable to locate at places that are closest to people’s homes or jobs and more desirable at places where people tend to pass by while travelling between other destinations. The potential of passersby at different locations of a spatial network can be estimated using a Betweenness measure [33].

The betweenness centrality, \( Betweenness^r[i] \), of a building \( i \) in graph \( G \) counts the number of times \( i \) lies on shortest paths between pairs of other reachable buildings in \( G \) that lie within the network radius \( r \). If more than one shortest path is found between two buildings, as is frequently the case in a rectangular grid of streets, then each of the equidistant paths is given equal weight such that the weights sum to unity. Betweenness, in the context of spatial networks, is thus defined as follows:

\[
Betweenness^r[i] = \sum_{j,k \in V(G) - \{i\}, d[i,j], d[i,k] \leq r} \frac{n_{jk}[i]}{n_{jk}} \cdot W[j]
\]

where \( n_{jk} \) is the number of shortest paths from building \( j \) to building \( k \) in \( G \), and \( n_{jk}[i] \) is the number of these paths that pass through \( i \), with \( j \) and \( k \) lying within the network radius \( r \) from \( i \), and \( W[j] \) is the weight of a particular destination \( j \). If the analysis is weighted by demographics of a certain type in the surrounding buildings, for instance, the Betweenness centrality can capture the potential number of passersby for that particular demographic at building \( i \).

Figure 9 illustrates the Betweenness measure applied on two common types of urban layouts: the grid and the cul-de-sac subdivision. For illustration purposes, both layouts have the same number of buildings. The analysis shows that the peak Betweenness values are twice as high in the cul-de-sac plan than in the grid, with the range of values also much wider in the former. Since the grid offers multiple routes between any pair of locations, not all paths need to pass a particular link or building. The Betweenness values are more equal and distributed in the grid, producing a lesser spatial hierarchy between different locations.

The Closeness centrality of a building \( i \) is defined as the inverse of the total distance required to reach from \( i \) to all surrounding destinations \( j \) within the given access radius \( r \) [34]:

\[
Closeness^r[i] = \frac{1}{\sum_{j \in V(G) - \{i\}, d[i,j] \leq r} (d[i,j] \cdot W[j])}
\]

The Closeness measure illustrates how close each of these locations is to all other surrounding locations within a given network access radius. The index is therefore best suited for analyzing the relative proximity of a set of locations to surrounding resources in a city.

Finally, the Straightness centrality \( Straightness^r[i] \) of a building \( i \) estimates how closely the shortest network distances between \( i \) and its surrounding buildings \( j \) that are reachable within radius \( r \), resemble straight Euclidean distances [35,15]:
$$Straightness[i]^r = \sum_{\delta[i,j] \in V(G)-i,j} \frac{\delta[i,j]}{d[i,j]} \times W[j]$$

where $\delta[i,j]$ is the as-a-crow-flies distance between buildings $i$ and $j$, $d[i,j]$ the shortest network distance between the same buildings, and $W[j]$ the weight of destination $j$. As a ratio between the Euclidian distance and the geodesic distance, Straightness can only be estimated if the units of impedance are in linear distance (e.g. miles), not time (e.g. minutes). The index can be used to estimate how directly a set of residential apartment buildings connect to their nearest bus stops, for instance.

**Areas of Application**

As urban networks typically represent the built environment in a way that ties activities and spaces together with a connecting web of travel paths, urban network analysis is well-suited to social, economic and transportation questions that involve accessibility and the movement of people, goods and information in a city. Unlike idealized Euclidian distances, proximity measurements on networks model the environment with the constraints imposed by the geometry of streets and other man-made or natural boundaries that closely approximate the human experience of navigating a city. When analyzing how close or accessible households are to public transit stations, for instance, conventional Euclidian buffer distances risk connecting places that are in reality separated by highways, water-bodies or fences, resulting in over-estimations of accessibility. An accurate network representation of the same area can integrate these constraints and provide a more reliable estimate.

Research has shown that urban network analysis measures can be useful predictors for a number of interesting phenomena. Porta and Sevtsuk have used network analysis to study the location patterns of retail commerce on urban street networks and found that retailers typically locate in places that are more “between” surrounding destinations [36], close to jobs, transit stations, and built density [23]. This research has developed a new direction for urban economics, moving from economic geography to economic geometry. A similar approach can be applied to studying the distribution of other land uses, land values, rents, commercial revenue and other spatial economic indicators.

A number of researchers have also studied how the geometry and topology of urban street networks affect pedestrian and motorized traffic patterns [37,13]. Using a third network element to describe buildings or establishments, shown above, allows an analyst to investigate how an addition of a new building or business affects the accessibility of existing buildings or businesses. It can be used to show, for instance, how much additional foot traffic a newly proposed development could add to individual street segments, or where opportunities might be created for new businesses.
**Future Trends**

From the point of view of planners, a central shortcoming of current network analysis methods is their inability to generate alternative geometric configurations. Most spatial network analysis approaches are good at quantifying existing geometric networks, but poor at suggesting alternative solutions that could improve an existing network with respect to given constraints [38,39]. Thus an analyst will typically test before and after scenarios of a proposed urban intervention and use network analysis to illustrate the improvements achieved by the proposed changes. It is less clear how the results could be used to improve the design. This is not a shortcoming of network analysis per se, but of all spatial analysis methods in general. Spatial analysis tools do not design, they analyze existing designs.

A potentially promising development on this issue has recently appeared in procedural urban models [40,41]. These models generate geometric configurations of urban form on a fly, based on a set of input parameters and are capable of illustrating the geometric results so as to achieve a more desirable combination of the input parameters. If one of the parameters of a residential development model is to achieve particular levels of network accessibility to surrounding jobs, for instance, then the model could iterate through a large number of possible geometric configurations and search for solutions that yield the highest home-to-jobs accessibility. This looks like a promising research direction in the coming years. But since all good urban design involves hundreds, if not thousands of variables that a designer intuitively balances, it is unlikely that any truly sophisticated generative models will appear in the near future.

**Cross-References**


**References**


