M-SA-20230301

Network Movement through Discrete Patronage Betweenness and Its Distribution Throughout Time

A part of Network Simulation based Analysis and Visualization

Kevin Sutjijadi @ 2023

Abstract

A derivative and modification of the set of tools from Urban Network Analysis (UNA), to simulate how much movement between set origins and destinations, and to extrapolate its density on a set of time given its distribution parameters.

Summary

Network Movement can be defined from the total amount on a cycle and its distribution. Calculation of said forms are separated into two calculations with the second dependent on the former.

As the function of the method is to calculate movement passing through on a part of a network, the object of the calculation is the segments, that are parts of paths that connect nodes/intersections. Said method will be mentioned as discrete patronage betweenness which is calculated by

$$P(i) = \sum_{j \in J} p_j \cdot \frac{\sum_{\{i\} \subset jk \in G, d_{jk} < (d_{jk\,min} \cdot r) < D} W_k \cdot \left(\frac{d_{jk\,min}}{d_{jk}}\right)^{1+\alpha}}{\sum_{jk \in G, d_{jk} < (d_{jk\,min} \cdot r) < D} W_k \cdot \left(\frac{d_{jk\,min}}{d_{jk}}\right)^{1+\alpha}}$$

Equation 1

i	= path segment
j	= origin point
p	= patronage capacity
J	= origins set
k	= destination point
K	= destinations set
G	= graph/network
jk	= paths(s) from origin to destination
d_{jk}	= distance from feature to location
d_{min}	= minimum distance of path
r	= detour ratio
α	= detour weighting exponent ratio
W	= weight value

Where the resulting value is the sum of function of paths that includes the segment divided by function of all possible paths, multiplied by the origin's weight. Meaning, that segments on all part of possible path alternative will equate up to complete weight, that proves the function.

Alternative paths, also known as detour ratio, is expressed by a float value of at least more than 1.0, where alternative pathways are still less than the closest path times detour ratio is still included in the summation.

A continuation for the Discrete Patronage Betweenness is its Distribution Throughout Time, where the number of sets of movement is distributed in a time frame as

$$F(i,h) = \sum_{m \in M} F_m(i,h) \qquad \textit{Equation 2}$$

Where it is the sum of the set of a distributed discrete patronage betweenness that is expressed as

$$F_m(i,h) = \int_{h-r}^{h+r} \frac{P_m(i)}{\omega_m} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\cdot\left(\frac{t-\xi_m}{\omega_m}\right)^2} \cdot \frac{1}{2}\left(1 + erf\left(\frac{\alpha \cdot \left(\frac{t-\xi_m}{\omega_m}\right)}{\sqrt{2}}\right)\right) dt$$

Equation 3

with the location (ξ) parameter defined as

$$\xi_m = \mu_m - \omega_m \cdot \left(\frac{2\alpha}{\sqrt{\pi(1+\alpha)^2}}\right) \qquad \textit{Equation 4}$$

i	= path segment
h	= time mark
r	= time spread of calculation
m	= a movement relation
M	= set of movement relation
P	= discrete patronage betweenness
ω	= distribution shape
ξ	= distribution location
α	= distribution skewness
μ	= distribution mean

Which is based on Normalized Skewed Distribution with added integral expression and weighting multiplication. These functions can be used to develop models of relations between features in an area generates movements, as well as help to extrapolate cases with known or set parameters on how the movement distributes within the area by amount, density, and consistent.

1. Introduction

Urban Network Analysis has become one of the more popular tools on urban analytics, as it provides a great set of tools and thinking that coincides with the new paradigm of planning and development, focusing on pedestrian movement and its further implications on structures, economies, and cities (Sevtsuk, 2021). This in-part has also been catalyzed by more accessible tools to try and demonstrate its capabilities on answering the related problems.

Adapted from social science, network analysis, is a segment of method to put relations into graph (Barrat, 2004). As the nature of streets, roads, circulation networks are already similar to the concept, Urban Network Analysis is used to tell relations between activity nodes, junctions, public space, and the movement of its users. This adaptation has been proven true, as centrality has been used as useful predictors of the importance of junctions in transportation networks (Sevtsuk, 2017; Garrison, 1960; Garrison and Marble, 1962; Kansky, 1963; Haggett and Chorley, 1969). As its application began to widen, both in its measurements and simulation scale, adjustments to set methods are required for larger scales of calculation to be feasible. As the current tools available are in the form of plugins of ArcGIS and Rhino (UNA Toolbox), although robust, presents limitations in computing scale and documentation of process.

With further application of the tool to become a way of measuring a region's quality, i.e. as formally acknowledged key performance indicators, a comprehensive but also simple-to-measure metrics will be developed. Speaking of movement, its well-known/intuitive basic unit is footfall (Ewing, Cervero, 2001), where it is a single person going through a defined segment in one way. This is suitable as a common metric as it is easy to measure and understand, where gravity and centrality metrics are more abstract and unmeasurable.

For the purpose of urban planning and design, the ability to calculate and manage to an extent the density and amount of footfall going through a segment will enable further detailed and accurate designs, where managing alternate paths, activity centers, and even unto pedestrian widths can be aided (Kim, 2010). Its distribution throughout time cycles are also important in that its consistency and characteristics (leisure or commuting), provides a clearer picture of how that region or area will operate and its significance throughout the region.

2. Discrete Patronage Betweenness

Calculating basic discrete footfall can be defined from the origin, how much capacity or discrete users it generates, plot their course into a set of destinations, each with different attraction weighting, and also the distances they have to cover to reach it on the circulation network. the closer the distance, the more luring it is, but with enough intrinsic attractiveness weighting, further features can also outweigh closer ones. To express this in mathematical terms, there are several factors and concepts to combine.

Before delving into calculations, as the base problem is centralized around graphs and network, pairs of origin-destination, the used distance are based within the graph, not by direct lines. Where graph theory is based on nodes and segments, spatial factors such as actual position and shape will be simplified as junctions become nodes, and polyline path segments will be extracted as straight segments with the value of segment distance weight, which can be modified by external weighting factors such as its quality, inclination, or other intrinsic values (Sevtsuk, 2010; Naismith, 1892; Scarf, Phillip, 2008). Then, possible paths have to be mapped using pathfinding algorithms, commonly using Dijkstra, A*, or other approaches (Dijkstra, 1959; Hart, Nilsson, Raphael, 1968).

Starting from the concept of betweenness, as it has been derived from Social Science, it expresses the importance of a node or a segment on how much connections pass through by the permutation of node pairs (Freeman, 1977), in the form of the function:

$$g(i) = \sum_{j,k \in G-\{i\}} \frac{n_{jk}[i]}{n_{jk}} \qquad \textit{Equation 5}$$

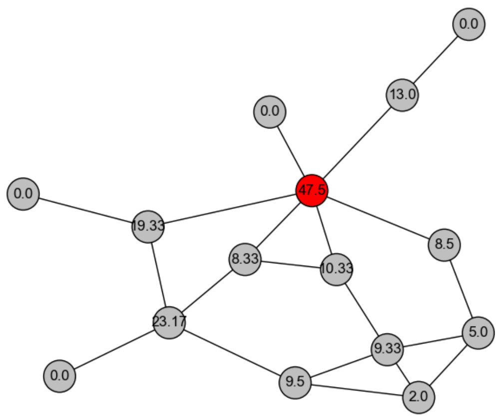

Figure 1 Betweenness Centrality by networkx (Source: Stack Overflow @SparkAndShine, 2016)

Where the segment or node value is the sum within permutation of a set of pairing of origin and destination within a graph of the number of possible paths that pass through the node or segment, divided by all other possible paths as shown in Figure 1.

However, within the stated desired factors of our function, distance and destination weighing also comes into play. with lower values over larger distances, incorporating inverse distance would solve this factor, which is expressed as

$$f(j,k) = \frac{1}{d_{jk}^{p}} \qquad \textit{Equation 6}$$

where it is simply one over the power (*p*) of a distance. Using this part results in possible values of infinite (at 0 distance), up to near 0 (at infinite distance), as well as changing values between different units of distances; As represented on Figure 2, where the red graph is where the power value is 1.0, and blue is 2.0.

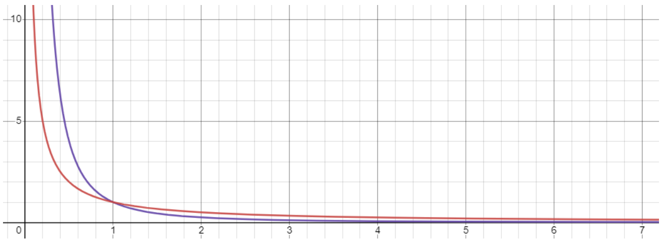

Figure 2 Graph of inverse distance function

To adjust for the context of the calculation, as further statements are involved, such as the total movement on all possible destinations or paths have to equal the patronage capacity of the origin, and alternative paths to alternative destinations are taken into account, the function for determining destination weights from the origin's perspective should be

$$f(j,k) = W_k \cdot \left(\frac{d_{jk\,min}}{d_{jk}}\right)^{1+\alpha}$$

where the distance of **a** pairing is divided by the set of alternative path distance, powered over an exponent value determined by the model. If exponent multiplier () equals to 0, then the inverse distance calculation will be linear, where any other values represent exponent mapping (except -1, where distance becomes a non-factor).

Combined, the function for discrete patronage betweenness is expressed as

$$P(i) = \sum_{j\in J} p_j \cdot \frac{\sum_{\{i\}\subset jk\in G, d_{jk}<(d_{jk\,min}\cdot r)<D} W_k \cdot \left(\frac{d_{jk\,min}}{d_{jk}}\right)^{1+\alpha}}{\sum_{jk\in G, d_{jk}<(d_{jk\,min}\cdot r)<D} W_k \cdot \left(\frac{d_{jk\,min}}{d_{jk}}\right)^{1+\alpha}}$$

Equation 7

i = path segment
j = origin point
p = patronage capacity
J = origins set
k = destination point
K = destinations set
G = graph/network
jk = paths(s) from origin to destination
d_{jk} = distance from feature to location
d_{min} = minimum distance of path
r = detour ratio
α = detour weighting exponent ratio
W = weight value

Where the resulting value is the sum of function of paths that includes the segment divided by function of all possible paths, multiplied by the origin's weight. Meaning, that segments on all part of possible path alternative will equate up to complete weight, that proves the function. Alternative paths, also known as detour ratio, is expressed by a float value of at least more than 1.0, where alternative pathways are still less than the closest path times detour ratio is still included in the summation.

3. Path computation and visualization

Said function have been applied into Python, developing a new library that houses other spatial network analysis functions that will be referred as SNAPy (Spatial Network Analysis Python). The library has dependencies on related libraries, such as Networkx, NumPy, and GeoPandas. As for visualization, for ease of use and flexibility, QGIS is used, as the links on sample data are stored in gpkg and automatic updating calculation results are possible.

The sample data used this exhibition is using Jakarta's road line, around the area of Kebayoran Baru. But, before application, the network data have been cleaned on some parts, as well as segmented to parts on junctions using a Grasshopper Script. Future development of SNAPy plans to also have this capability through using shapely. For feature points of destination and origin, its location and weighing are not representative of the condition, further elaboration and visualization are only used as examples of the application of the function.

The first example (Figure 3), a simple pathfinding of a single point of origin unto a single point of destination with 1.0 detour ratio. This results as the only recorded path is the closes one between the origin and the destination.

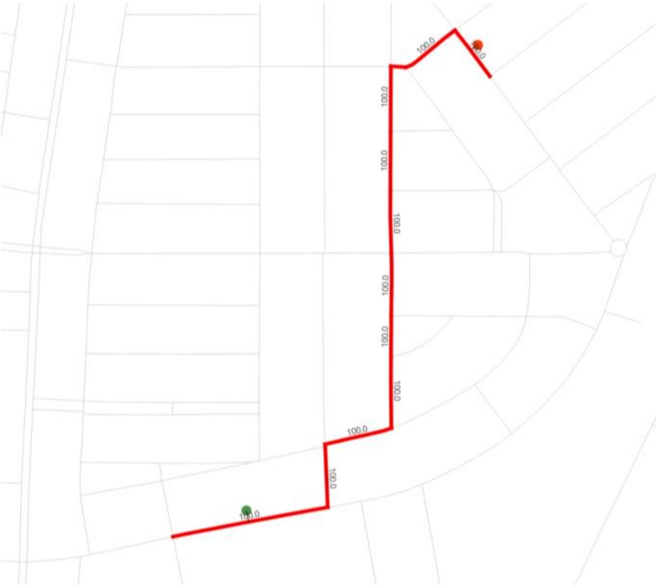

Figure 3 SNAPy pathfinding sample

By adding detour ratio of 1.5 with exponent of 0.1, the results of the function branches into multiple paths, with the furthest distance taken is limited to 1.5 times of the shortest path found (Figure 4).

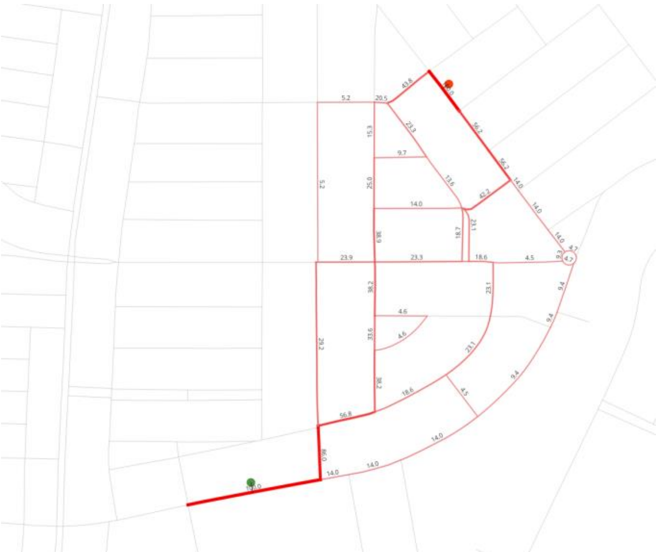

Figure 4 SNAPy Patronage Betweeness dr:1.5, exp:0.1

As for multiple destinations, and different weight comes into play, as exhibited on Figure 5, although the smaller weighted destination is closer, most (60%) of walks are directed into the further located destination. There are some segments that houses multiple paths towards multiple destinations, which is summed as stated in the mathematical function.

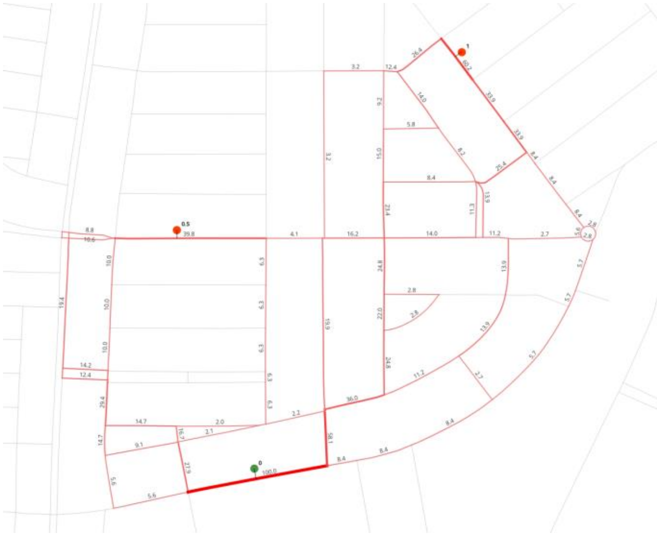

Figure 5 SNAPy multiple valued destinations

Calculations of these paths are done efficiently enough, with a small sample and single threading, it is done in 0.075 seconds. Multi-threading capabilities are also available in the library, separating chunks by origin points, and also graph segment index based pathfinding, and compilation per origin point may provide a fast and memory efficient calculation for larger simulations.

4. Distribution on time

From prior function, a movement set can be somewhat determined, i.e. morning commute to available public transport, schools, etc. But this movement happen on a spread of time. Varying movement sets such as office-parking, commercial-parks, etc. have their spread and common distribution throughout the day (Sevtsuk, 2021). Another example can be obtained in real time, where on Melbourne, some corridors provide real-time pedestrian count data, and it shows that the spread of these movement have common peaks, and can be different on different locations depending on its location, as on Figure 6.

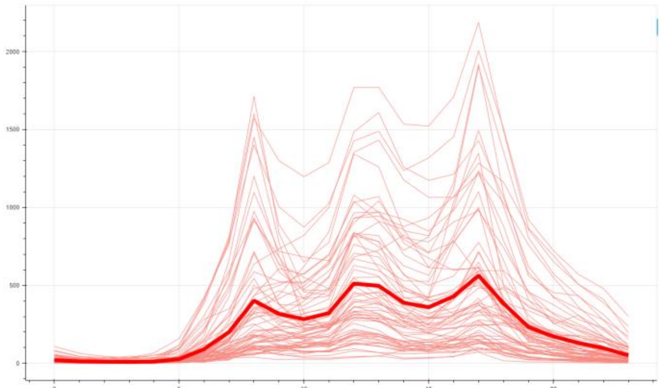

Figure 6 Melbourne pedestrian 30-day average 2022-12

The spread of this sets of movement, as most statistical distribution are on the natural world, is made from numerous overlapping and maybe even interacting asymmetrical distribution. As the total movement set amount is known, skewed normal distribution suits this context best (Azzalini, Capitanio, 2014), which is expressed as

$$f(x) = \frac{2}{\omega_m}\cdot\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\xi_m}{\omega_m}\right)^2}\cdot\frac{1}{2}\left(1+erf\left(\frac{\alpha\cdot\left(\frac{x-\xi_m}{\omega_m}\right)}{\sqrt{2}}\right)\right)$$

Equation 8

with the location (ξ) parameter defined as

$$\xi_m = \mu_m - \omega_m\cdot\left(\frac{2\alpha}{\sqrt{\pi(1+\alpha)^2}}\right)$$ *Equation 9*

Since there are multiple movement sets, some summation function will also be applied. Furthermore, as the desired readability of results, the output value can be expressed in the amount of movement per a range of time, such as hour or several minutes, which the resulting function as

$$F(i,h) = \sum_{m\in M}\int_{h-r}^{h+r}\frac{P_m(i)}{\omega_m}\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\left(\frac{t-\xi_m}{\omega_m}\right)^2}\cdot\frac{1}{2}\left(1+erf\left(\frac{\alpha\cdot\left(\frac{t-\xi_m}{\omega_m}\right)}{\sqrt{2}}\right)\right)dt$$

Equation 10

i = path segment
h = time mark
r = time spread of calculation
m = a movement relation
M = set of movement relation
P = discrete patronage betweenness
ω = distribution shape
ξ = distribution location
α = distribution skewness
μ = distribution mean

with the location (ξ) parameter defined as stated on Equation 9. An example with three movement sets with varying distribution parameters are shown on Figure 7, with the green line as the sum of all movement set, and the x axis represents the time.

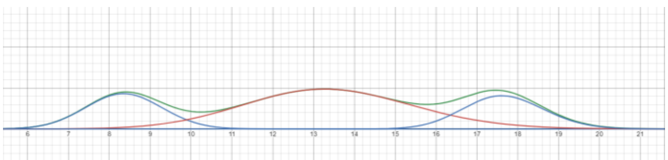

Figure 7 Sample accumulated skewed distribution

A reminder, note that the desired graph and output value should already be integrated into timeframes to be readable. The graph shown on Figure 7 are the accumulation of distribution function without integration.

5. Distribution calculation and visualization

The means for calculating the distribution is also integrated within the SNAPy library, using scipy and numpy to provide a faster calculation, as this extrapolation of movement set can result in a substantial amount of data, where each segment from the network can be extrapolated into the amount of keyframes for each set of time calculated as seen in figure 8, where each blue line represents a segment, and the red line is the region's average.

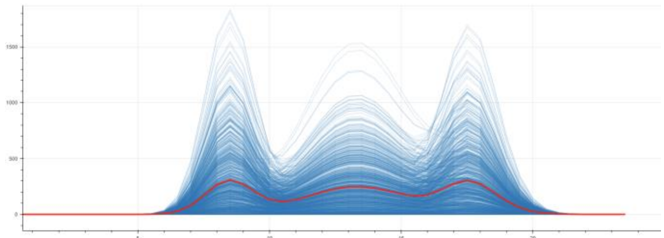

Figure 8 Plotted of a regions movement spread using SNAPy time distribution

6. Notes and usage

This method is formed as a conjecture, and still bears the burden of proof on future study and application; As this is made by putting a common thought process into a quantitative framework.

These new developed methods are a part of Spatial Network Analysis Python (SNAPy), which during the writing of this method is still in development. The current state have been personally assessed to have an easy enough application capabilities, though some data preparation with other programs such as Rhino and Grasshopper is still required.

References

Sevtsuk A (2021) We shape our buildings, but do they then shape us? A longitudinal analysis of pedestrian flows and development activity in Melbourne. PLoS ONE 16(9):e0257534

Barrat, A; et al (2004) The architecture of complex weighted networks

Garrison, W. L., Marble, D. F., (1962) The Structure of Transportation Networks, U.S.A.T. Command, U.S. Army Transportation Command Technical Report.

Garrison, W. L. (1960), Connectivity of the Interstate Highway System, Regional Science Associations, Papers and Proceedings, 6, pp. 121-137.

Haggett, P., & Chorley, J. C. (1969) Network Analysis in Geography. London: Butler & Tanner Ltd.

Ewing R, Cervero R. (2001) Travel and the Built Environment: A Synthesis. Transp Res Rec. pp.87-114.

Kim S., et al (2010) Determining the Sidewalk Pavement Width by Using Pedestrian Discomfort Levels and Movement Characteristics

Freeman, Linton (1977) A set of measures of centrality based on betweenness, Sociometry

Sevtsuk, A (2010) Path and Place: A Study of Urban Geometry and Retail Activity in Cambridge and Somerville, MA.

Naishmith, W. W. (1892) Excursions. Cruach Ardran, Stobinian, and Ben More. Scorrish Mountaineering Club Journal

Scarf, Philip (2008) A mathematical excursion in the isochronic hills. Mathematics Today

Dijkstra, E W (1959) A note on two problems in connexion with graphs

Hart, P.E; Nillson, N.J.; Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics

Pedestrian.melbourne.vic.gov.au, Real-time pedestrian count

Azzalini A.; Capitanio A. (2014) The skew-normal and related families