

# sDNA: how and why we reinvented Spatial Network Analysis for health, economics and active modes of transport

Crispin HV Cooper<sup>\*1</sup>, Alain J Chiaradia<sup>†1</sup>

<sup>1</sup>Sustainable Places Research Institute, Cardiff University

<sup>2</sup>School of Planning and Geography, Cardiff University

October 31, 2015

## Summary

We introduce sDNA, a GIS/CAD tool and methodology for analysis of spatial networks. The design decisions behind the tool are documented, in particular the choice of standardizing on the network link in order to match existing data standards and increase computational efficiency. We explore the effects of this decision on algorithm design, and present results that validate our decision to depart from a recent tradition and revive a much older one.

**KEYWORDS:** network analysis, space syntax, health, transport, pedestrian

## 1. Introduction

sDNA is both a methodology and a GIS/CAD plug-in for the analysis of spatial networks, compliant with widely accepted data standards. It has recently been used to provide environmental morphometrics for the 500,000-point UK BioBank database (UK Biobank 2013) – a major national health resource - in a project which won the 2014 RTPI award for excellence in spatial planning research (RTPI 2014). Its uses are not restricted to health, however; sDNA has also been used in mass transport investment option analysis (for Shanghai, 2014), environmental footprinting (Collins and Cooper 2014) and social cohesion studies (Cooper, Fone, and Chiaradia 2014) as well as numerous planning consultancy projects (for Trowbridge 2012, London Borough of Merton 2013, City of London 2014, Paris 2010-12). The plug-in is made available both freely and commercially.

In this short paper we discuss why we thought spatial network analysis needed reinventing (a risky undertaking given the software development time involved); how our design criteria helped to shape the new approach, and how the new approach has been validated. It thus serves of a record as to why the framework we offer is structured as it is.

## 2. Background

Spatial network analysis is an old discipline, dating back at least to Euler who in 1736 solved the problem of the Seven Bridges of Königsberg (Coupy 1851). This single piece of work necessitated the invention of both network codification and network generalisation and was the beginning of modern graph theory (Biggs, Lloyd, and Wilson 1986). Closeness - a mean shortest path measure on networks - and betweenness - a measure of flow derived from closeness, were defined in parallel in sociology (Bavelas 1948; Freeman 1977), communications networks (Shimbel 1953), transport network analysis (Garrison and Marble 1962; Ford and Fulkerson 1962; Kansky 1963), and geography (Haggett and Chorley 1969). Today these are applied in social network analysis research

---

\* CooperCH@cardiff.ac.uk

† ChiaradiaAJ@cardiff.ac.uk

(Freeman 2004), social media (Facebook) and Google's Page rank can be conceptualised as a re-invention of accessibility weighted by opportunities importance (Hansen 1959). The concept of mean shortest path was also used by Christaller (1933) and in Reilly's (1931) gravity model following Ravenstein (1885; 1889). Mean shortest path is in turn part of the early definition of accessibility, either weighed by the importance of destinations (Hansen 1959), or unweighted (Ingram 1971), which also underpin most of the last 50 years of transportation modelling.

More recently, spatial network analysis at a highly spatially desegregated (street link) level has been applied to a variety of problems in economic and transport modelling. Such analyses can be predicated on the existence of origin/destination data points (MIT 2011; ESRI 1999), or alternatively can dispense with collection of such data, instead studying only the structure of the network itself. Haggett and Chorley (1969), Kansky (1963), early space syntax research (Hillier and Hanson 1984) and place syntax (Stahle, Marcus, and Karlstrom 2008) take this latter approach. Space Syntax used an axial line network codification; closeness and betweenness on axial lines gave good correlations with measured pedestrian and vehicle flows (Hillier et al. 1993; Penn et al. 1998). This was subsequently adapted to segmented axial line (Hillier and Iida 2005) and segmented link (Turner 2007).

### **3. Rationale**

Our aim in developing sDNA was to adapt such techniques to current cartography standards and datasets. On the network mapping side, most cartography uses the link-node standard: e.g. OS ITN (Ordnance Survey 2011a), the European road representation standard (ISO 2011), US Tiger lines, OpenStreetMap and indeed the data underlying any modern satnav system. It thus made sense to make link representation an intrinsic part of the algorithm in sDNA. There are additional reasons for doing so, however. Density of links - rather than of segments or network length - can give a good proxy for origins and destinations in the absence of such data (Chiaradia et al. 2012). For a given level of detail, the links in a network are uniquely defined (unlike axial lines - see Ratti 2004) and hence make sense as a mapping standard. Finally using whole links, rather than link segments, reduces the number of entities in an analysis thus increasing computational efficiency.

Basing sDNA on established data models has the further advantage of allowing us to leverage additional data such as OS Mastermap topography layer (Ordnance Survey 2011b), addressing layer (Address Layer2) as well as survey or census data. Currently this is done by joining related data to each network link, leaving network links as the fundamental unit of analysis - although in cases where greater precision is required, links may be split to accommodate data points. We take an approach that offers a spectrum of options between all- and no-data alternatives, both for origin and destination weights and for what is variously called cost or impedance: the quantity which is minimized in selection of shortest paths. (We prefer the simpler term 'distance', which may be defined differently depending on the analysis).

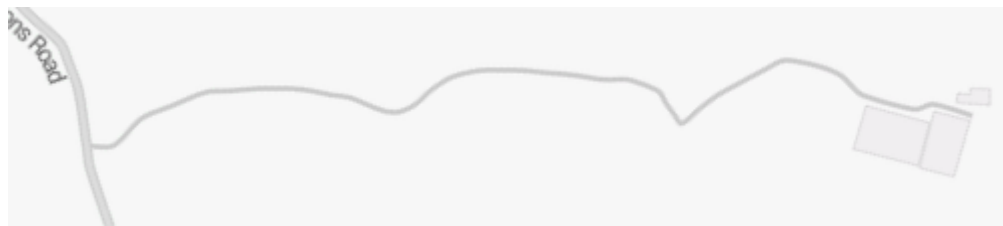
### **4. Methodology**

Standardizing on the network link required a new algorithmic approach, in particular with respect to determining the locality of analysis. As axial lines are a cognitive representation of space based on lines of sight, it is reasonable to treat them as discrete entities due to the cognitive cost of traversing between them. This is also true of angular segments, based on the cognitive and time costs incurred by turning corners (Caldwell 1961; Kirby 1966; Turner 2001). The same cannot be said of network links: while a user of a high street could probably tell you how many corners - if any - the street has, or how long it is, the number of links that constitute its length is likely of no direct relevance to them (illustrated in Figure 1 - although indirectly, the link density probably correlates to their reason for being there). Thus, while the approach taken by e.g. Hillier and Iida (2005) is to measure locality as a step radius, in terms of the number of axial lines or segments that are traversed, we choose instead to follow those who measure locality of analysis in Euclidean network terms (Kansky 1963; Haggett and Chorley 1969; Sheffi 1985; Porta, Crucitti, and Latora 2006; Turner 2007), for example considering

all links within 1 km of each origin as measured along the network. We refer to such a locality as a network radius. But the similarity of approaches ends here, because links – unlike axial lines or angular segments – are not only perceptually non-discrete, but also cannot be treated as indivisible from this standpoint (Figure 2).



**Figure 1** Queen Street, Cardiff is composed of 8 network links although perceptually is close to being a single straight line. Base map © OpenStreetMap contributors



**Figure 2** A long rural lane with multiple corners consists of only one link, but perceptually consists of multiple segments. © OpenStreetMap contributors

As traversing an exact distance from each origin is likely to result in the inclusion of partial portions of links, we define two modes for handling this situation. In *discrete space*, we choose to include a link in the analysis if its centre falls within the network radius, and exclude it otherwise. In *continuous space*, our algorithm dynamically cuts each link at the point where the edge of the radius is reached, and in subsequent analysis the partial link receives a partial weighting (as shown in Equation 1) in any subsequent calculation of closeness, betweenness, network quantity or averaging of other network qualities such as diversion ratio (Chiaradia, Cooper, and Webster 2012):

$$P(L) = \frac{\text{length of } L \text{ falling within radius}}{\text{length of } L} \quad (1)$$

On large scales, the difference between the two is minimal as the portions of network ‘wrongly’ included or excluded in discrete space are small in comparison to the network which is wholly within the radius. On short scales however, particularly those relevant to pedestrian transport, continuous space analysis can produce a significantly more accurate result.

## 5. Results and Outcomes

Convenient though it may be to adapt spatial network analysis to match modern data representations, there was a risk was that the resulting models would not be representative of human behaviour. Fortunately, empirical testing proved their worth in validation against measured pedestrian and vehicle flows. Table 1 presents some correlations between sDNA measures and measured flows for the classic space syntax dataset in central London. This micro to macro behavioural performance is

what underpins sDNA's applicability and reliability in the studies we cited in the introduction. The theoretical implications will be discussed in a forthcoming paper.

**Table 1** Correlations between sDNA network variables and measured flows

Network variable		Num. measured points (n)	Mean Angular Distance (Closeness)	Angular Betweenness
r <sup>2</sup> with vehicle flow	Barnsbury	82	0.73****	0.74****
	Clerkenwell	42	0.82****	0.80****
	South Ken.	46	0.82****	0.76****
	Brompton	61	0.60****	0.58****
	(mean)		0.74	0.72
r <sup>2</sup> with ped. flow	Barnsbury	102	0.69****	0.60****
	Clerkenwell	51	0.78****	0.73****
	South Ken.	62	0.49****	0.58****
	Brompton	85	0.56****	0.47****
	(mean)		0.63	0.60

\*\*\*\* indicates significance P<0.0001%

Recent adaptations to the software enable (1) analysis in full 3-d, (2) choice of theoretic hybrid distance metrics based on pedestrian route choice analysis (3) output of shortest paths and network buffers, (4) large scale network analysis (e.g. whole UK, over 10<sup>6</sup> links). As an example of a 3d hybrid metric, height changes can be valued differently according to whether they take place on stairs, escalator or elevator; and these factors in turn valued in proportion to Euclidean distance and angular change. Such features have been already been shown to be useful in modelling pedestrians in complex multi-level urban environments, and cyclist travel mode and route choice in relation to slope, with further applications in transport and health research as well as land use planning.

## 6. Biographies

Crispin Cooper is lead developer for the sDNA software, and also researches applications of network analysis in health, community cohesion and active modes of transport. Crispin has a Master's degree in Computer Science (University of Cambridge, 2002) and a PhD in City & Regional Planning (Cardiff University, 2010).

Alain Chiaradia is the concept lead for sDNA software design. He has 20 years' experience in software design and use of spatial design network analysis as honorary senior research fellow at UCL and as consultant to local authorities, developer, and urban designer in the UK and internationally. His research interest spans from economic geography to urban design economics.

## 7. Author contributions

AC developed the initial concept and conducted most of the literature review. CC worked on the rationale, software development and analysis. Both authors reviewed and edited the final version.

## References

- Bavelas, Alex. 1948. "A Mathematical Model for Small Group Structures." *Human Organisation* 7: 16–30.
- Biggs, N L, E Keith Lloyd, and Robin J Wilson. 1986. *Graph Theory 1736-1936*. Oxford: Oxford University Press.

- Caldwell, T. 1961. "On Finding Minimum Routes in a Network With Turn Penalties." *Communications of the ACM* 4: 107–8.
- Chiaradia, Alain, Crispin Cooper, and Chris Webster. 2012. *sDNA a Software for Spatial Design Network Analysis, Specifications*. UK: Cardiff University. <http://www.cf.ac.uk/sdna/wp-content/downloads/documentation/Detailed%20sDNA%20measure%20descriptions.pdf>.
- Chiaradia, Alain, Bill Hillier, Christian Schwander, and Martin Wedderburn. 2012. "Compositional and Urban Form Effects in Centres in Greater London." *Urban Design and Planning - Proceedings of the ICE* 165: 21–42.
- Christaller, W. 1933. *Die Zentralen Orte in Suddeutschland - The Central Places of Southern Germany*. Jena - Englewood Cliffs, NJ: Gustav Fischer Verlag - Prentice-Hall.
- Collins, Andrea, and Crispin H V Cooper. 2014. *The Environmental Impacts of Festivals: Reflections on the 2012 Hay Literature Festival, Wales*.
- Cooper, Crispin H. V., David L. Fone, and Alain Chiaradia. 2014. "Measuring the Impact of Spatial Network Layout on Community Social Cohesion: A Cross-Sectional Study." *International Journal of Health Geographics* 13 (1): 11. doi:10.1186/1476-072X-13-11.
- Coupy, E. 1851. "Solution D'un Problème Appartenant À La Géométrie de Situation, Par Euler." *Nouvelles Annales de Mathématiques Ire Série* 10: 106–19.
- ESRI. 1999. *ArcGIS Network Analyst*. <http://www.esri.com/software/arcgis/extensions/networkanalyst>.
- Ford, L R, and D R Fulkerson. 1962. *Flows in Networks*. 2nd ed. Princeton: Princeton University Press.
- Freeman, L C. 1977. "A Set of Measures of Centrality Based on Betweenness." *Sociometry* 40 (1): 35–41.
- . 2004. *The Development of Social Network Analysis*. North Charleston: BookSurge LLC.
- Garrison, W L, and D F Marble. 1962. *The Structure of Transportation Networks*.
- Haggett, Peter, and Richard J Chorley. 1969. *Network Analysis in Geography*. London, UK: Hodder & Stoughton Educational.
- Hansen, W G. 1959. "How Accessibility Shape Land Use." *Journal of American Institute of Planners* 25: 73–76.
- Hillier, Bill, and Julienne Hanson. 1984. *The Spatial Logic of Space*. Cambridge: Cambridge University Press.
- Hillier, Bill, and S Iida. 2005. "Network and Psychological Effects: A Theory of Urban Movement." In *Proceedings of the 5th International Space Syntax Symposium*. TU Delft: Techne Press.
- Hillier, Bill, Alan Penn, T Grajewski, and J Xu. 1993. "Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement." *Environment and Planning B: Planning and Design* 20: 29–66.
- Ingram, D R. 1971. "The Concept of Accessibility: A Search for an Operational Form." *Regional Studies* 5: 101–7.

- ISO. 2011. *Intelligent Transport Systems - Geographic Data Files (GDF) - GDF5.0*. 14825. [http://www.iso.org/iso/home/store/catalogue\\_ics/catalogue\\_detail\\_ics.htm?csnumber=54610](http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=54610).
- Kansky, K J. 1963. *Structure of Transport Networks: Relationships between Network Geometry and Regional Characteristics*. Chicago: University of Chicago.
- Kirby, R F. 1966. "A Minimum Path Algorithm for a Road Network with Turn Penalties." In Sydney: 3rd Australian Road Research Board (ARRB) Conference.
- MIT. 2011. *Urban Network Analysis*. <http://cityform.mit.edu/projects/urban-network-analysis-toolbox>.
- Ordnance Survey. 2011a. *OS Mastermap ITN Layer Urban Paths Theme*. <http://www.ordnancesurvey.co.uk/docs/user-guides/ITN-paths-tech-userguide.pdf>.
- . 2011b. *OS Mastermap Topography Layer User Guide and Specification*. <http://www.ordnancesurvey.co.uk/docs/user-guides/os-mastermap-topography-layer-user-guide.pdf>.
- Penn, Alan, Bill Hillier, David Banister, and J Xu. 1998. "Configurational Modelling of Urban Movement Networks." *Environment and Planning B: Planning and Design* 25: 59 – 84.
- Porta, S, P Crucitti, and V Latora. 2006. "The Network Analysis of Urban Streets: A Primal Approach." *Environment and Planning B: Planning and Design* 33: 705–25.
- Ratti, Carlo. 2004. "Space Syntax: Some Inconsistencies." *Environment and Planning B: Planning and Design* 31 (4): 487–99. doi:10.1068/b3019.
- Ravenstein, George, Ernest. 1885. "The Laws of Migration." *Journal of the Statistical Society of London* 48: 167–235.
- . 1889. "The Laws of Migration." *Journal of the Royal Statistical Society* 52: 241–305.
- Reilly, W J. 1931. *The Law of Retail Gravitation*. New York: Knickerbocker Press.
- RTPI. 2014. *Excellence in Spatial Planning Research Awards*. <http://www.rtpi.org.uk/briefing-room/news-releases/2014/september/excellence-in-spatial-planning-research-awards-winners/>.
- Sheffi, Yosef. 1985. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs: Prentice Hall.
- Shimbel, A. 1953. "Structural Parameters of Communication Networks." *The Bulletin of Mathematical Biophysics* 15 (4): 501–7.
- Stahle, A, L Marcus, and A Karlstrom. 2008. "Place Syntax: Geographic Accessibility with Axial Lines in GIS." In Delft: Proceedings in 5th Space Syntax Symposium.
- Turner, Alasdair. 2001. "Angular Analysis." In *3rd International Space Syntax Symposium*. Atlanta.
- . 2007. "From Axial to Road-Centre Lines: A New Representation for Space Syntax and a New Model of Route Choice for Transport Network Analysis." *Environment and Planning B: Planning and Design* 34 (3): 539–55.
- UK Biobank. 2013. *BioBank*. <http://www.ukbiobank.ac.uk/>.