Markov Chain Topological Route Selection

E.J. Manley[1], T. Cheng[1], J. Haworth[1]

[1] SpaceTimeLab, University College London,
Gower Street, London, United Kingdom.
Telephone: +44 (0) 20 7679 7224
Email: Edward.Manley.09@ucl.ac.uk

1. Introduction

Urban transport systems are intrinsically defined by the behaviours of many thousands or millions of individuals. It is thus vital, that where one seeks to model transport systems under change or shock, individual behaviour is well understood. Yet recent research into one core spatial behaviour – route selection – suggests that conventional models of such behaviour are poorly conceived (Manley et al. 2012b).

Existing models of route-choice suggest that individuals select routes based predominantly on a principle of economic rationality, namely that they will always aim to minimise travel time or distance. However, research is growing that suggests that such assumptions are unrealistic (Golledge 1995, Wiener et al. 2004, Zhu & Levinson 2010), and result in inaccurate predictions of road transport flows (Manley et al. 2012a). Rather, behaviours are influenced by subjective representations of space and estimations of prospective travel times. Furthermore, it has been established that individuals do not necessarily remember specific routes or paths, but rather maintain a topological network within the brain of salient places through which a route is constructed (Golledge 1978, Passini 1984, Winter et al. 2007). Such findings help better explain the observable heterogeneity in behaviours observed in route selection, and the presence of non-linearity in choice, in that certain areas of the network appear to attract and repel traffic significantly more than might be expected (Manley et al. 2012b).

In aiming to more accurately model road transport flows, one must therefore seek to incorporate the full extent of this behavioural heterogeneity within a model of route-choice. The Markov chain Monte Carlo (MCMC) method is a statistical approach to modelling and predicting behavioural heterogeneity. Through MCMC, rather than specifying a set of parameters and making predictions on the relative contribution of each in determining behaviour, prior actions – described through data – are used to make predictions on future behaviours. In this research, an MCMC model is applied to describing route-choice behaviour across a topological representation of the road network. In using the MCMC approach, we are able to capture all relevant behaviours that may be difficult to accurately specify using alternative methods. While this approach does not offer any explanatory power in describing individual route selection, it does enable the improved statistical description of the full heterogeneity of choice among a population of individual travellers.

2. MCMC Route Choice Structure and Definition

The traditional Markov chain model can be thought of as a network of nodes, connected according to their probability of sequential selection. The probability of moving from one node to another is defined according prior observed behaviours, usually captured through data. In the case of the route-choice model described here, the network nodes are intended to represent salient locations on the road network, and the probability of moving between two given points specified according to a large dataset of cab routes in London. For the purposes of this model, an adapted version of the traditional Markov chain is developed, incorporating higher-level strategic choice behaviours that shape movement towards a target destination.

2.1. Node Specification

The MCMC model is applied to a topological representation of the road network. The nodes within the MCMC model are intended to represent locations on the road network at which route selection choices are likely to be made. As such, a structure of nodes is developed from the points of junctions between major thoroughfares. It may be assumed that, as these locations represent points of likely route deviation, individuals are more likely to maintain a memory of them. Route choice, therefore, rather than being executed on a link-by-link basis, is modelled through a node-to-node selection process.

2.2. Inter-Nodal Probability Specification

The probability of connection between given nodes on the road network are drawn from the route choice behaviours during around 700000 private hire cab journeys through London, United Kingdom. In our approach, unlike traditional Markov chain representations, node selection is dependent on the connection with the current node *given* the node selected prior to the current node. This allows for a sense of linear path continuity to be incorporated into the selection process, so that if an individual is approaching a node from one direction, they are most likely to continue broadly in same direction. This process is best demonstrated in the model in Figure 1, where an individual's choice from node j is shaped by the fact that they'd previously visited node k.

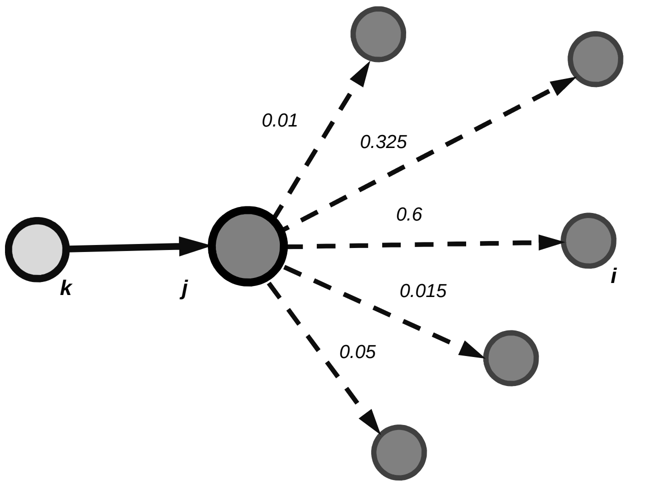

Figure 1. Markov chain network structure, where k is the previous node, j is the current node and i is an optional next node, showing the probability of connections between node i and a range of sequential nodes.

Using the routes identified within the dataset, the inter-nodal probabilities of connection are established. As such, the probability of selecting node *i* from node *j* represents the proportion of trips from node *j* that continue on to node *i* where the previous node was node *k*.

2.3. Target Focussing

Node specification and the identification of probability of connections between nodes does not present complete picture of routing behaviour, however. While the choice process is already abstracted to a coarser spatial scale, in line with prior research findings, it is apparent that higher-level strategic behaviours are also at play. This approach therefore implements an additional facet of behaviour, describing the probability of deviating away from the direction towards an individual's target destination. A probability is generated again from the data, describing the probability of selecting a node of a given deviation from the straight line towards a target. Figure 2 describes the relationship between angular deviation and probability of node selection.

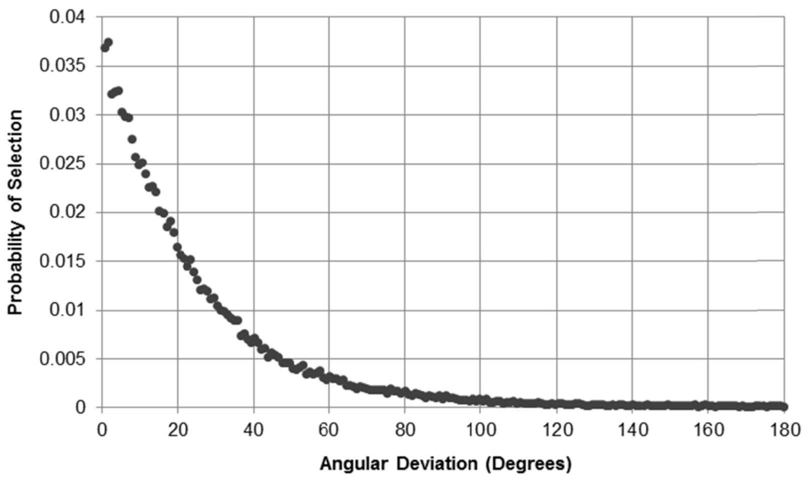

Figure 2. Graph demonstrating the probability of selecting an option node at given angular deviation from target.

This probability is then used to weight the prior probability of selection of a node, as follows, where P_i^N represents the prior probability of nodal connection, P_i^T is the probability of selection given node deviation from target angle, and P_j represents probabilities for all alternatives from the current node:

$$P_i = \frac{P_i^N P_i^T}{\sum_j P_j^N P_j^T}$$

In certain circumstances then, while one inter-nodal connection may usually be probable from a given node, the influence of the target direction will weight against that selection if it not running in the direction towards the target. The combination of the two probabilities therefore represents both the likelihood of nodal coupling along with the preference for a straight-line path.

3. Model Output

The model was applied to the prediction of a number of origin-destination pairs in London. The route selection process proceeds in a node-to-node fashion from origin to destination. Nodes are selected through uniform weighted random selection, according to the weighted probabilities of each option. For each origin-destination pair the selection process is executed 1000 times, ensuring full exploration of heterogeneous behaviours.

At this initial stage, in order to demonstrate with clarity the selection process, the subsequent road-level selections, that would ordinarily follow the node selections, is not incorporated. The results therefore show the percentages of trips selecting a node-by-node path between an origin and destination. As such, without comparable datasets, the results provided here are merely intended to demonstrate the ability of this approach to identify the range of feasible route selections from an origin to a destination.

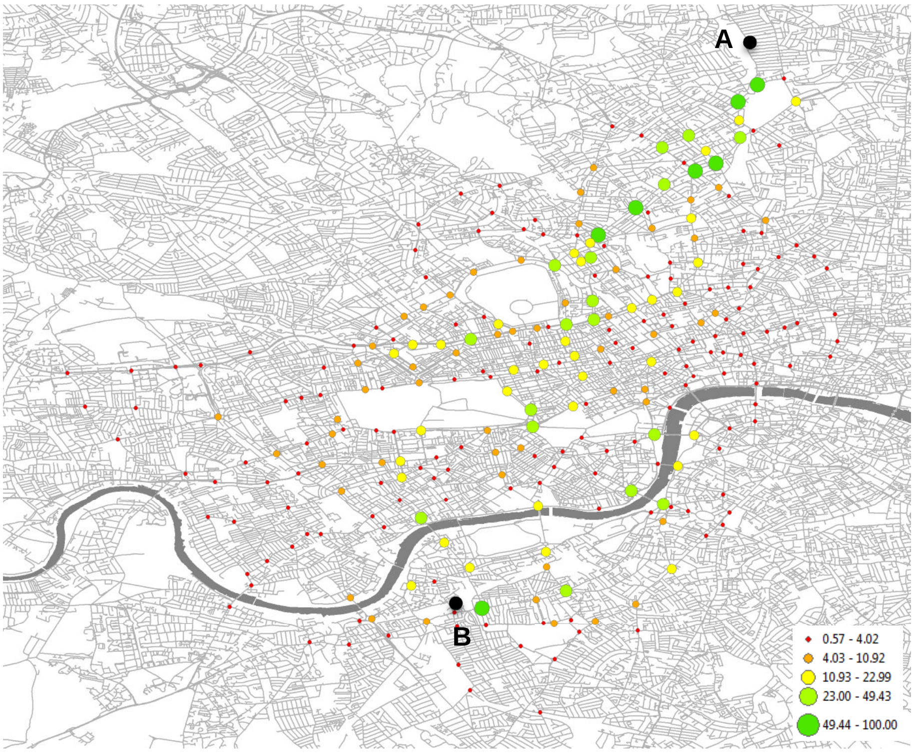

Figure 3. Node selection percentages from location A to location B, in London, United Kingdom.

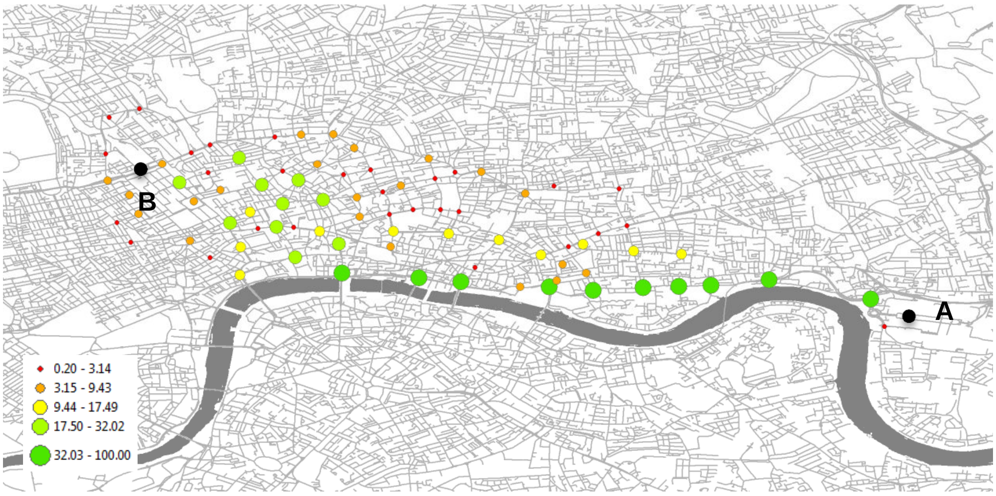

Figure 4. Node selection percentages from location A to location B, in London, United Kingdom.

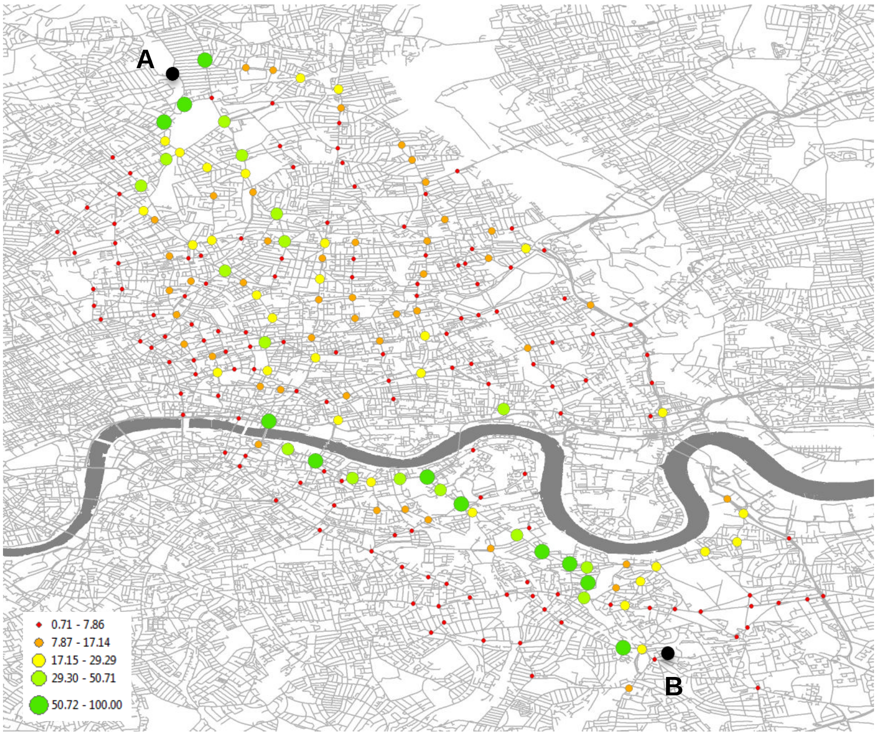

Figure 5. Node selection percentages from location A to location B, in London, United Kingdom.

4. Discussion and Conclusions

The process by which an individual selects a route from an origin to a destination is highly complex, influenced by experience, subjectivity and spatial non-linearity. What is clear from previous work is that this decision process takes place within a topological framework, whereby salient locations are linked within a network structure. The route choice model presented in this paper represents a movement towards a realistic representation of this cognitive mechanism.

Through the specification of the Markov chain model, using a rarely seen wealth in routing data, we have been able to effectively describe the heterogeneity inherent across a population of individuals making a route selection between an origin and destination. Although this approach does not yet offer any explanatory power as to the nature of the decisions being taken by the individual, it does provide a picture of population route choice that – given it is drawn from previous route selections – must broadly represent the choices of a population of individuals.

There are some clear avenues for further work with respect to this research. Firstly, there must be some work carried out into the validation of this approach, with identification of whether the selections are broadly in line with the real data for specific origin-destination pairs. Second, further work should be carried out into identifying how effective the current node specification is in capturing key route decision points, and whether certain nodes may be added or removed from this representation. And finally, as hinted at above, the model currently lacks explanatory power, thus further analysis should be carried out into identifying the important influencing parameters causing node-to-node linkage.

5. Acknowledgements

This work is part of the STANDARD project – Spatio-Temporal Analysis of Network Data and Road Developments, supported by the UK Engineering and Physical Sciences Research Council (EP/G023212/1) and Transport for London (TfL). The data for this work was kindly provided by Addison Lee Private Hire Taxi Company.

6. References

Golledge, R.G. 1978. Learning about urban environments. In T. Carlstein, D.N. Parkes, & N.J. Thrift (eds). *Assessing the Economic Impact of Retail Centres: Issues, Methods and Implications for Government Policy.*

Golledge, R.G., 1995. Path Selection and Route Preference in Human Navigation: A Progress Report (UCTC No. 277). Berkeley, California: The University of California Transportation Center.

Manley, E.J., Cheng, T., Penn. 2012a. *Modelling Movement in the City: The Influence of Individuals*. AGILE Workshop on Complexity Modelling for Urban Structure and Dynamics. April 2012.

Manley, E.J., Addison, J.D., Cheng, T., Penn, A. 2012b. *Understanding Urban Traffic Patterns Using (Big) Data and Agent-Based Simulation.* COSMIC Satellite Meeting at ECCS, September 2012.

Passini, R. 1984. Spatial Representations, a Wayfinding Perspective. *Environmental Psychology*. 4:152-164.

Wiener, J.M., Schnee, A., & Mallot, H.A. 2004. Use and interaction of navigation strategies in regionalized environments. *Environmental Psychology*. 24. 475-493.

Winter, S., Tomko, M., Elias, B., Sester, M. 2007. Landmark Hierarchies in Context. *Environment and Planning B*. 35:381-398.

Zhu, S., & Levinson, D. 2010. Do people use the shortest path? An empirical test of Wardrop’s first principle. Working paper. Available at: http://nexus.umn.edu/Papers/ShortestPath.pdf.