ART-P-MAP Neural Network Modelling of Spatial Choices

Z. Gong*[1] and J-C. Thill[2]

[1]School of Geography, Earth and Environmental Sciences, University of Birmingham
[2]Department of Geography & Earth Sciences, University of North Carolina at Charlotte
*Email: z.gong@bham.ac.uk

Abstract

The modelling of spatial choices is solidly grounded in the behavioural theory of discrete choices, which itself conceptualizes spatial choices as the result of a process consistent with random utility theory. Utility-based discrete choice models provide the primary framework of analysis of spatial choices. More recently computational models and machine learning techniques have also been shown to be quite effective at comprehending the factors of spatial choices. They have been justified on conceptual grounds based on alternative principles of human decision making and information processing. Within this context, this paper proposes a semi-supervised neural network learning system, ART-P-MAP, for spatial choice modelling. A behavioural foundation for this model is provided aiming to alleviate some of the limitations of conventional modelling approaches. Due to its inherent model structure and learning mechanism, the new approach allows various factors to be taken into consideration and more complex structures to be captured from the empirical data. An illustrative case in Minneapolis-St. Paul, MN, is provided for evaluating the proposed model's performance with a comparison to a decision tree model.

Keywords: ART-P-MAP, Neural networks, Spatial choice modelling.

1. Introduction

Spatial choice theory has been dominated by the paradigm of a two-stage process (Timmermans and Golledge, 1992). The universe of alternatives is reduced in the first stage to a smaller set called the choice set whose construction depends upon one’s knowledge and awareness of choice alternatives, the feasibility and accessibility of alternatives as well as their perceived characteristics at decision time. Existing modelling frameworks were usually formalized in the context of random utility choice modelling (Fotheringham and O’Kelly, 1989). Choice probabilities are decomposed into probabilities for obtaining given choice sets and choice probabilities conditional upon choice sets. Numerous approaches have been proposed for spatial choice set modelling such as nested logit and competing destinations. However, they are still subjected to limitations including predefined clusters of alternatives or pre-specified functional forms for choice set estimation. In addition, the assumption of identical individual made in random utility theory and the related concept of choice set restricts themselves to taking no account of the variation of individuals’ decision patterns across individuals because of a large number of decision makers and the variety of circumstances that regulate the delineation of their choice sets. Recognizing the capability of computational intelligence techniques, which provides the mechanisms by which knowledge is acquired inductively by extracting useful information from raw datasets, this study is to incorporate these techniques into the existing choice set modelling framework to overcome some limitations imposed on the conventional methods.

In previous research by Thill and Wheeler (2000), an inductive decision tree was used to fit a destination choice dataset (including decision makers' socio-economic characteristics, separation between destinations and decision makers, and characteristics of destinations) to derive the decision rules that reflect factors influencing destination choice. Although this model is capable of estimating the choice set by extracting information from the empirical data, it assumes a uniform hierarchical structure of preferences for all decision makers' information processing strategies to delineate choice sets and make decisions. This is not appropriate for realistic situations where the hierarchical structures may well vary across decision makers.

2. Spatial decision pattern and ART-P-MAP

This study aims at forming a concept of spatial decision pattern that is an alterative to the concept of choice set. Decision patterns can be represented as hypothetical clusters pertaining to choices made by individuals. These hypothetical clusters mean to capture various and complex structures (not only hierarchical) of information processing strategies of decision makers from empirical data. These clusters can be extracted by learning models in a multi-attribute space constructed from utility related factors, including characteristics of individuals, characteristics of alternatives and their separation; they can also be utilized to predict the probability for a specific choice. The challenge is that these hypothetical clusters need to be detected, confirmed, and accommodated automatically by a learning model, which means that there is no assumption or pre-specification regarding the property of these clusters (e.g., number and structure).

To address the above issue, this study proposes to adopt ART-P-MAP neural networks, a general incremental learning model developed by Gong et al. (2015) for spatial modelling involving discrete choices. ART-P-MAP has the capability to automatically accommodate heterogeneous decision patterns based on adaptive resonance theory (Gong et al., 2015) and to make inference regarding choice probabilities based on Bayesian decision theory. In other words, it can adaptively formulate its network structure of hypothetical clusters during the training phase; more importantly, it explicitly makes probabilistic inference in response to heterogeneous data environments during prediction phase. For specifics on training and prediction, please refer to Gong et al. (2015). In prediction, the probability for a choice o among the choice universe O given an input data I is computed as:

$$P(o|I) = \frac{\sum_{h_j \in \{C_j \geq \gamma\}} F_{jo} C_j}{\sum_{o \in O} \sum_{h_j \in \{C_j \geq \gamma\}} F_{jo} C_j} \qquad \text{Equation 1}$$

where F_{jo} is the frequency of inputs with a choice o recorded for j th hypothetical cluster h_j, γ is a threshold for a match function $C_j = \frac{|I \wedge W_j|}{|W_j|}$, which measures the degree of match between an input I and a hypothetical cluster h_j using weights W_j that characterize h_j.

3. Case study

The data used for evaluating the proposed approach is from the 1990 Minneapolis-St. Paul travel behaviour inventory, which is same dataset used in Thill and Wheeler (2000). The dataset consisting of 667 trips extracted from the original database meets the following conditions such that trips are home-based, trip origins and destinations are georeferenced to corresponding traffic analysis zones (TAZ) that are within the metropolitan area, trips are not part of a multistep tour, their purpose is shopping, and travel is by car. In total, 1,165 TAZs form the universal destination set for the spatial choice problem. A training set is generated as follows. For each of the 667 realized trips, 100 unchosen destinations are randomly selected from the universal destination set, which are combined with the origin of the realized trip. Thus, the training set consists of 67,367 records of combined origins and destinations, among which only 667 are realized trips. 19 factors defining the attribute space are included in the data (Table 1).

Factor	Definition
DISTANCE	Travel distance
TIME	Travel time
POP90	1990 employment count at destination
RET_EM	1990 employment in retailing at destination
PERSERV_EM	1990 employment in personal services at destination
MALL	1 if regional mall at destination; 0 otherwise
AREA_TYPED1	1 if trip destination is in a developed area; 0 otherwise
AREA_TYPED2	1 if trip destination is in central city/CBDs; 0 otherwise
AREA_TYPED3	1 if trip destination is in an outlying business district; 0 otherwise
AREA_TYPEO1	1 if trip destination is a developed area; 0 otherwise
AREA_TYPEO2	1 if trip origin is in central city/CBDs; 0 otherwise
AREA_TYPEO3	1 if trip origin is in an outlying business district; 0 otherwise
PCOMLU	Percent of destination area occupied by commercial/service land use
AGE	Age of the decision-maker
GENDER	1 if decision-maker is male; 0 if female
INCOME	1 if annual house hold income is over $35,000; 0 otherwise
HHLDSIZE	Number of members in the decision-maker's household
INFANTS	Number of infants under 5 in household
CARS	Number of cars in household

Table 1: Factor description.

4. Results and discussion

5-fold cross-validation is applied to evaluate the accuracy of the proposed modelling approach. Results (Appendix 6.1 – 6.5) show that the performance for the proposed model is quite consistent. The total average accuracy (Table 2) is 98.4%; the average accuracies for Chosen and UnChosen categories (Table 2) are 19.4% and 99.2%, respectively. This contrast between accuracies of two categories is not surprising since the Chosen category only occupies a very small proportion of the entire dataset. Therefore, it has less chance to predict the Chosen category. This problem may result from the nature of frequency-based statistical approach embedded in the proposed model.

Category	Average Accuracy
Chosen	0.193692
UnChosen	0.991933
Total	0.984028

Table 2: Model accuracy for ART-P-MAP.

For comparison purpose, Table 3 shows results from a decision tree model used by Thill and Wheeler (2000) applied to the same dataset with 3/5 of all records for training and the rest for validation. First, it is noted that the decision tree achieves much higher sensitivity with 28.8% (proportion of the actual Chosen captured). This is because decision tree predicts much more records to be the Chosen category. In other words, it overestimates the chance of a destination to be chosen. Although it captures many true Chosen records, it loses more true UnChosen records (low accuracy 82.4%). On the other hand, when it over-predicts the Chosen category, the proportion of correct hits is very low (precision of 1.6%). In all, the 81.9% total accuracy for decision tree is much lower than for ART-P-MAP (98.4%).

Error Matrix		Actual			
		Chosen	UnChosen	Total	
Predicted	Chosen	77	4690	4767	0.016153
	UnChosen	190	22010	22200	0.991441
	Total	267	26700	26967	
		0.28839	0.824345		0.819038

Table 3: Error matrix for decision tree model.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
AUR	0.864166	0.83923	0.867021	0.834223	0.880705	0.857069

Table 4: Areas under ROC curves for ART-P-MAP.

To further examine the performance of the proposed model, the ROC curve and AUR (Area under ROC Curve) for the 5-fold cross validation and their average are provide (Table 4 and Figure 2). Generally, the performance for each fold is consistently similar based on the curve and AUR. On average, 85.7% of AUR is obtained, indicating that the model achieves high performance. In sum, according to the performance evaluation and comparison on the exemplar travel behaviour dataset, the ART-P-MAP model surpasses the decision tree model for destination choice modelling.

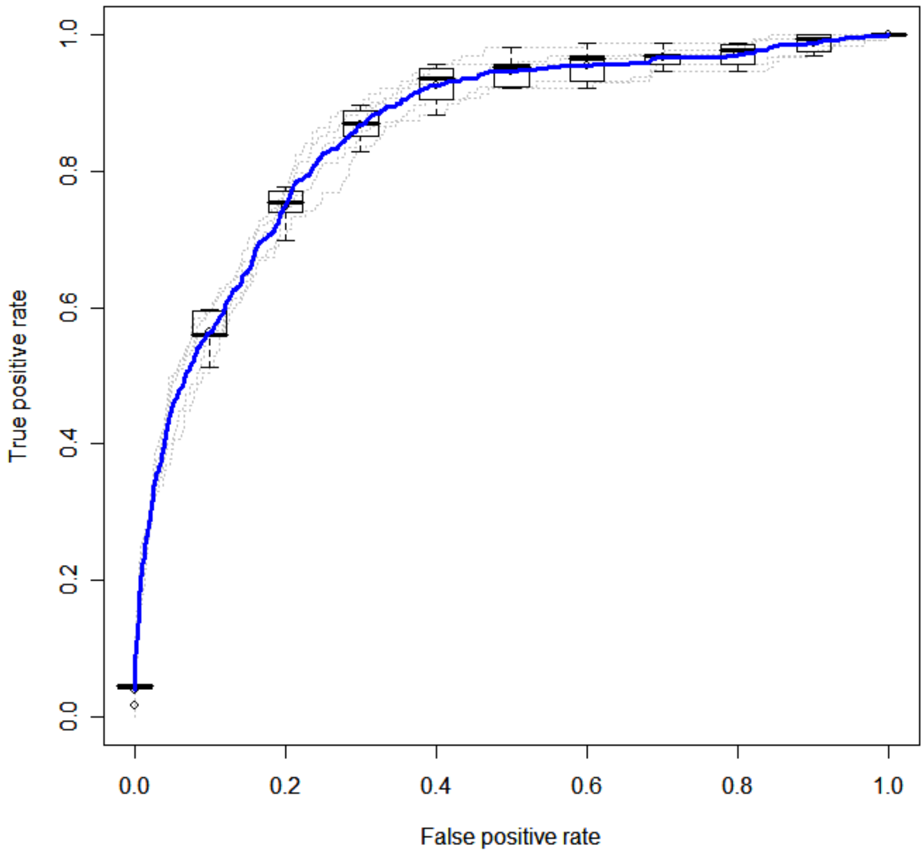

Figure 2: ROC curves for 5-fold cross-validation of ART-P-MAP.

5. References

Fotheringham, A.S. and O'Kelly, M.E. 1988. *Spatial interaction models: formulations and applications*. Netherlands: Springer.

Gong, Z., Thill, J.C. and Liu, W. 2015. ART-P-MAP neural networks modeling of land-use change: accounting for spatial heterogeneity and uncertainty. *Geographical Analysis*. **47**(4), pp.376-409.

Thill, J.-C. and Wheeler, A. 2000. Tree induction of spatial choice behavior. *Transportation Research Record: Journal of the Transportation Research Board*. **1719**, pp. 250-258.

Timmermans, H. and Golledge, R.G. 1990. Applications of behavioural research on spatial problems II: preference and choice. *Progress in Human Geography*. **14**(3), pp.311-354.

6. Appendix

6.1 Fold 1 error matrix for ART-P-MAP

Error Matrix	Actual	

		Chosen	UnChosen	Total	
Predicted	Chosen	23	102	125	0.184
	UnChosen	102	13247	13349	0.992359
	Total	125	13349	13474	
		0.184	0.992359		0.98486

6.2 Fold 2 error matrix for ART-P-MAP

Error Matrix		Actual			
		Chosen	UnChosen	Total	
Predicted	Chosen	27	89	116	0.232759
	UnChosen	89	13269	13358	0.993337
	Total	116	13358	13474	
		0.232759	0.993337		0.986789

6.3 Fold 3 error matrix for ART-P-MAP

Error Matrix		Actual			
		Chosen	UnChosen	Total	
Predicted	Chosen	22	116	138	0.15942
	UnChosen	116	13219	13335	0.991301
	Total	138	13335	13473	
		0.15942	0.991301		0.98278

6.4 Fold 4 error matrix for ART-P-MAP

Error Matrix		Actual			
		Chosen	UnChosen	Total	
Predicted	Chosen	23	104	127	0.181102
	UnChosen	104	13242	13346	0.992207
	Total	127	13346	13473	
		0.181102	0.992207		0.984562

6.5 Fold 5 error matrix for ART-P-MAP

Error Matrix		Actual			
		Chosen	UnChosen	Total	
Predicted	Chosen	34	127	161	0.21118
	UnChosen	127	13185	13312	0.99046
	Total	161	13312	13473	
		0.21118	0.99046		0.981147