Personalized Tour Recommendation Based on UserInterests and Points of Interest Visit Durations

Kwan Hui Lim, Jeffrey Chan, Christopher Leckie and Shanika Karunasekera

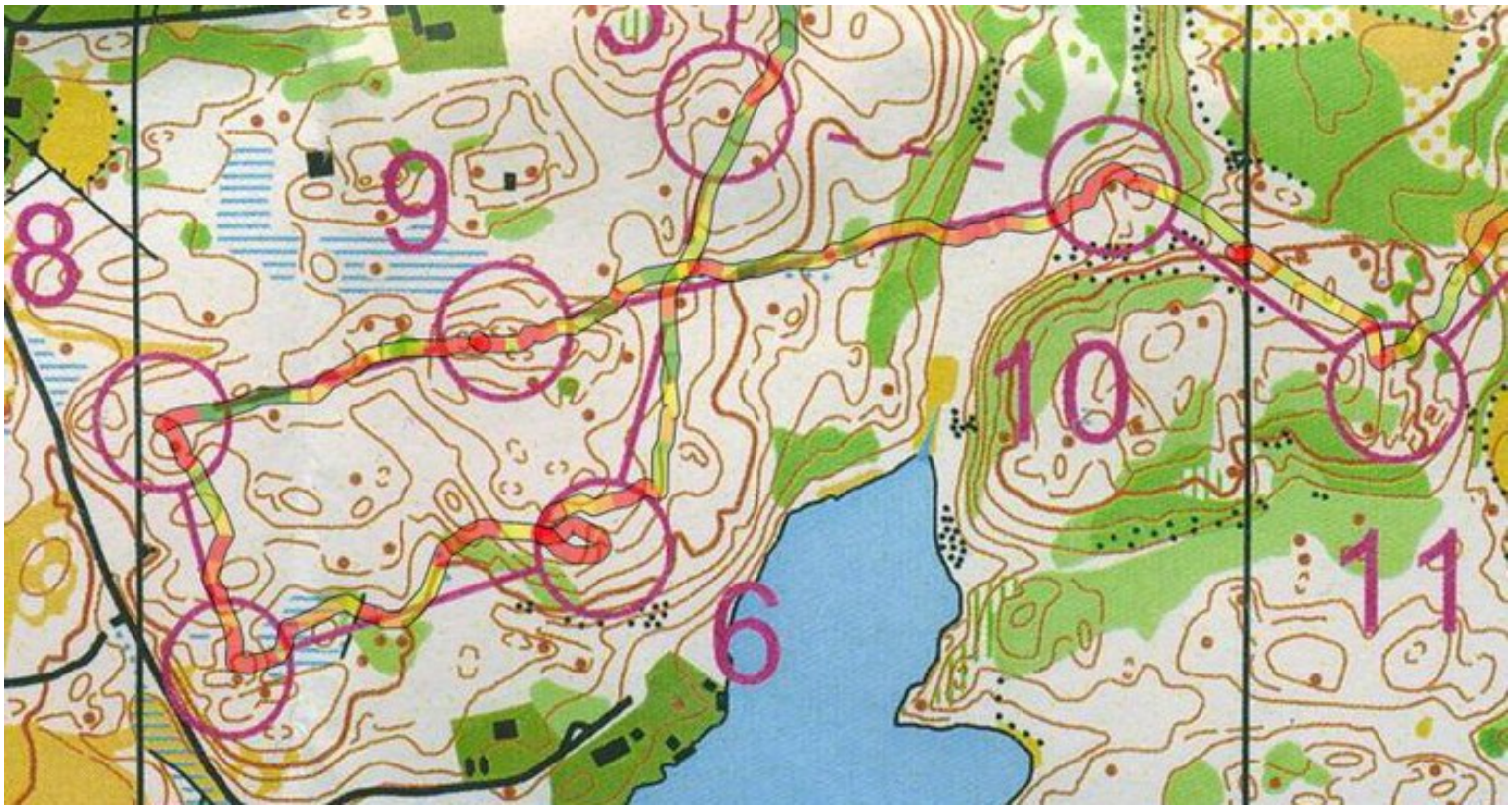
November 3, 2016

Outline

- Context
- Definitions
- Framework
- Methodology
- Results

Context

Based on the Orienteering problem [Tsiligirides, 1984].



Definitions

Given a city P with m POIs, each one having a category Cat_p.
The popularity of p, Pop_p is given by the ammount of visits to p.
For a user u, the travel history is given by

$$S_u = \Big((p_1, t^a_{p_1}, t^d_{p_1}), ..., (p_n, t^a_{p_n}, t^d_{p_n}) \Big) \tag{1}$$

Definitions
Average POI Visit duration

For all users U, the average visit duration for a POI p is given by:

$$\bar{V}(p) = \frac{1}{n_p} \sum_{u \in U} \sum_{p_x \in S_u} (t^d_{p_x} - t^a_{p_x}) \delta(p_x = p), \tag{2}$$

where n_p is the total number of visits of p.

Definitions
Time-based User Interest

For every user u, the personal interest in the category c is given by:

$$Int_u^{Time}(c) = \sum_{p_x \in S_u} \frac{(t_{p_x}^d - t_{p_x}^a)}{\bar{V}(p_x)} \delta(Cat_{p_x} = c) \tag{3}$$

Definitions
Personalized POI Visit Duration

Based on the user preferences in Cat_p.

$$T_u^{Visit}(p) = Int_u^{Time}(Cat_p) \times \bar{V}(p) \tag{4}$$

Definitions

For every travel sequence, the function to maximize is

$$\sum_{i=2}^{N-1} \sum_{j=2}^{N} x_{i,j} \left(\eta Int(Cat_i) + (1-\eta) Pop(i)\right), \tag{5}$$

where $x_{i,j} = 1$ if i and j are visited in sequence, 0 otherwise.

Definitions
Constrains

$$\sum_{j=2}^{N} x_{1,j} = \sum_{i=1}^{N-1} x_{i,N} = 1 \tag{6}$$

$$\sum_{i=1}^{N-1} x_{i,k} = \sum_{j=2}^{N} x_{k,j} \leq 1 \quad \forall k = 2, ..., N-1 \tag{7}$$

$$\sum_{i=2}^{N-1} \sum_{j=2}^{N} Cost(i,j) x_{i,j} \leq B \tag{8}$$

$$2 \leq p_i \leq N \quad \forall i = 2, ..., N \tag{9}$$

$$p_i - p_j + 1 \leq (N-1)(1 - x_{i,j}) \quad \forall i, j = 2, ..., N \tag{10}$$

Framework
Data

POI obtained from Wikipedia with coordinates and categories. Photos obtained from Yahoo! Flickr Creative Commons 100M dataset (YFCC100M), with timestamp and coordinates.

City	**# Photos**	**# Users**	**# POI Visits**	**# Seq**
Toronto	157k	1395	39k	6k
Osaka	392k	450	7k	1k
Glasgow	29k	601	11k	2k
Edimburgh	82k	1454	34k	5k

Framework

1. Determine POI visits. Distance from each photo to the POI less than 200m
2. Construct travel history. An individual sequence doesn't have visits separated by more than 8 hours.
3. Recommend Tours. Use different values for η and Int_u.

Experiments

Evaluate how the interest function and weight of user interest affects the recommendations.

- $\eta = 0$.
- Int_u^{Time} and $\eta = 0.5$.
- Int_u^{Freq} and $\eta = 0.5$.
- Int_u^{Time} and $\eta = 1$.
- Int_u^{Freq} and $\eta = 1$.

Experiments
Baseline

- Greedy Nearest (GNear): Select at random from 3 nearest unvisited POIs.
- Greedy Most Popular (GPop): Select at random from 3 most popular unvisited POIs.
- Random (Rand): Select at random from all unvisited POIs.

Results

Time-based vs Freq-based

Toronto			
Algo.	*Recall*	*Precision*	F_1-*score*
PT-.5F	.760±.009	.679±.013	.708±.012
PT-.5T	**.779±.010**	**.706±.013**	**.732±.012**
PT-1F	.737±.010	.682±.013	.698±.012
PT-1T	**.744±.011**	**.710±.013**	**.718±.012**
GNEAR	.501±.010	.512±.015	.487±.011
GPOP	.440±.009	.623±.015	.504±.011
RAND	.333±.007	.495±.011	.391±.009

Osaka			
Algo.	*Recall*	*Precision*	F_1-*score*
PT-.5F	.757±.025	.645±.037	.687±.032
PT-.5T	**.759±.026**	**.662±.037**	**.699±.033**
PT-1F	.679±.023	.582±.032	.616±.027
PT-1T	**.683±.025**	**.622±.032**	**.641±.029**
GNEAR	.478±.026	.433±.038	.441±.030
GPOP	.439±.034	.649±.038	.517±.035
RAND	.354±.021	.488±.032	.406±.024

Glasgow			
Algo.	*Recall*	*Precision*	F_1-*score*
PT-.5F	.819±.017	.758±.024	.780±.021
PT-.5T	**.826±.017**	**.782±.022**	**.798±.020**
PT-1F	**.748±.017**	.728±.022	.726±.019
PT-1T	.739±.018	**.736±.021**	**.728±.019**
GNEAR	.498±.020	.519±.028	.490±.022
GPOP	.418±.015	.592±.024	.480±.017
RAND	.340±.012	.462±.017	.386±.013

Edinburgh			
Algo.	*Recall*	*Precision*	F_1-*score*
PT-.5F	.740±.006	.607±.010	.654±.009
PT-.5T	**.740±.007**	**.633±.010**	**.671±.008**
PT-1F	**.678±.007**	.572±.009	.605±.008
PT-1T	.668±.007	**.601±.009**	**.618±.008**
GNEAR	.471±.007	.429±.010	.427±.008
GPOP	.486±.008	.640±.010	.539±.008
RAND	.336±.006	.479±.009	.384±.006

Personalized vs Non-personalized

Toronto		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**147.57±10.85**
	Non-personalized	152.44±9.84
PT-.5F	Personalized	**146.33±10.85**
	Non-personalized	152.61±10.09
PT-.5T	Personalized	**143.56±10.89**
	Non-personalized	150.65±10.09
PT-1F	Personalized	**137.07±11.40**
	Non-personalized	145.54±10.78
PT-1T	Personalized	**145.20±11.79**
	Non-personalized	148.18±11.29

Osaka		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**51.35±11.41**
	Non-personalized	54.91±11.91
PT-.5F	Personalized	**56.71±12.43**
	Non-personalized	60.06±13.09
PT-.5T	Personalized	57.09±12.39
	Non-personalized	**55.84±13.18**
PT-1F	Personalized	**56.62±13.21**
	Non-personalized	62.24±14.60
PT-1T	Personalized	**53.44±13.05**
	Non-personalized	58.88±14.63

Glasgow		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**75.98±11.53**
	Non-personalized	85.76±12.07
PT-.5F	Personalized	**88.20±13.03**
	Non-personalized	92.71±12.92
PT-.5T	Personalized	**76.40±11.34**
	Non-personalized	90.33±12.35
PT-1F	Personalized	**79.67±12.27**
	Non-personalized	86.24±12.85
PT-1T	Personalized	**73.29±11.94**
	Non-personalized	91.06±13.45

Edinburgh		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**91.08±4.85**
	Non-personalized	113.15±5.21
PT-.5F	Personalized	**84.56±4.96**
	Non-personalized	99.54±5.14
PT-.5T	Personalized	**89.76±5.85**
	Non-personalized	100.15±5.27
PT-1F	Personalized	**69.61±5.04**
	Non-personalized	78.89±5.31
PT-1T	Personalized	**72.11±6.09**
	Non-personalized	74.48±5.29

Personalized vs Non-personalized

Toronto		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**147.57±10.85**
	Non-personalized	152.44±9.84
PT-.5F	Personalized	**146.33±10.85**
	Non-personalized	152.61±10.09
PT-.5T	Personalized	**143.56±10.89**
	Non-personalized	150.65±10.09
PT-1F	Personalized	**137.07±11.40**
	Non-personalized	145.54±10.78
PT-1T	Personalized	**145.20±11.79**
	Non-personalized	148.18±11.29

Osaka		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**51.35±11.41**
	Non-personalized	54.91±11.91
PT-.5F	Personalized	**56.71±12.43**
	Non-personalized	60.06±13.09
PT-.5T	Personalized	57.09±12.39
	Non-personalized	**55.84±13.18**
PT-1F	Personalized	**56.62±13.21**
	Non-personalized	62.24±14.60
PT-1T	Personalized	**53.44±13.05**
	Non-personalized	58.88±14.63

Glasgow		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**75.98±11.53**
	Non-personalized	85.76±12.07
PT-.5F	Personalized	**88.20±13.03**
	Non-personalized	92.71±12.92
PT-.5T	Personalized	**76.40±11.34**
	Non-personalized	90.33±12.35
PT-1F	Personalized	**79.67±12.27**
	Non-personalized	86.24±12.85
PT-1T	Personalized	**73.29±11.94**
	Non-personalized	91.06±13.45

Edinburgh		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**91.08±4.85**
	Non-personalized	113.15±5.21
PT-.5F	Personalized	**84.56±4.96**
	Non-personalized	99.54±5.14
PT-.5T	Personalized	**89.76±5.85**
	Non-personalized	100.15±5.27
PT-1F	Personalized	**69.61±5.04**
	Non-personalized	78.89±5.31
PT-1T	Personalized	**72.11±6.09**
	Non-personalized	74.48±5.29

Comparison vs Baselines

Toronto

Algo.	*Popularity*	*Interest*	*Rk*
PT-0	2.204±.069 (1)	0.904±.048 (5)	3
PT-.5F	2.053±.063 (2)	1.088±.060 (4)	3
PT-.5T	1.960±.064 (3)	1.223±.061 (2)	**2.5**
PT-1F	1.583±.048 (4)	1.137±.061 (3)	3.5
PT-1T	1.419±.044 (7)	1.351±.069 (1)	4
GNEAR	1.424±.049 (6)	0.773±.054 (6)	6
GPOP	1.566±.050 (5)	0.443±.029 (8)	6.5
RAND	0.581±.032 (8)	0.467±.037 (7)	7.5

Osaka

Algo.	*Popularity*	*Interest*	*Rk*
PT-0	1.263±.094 (1)	0.791±.166 (6)	3.5
PT-.5F	1.126±.095 (3)	1.151±.213 (3)	3
PT-.5T	1.144±.093 (2)	1.171±.206 (2)	**2**
PT-1F	0.809±.075 (5)	1.137±.211 (4)	4.5
PT-1T	0.737±.067 (6)	1.205±.211 (1)	3.5
GNEAR	0.500±.059 (7)	0.853±.183 (5)	6
GPOP	0.837±.062 (4)	0.223±.066 (8)	6
RAND	0.433±.055 (8)	0.305±.089 (7)	7.5

Glasgow

Algo.	*Popularity*	*Interest*	*Rk*
PT-0	1.701±.101 (1)	0.459±.069 (5)	3
PT-.5F	1.562±.089 (3)	0.563±.091 (3)	3
PT-.5T	1.601±.089 (2)	0.625±.084 (2)	**2**
PT-1F	1.128±.069 (5)	0.562±.090 (4)	4.5
PT-1T	1.001±.052 (6)	0.676±.096 (1)	3.5
GNEAR	0.874±.064 (7)	0.339±.070 (6)	6.5
GPOP	1.399±.075 (4)	0.217±.049 (8)	6
RAND	0.483±.048 (8)	0.229±.041 (7)	7.5

Edinburgh

Algo.	*Popularity*	*Interest*	*Rk*
PT-0	2.269±.046 (1)	1.047±.053 (5)	3
PT-.5F	2.016±.042 (2)	1.383±.068 (4)	3
PT-.5T	2.012±.043 (3)	1.579±.069 (2)	**2.5**
PT-1F	1.541±.038 (5)	1.430±.070 (3)	4
PT-1T	1.336±.034 (6)	1.722±.076 (1)	3.5
GNEAR	1.269±.033 (7)	0.939±.054 (6)	6.5
GPOP	1.775±.039 (4)	0.577±.033 (7)	5.5
RAND	0.656±.025 (8)	0.526±.033 (8)	8

Conclusions

- Time based outperforms frequency based interests. It represents better real life scenarios.
- This method gives personalized visit durations based on personal interests.

References

Tsiligirides, T. (1984).

Heuristic methods applied to orienteering.

Journal of the Operational Research Society, 35(9):797–809.