# Personalised Pathway Prediction

Fabian Bohnert and Ingrid Zukerman

Faculty of Information Technology, Monash University
Clayton, VIC 3800, Australia
{fabian.bohnert,ingrid.zukerman}@infotech.monash.edu.au

**Abstract.** This paper proposes a personalised frequency-based model for predicting a user's pathway through a physical space, based on non-intrusive observations of users' previous movements. Specifically, our approach estimates a user's transition probabilities between discrete locations utilising personalised transition frequency counts, which in turn are estimated from the movements of other similar users. Our evaluation with a real-world dataset from the museum domain shows that our approach performs at least as well as a non-personalised frequency-based baseline, while attaining a higher predictive accuracy than a model based on the spatial layout of the physical museum space.

## 1 Introduction

This paper proposes a personalised frequency-based model for predicting a user's pathway through a physical space, called *Personalised Transition Model (PTM)*. Our model utilises non-intrusive observations of users' previous movements to generate a pathway prediction, making the assumption that the movements of other like-minded users are more indicative of a target user's behaviour than those of dissimilar users. Specifically, *PTM* utilises personalised transition frequency counts to estimate a target user's transition probabilities between locations. We apply our model to a real-world scenario from the museum domain, by utilising it to predict a visitor's next few exhibits. Our results show that in our domain, *PTM* performs at least as well as a non-personalised frequency-based baseline. Additionally, our model attains a higher predictive accuracy than a transition model based on the spatial layout of the physical museum space.

Our application scenario is motivated by the need to automatically recommend exhibits to museum visitors, based on non-intrusive observations of visitors' movements in the physical space. Employing recommender systems in the museum domain is challenging, as predictions differ from recommendations (we do not want to recommend exhibits that visitors are going to see anyway). We will address this challenge by combining the prediction of a visitor's pathway through the museum with the exhibits predicted to be of interest to the visitor, e. g., [1]. This supports the recommendation of personally interesting exhibits that may be overlooked if the predicted pathway is followed.

## 2 Related Research

*PTM* extends our previous research on predicting a visitor's pathway through a physical museum space [2] by using *personalised* transition frequency counts to estimate a

P. De Bra, A. Kobsa, and D. Chin (Eds.): UMAP 2010, LNCS 6075, pp. 363–368, 2010.

visitor's transition probabilities from non-intrusive observations of visitors' previous movements. Other research projects that investigate techniques for personalising the museum experience include *PEACH* [3] for content presentation, *CHIP* [4] for exhibit recommendations based on explicit user input, and the system of Cantino *et al.* [5], which uses Markov decision processes to generate personalised tour proposals for museum visitors. Additional systems for predicting people's future pathways include [6–8]. Specifically, Han and Cho model and predict users' movements by combining Markov models with recurrent self-organising maps [6], Krumm's system uses a Markov model to make short-term route predictions for vehicle drivers [7], and Krumm and Horvitz's system utilises Bayesian inference to predict where a driver is going as a trip progresses, based on a history of the driver's destinations and data about driving behaviours [8].

## 3 Personalised Pathway Prediction from Non-intrusive Observations

Our *Personalised Transition Model (PTM)* predicts a user's next few locations (e. g., museum exhibits) from non-intrusive observations of users' previous movements. The model utilises a transition probability vector $\boldsymbol{p}^k$ which approximates the probabilities of moving between locations ($k$ denotes the number of previously visited locations). Specifically, $\boldsymbol{p}^k$'s element $\boldsymbol{p}^k(i)$ represents the probability of a user going from a current location $i_k$ to location $i$ (for all $i = 1, \ldots, n$, where $n$ is the cardinality of the set $I$ of all locations). More formally, the transition probability $\boldsymbol{p}^k(i)$ approximates $\Pr(X_{k+1} = i \mid I_a^k)$, i. e., the probability that the current user $a$'s $(k+1)$-th location is location $i$ (where $I_a^k$ is the sequence of locations visited by user $a$). The transition probabilities are updated whenever the current user moves to a new location. This section describes our approach of estimating $\boldsymbol{p}^k$ in a personalised frequency-based fashion.

For estimating the transition probabilities, we start by calculating the following similarity-weighted personalised frequency counts (for all $i = 1, \ldots, n$):

$$x_{i_k,i} = |N(a)| \, \frac{\sum_{u \in N(a)} sim_k(a,u) \; \mathbb{1}_u(i_k \to i)}{\sum_{u \in N(a)} sim_k(a,u)}, \tag{1}$$

where $N(a)$ is the set of nearest neighbours, $sim_k(a,u)$ is the similarity between users $a$ and $u$, and $\mathbb{1}_u(i_k \to i)$ indicates whether user $u$ went from location $i_k$ to location $i$.

We calculate $sim_k(a,u)$ by comparing the sequences of visited locations of users $a$ and $u$ (called $I_a^k$ and $I_u$ respectively). To this effect, we first determine whether user $u$ has visited user $a$'s current location $i_k$. If this is the case, we identify the transitions between pairs of locations that occur in both $I_a^k$ and $I_u^l$, where $I_u^l$ denotes the beginning of $I_u$ up to and including location $i_k$ (otherwise, if $i_k \notin I_u$, we set $sim_k(a,u) = 0$). We then calculate a discounted count of the identical transitions, where a transition is discounted according to how long before location $i_k$ the transition occurred in $I_a^k$ and $I_u^l$ (we use the inverse of the product of the number of visited locations from the transition until the current location $i_k$ in $I_a^k$ and $I_u^l$ as the discounting factor). The discounting is motivated by the fact that identical transitions immediately preceding the current location $i_k$ in $I_a^k$ and $I_u^l$ are more indicative of the users' pathway similarity $sim_k(a,u)$ around location $i_k$ than identical transitions that occurred earlier. Finally, we normalise the resultant sum to the interval $[0,1]$.[1]

The current user $a$'s set of nearest neighbours $N(a)$ is constructed by selecting up to $K_{\mathrm{NN}}$ users that are most similar to the current user $a$, from those users whose similarity $sim_k(a,u)$ is above a certain non-negative threshold $S$.

To smooth out outliers, we apply *additive smoothing* to the personalised frequency counts $x_{i_k,i}$ (Equation 1) by adding a smoothing constant $\alpha > 0$ (except for $x_{i_k,i_k}$, which is 0). Further, we set to 0 the smoothed personalised frequency counts that correspond to the visited locations, and normalise the values so that their sum is 1. By doing this, we focus on unseen locations (e. g., museum visitors rarely return to previously viewed exhibits). The resultant normalised values correspond to the personalised transition probabilities $\tilde{\boldsymbol{p}}^k(i)$, where $i = 1, \dots, n$.

We employ *shrinkage to the mean* to regularise the personalised transition probabilities $\tilde{\boldsymbol{p}}^k(i)$ by combining them with the transition probabilities $\boldsymbol{p}_{TM}^k(i)$ delivered by a non-personalised frequency-based *Transition Model (TM)* [2]. This yields *PTM*'s shrunken personalised transition probabilities $\boldsymbol{p}^k(i)$ as follows:

$$\boldsymbol{p}^k(i) = \boldsymbol{p}_{TM}^k(i) + \omega\left(\tilde{\boldsymbol{p}}^k(i) - \boldsymbol{p}_{TM}^k(i)\right),$$

where $\omega \in [0,1]$ is the shrinkage weight. If the set of nearest neighbours is empty (i. e., a similarity-weighted prediction $\tilde{\boldsymbol{p}}^k$ is not possible) or the current user $a$ has visited less than $M$ locations, we estimate the probabilities $\boldsymbol{p}^k(i)$ using simply $\boldsymbol{p}_{TM}^k(i)$.

In summary, *PTM* uses the following adjustable parameters when estimating $\boldsymbol{p}^k$: (1) smoothing constant $\alpha$, (2) the minimum number of visited locations $M$ (personalised prediction), (3) the minimum similarity $S$ (nearest neighbour), (4) the maximum number of nearest neighbours $K_{\mathrm{NN}}$, and (5) shrinkage weight $\omega$.

Having (re-)calculated the transition probabilities $\boldsymbol{p}^k(i)$ after every move, we predict a user's next $K$ locations using the *Sequence* $K$ approach, which finds the sequence of the $K$ unvisited locations $i_{k+1}, \dots, i_{k+K} \in I \backslash I_a^k$ that maximises the probability

$$\begin{aligned} &\Pr\left(X_{k+1} = i_{k+1}, \dots, X_{k+K} = i_{k+K} \mid I_a^k\right) \\ &= \prod_{m=1}^{K} \Pr\left(X_{k+m} = i_{k+m} \mid I_a^{k+m-1}\right) = \prod_{m=1}^{K} \boldsymbol{p}^{k+m-1}(i_{k+m}). \end{aligned}$$

Factorising the joint probability is possible due to $X_{k+m}$ depending only on the past. This enables maximisation by recursively spanning a search tree of depth $K-1$, and performing a search for a maximising path from its root to one of the leaves (we pruned the search tree by removing unlikely paths).

## 4 Evaluation

This section evaluates *PTM* with a real-world dataset of visitor pathways, which was obtained by manually tracking visitors at Melbourne Museum (Melbourne, Australia)

[1] We also experimented with similarity measures based on Hamming distance and Levenshtein distance, but their performance was inferior.

from April to June 2008. Specifically, we recorded 158 visitor pathways in the form of time-annotated sequences of visited exhibit areas, providing information of the type that may be automatically inferred from sensors. In total, the dataset (described in detail in [1]) contains 8327 stops at the 126 exhibit areas of Melbourne Museum.[2]

### 4.1 Experimental Setup

We implemented two baseline models to evaluate *PTM*'s performance: (1) a *Physical Distance Model (PDM)*, using the spatial layout of the museum space to estimate the transition probabilities (making the assumption that transitions to spatially close exhibits are exponentially more likely than those to exhibits that are farther away); and (2) a non-personalised frequency-based approach for estimating a visitor's location probabilities, called *Transition Model (TM)* [2].

Employing leave-one-out cross validation, we tested thousands of configurations of *PDM*, *TM* and *PTM* to assess the influence of the different parameters on the predictive performance of the models, and to determine the best-performing variants. Specifically, model assessment was done by comparing the *negative log probability (NLP)* scores of the various configurations (the NLP score represents the average of the negative logs of the probabilities with which the exhibits actually viewed next were predicted).[3]

We conducted two types of experiments for assessing the predictive accuracy of the best-performing model configurations in the *Sequence* $K$ prediction mode (as above, we used leave-one-out cross validation):

- **Overall Visit (OV).** *OV* evaluates *overall* performance for a museum visit. For each visitor, we started with an empty visit, and iteratively added each viewed exhibit to the visit history. For each iteration, we predicted the next $K$ exhibits, and added these predicted exhibits to a global set of predicted exhibits (ignoring duplicate predictions). At the end of a visit, we calculated *precision (Pre)*, *recall (Rec)* and *F-score* for the entire visit by comparing the accumulated set of all predicted exhibits to the set of actually viewed exhibits. The resultant values were averaged over all visitors.
- **Progressive Visit (PV).** *PV* evaluates *immediate* predictive model performance with the progression of a visit, i. e., as the number of viewed exhibit areas increases. For each *fraction of a visit*, we first predicted the next $K$ exhibits (we used visit fractions rather than the actual number of viewed exhibits, because different visits have different lengths). We then measured immediate classification accuracy by calculating $CA(K) = |\mathcal{K} \cap \mathcal{M}|/K$ (i. e., the proportion of the predicted sequence $\mathcal{K}$ of the next $K$ exhibits that appears in the sequence $\mathcal{M}$ of the next $K$ actually viewed exhibits; this measure equals immediate recall and precision), and averaged the resultant values over all visitors for each visit fraction.

[2] For our experiments, we ignore travel time between exhibit areas, and collapse multiple viewing events of one area into one event.

[3] The *PTM* configuration that achieves the minimum NLP score is $\{\alpha = 0.04,\ M = 3,\ S = 0.00,\ K_{\text{NN}} = 58,\ \omega = 0.40\}$, where the symbols are explained at the end of Sect. 3. We omit further results of a sensitivity analysis of *PTM*'s parameters due to space limitations.

**Table 1.** Model performance for the *OV* and *PV* experiments

| | *Sequence* 1 | | | | *Sequence* 3 | | | |
|---|---|---|---|---|---|---|---|---|
| | Pre | Rec | F-score | *CA*(1) | Pre | Rec | F-score | *CA*(3) |
| ***PDM*** | 59.03% | 46.13% | 51.73% | 45.06% | 54.27% | 65.28% | 59.08% | 45.93% |
| ***TM*** | 65.58% | 56.87% | 60.87% | 54.29% | 58.59% | 71.52% | 64.19% | 52.59% |
| ***PTM*** | **67.09%** | **58.33%** | **62.35%** | **55.97%** | **59.12%** | **72.90%** | **65.06%** | **54.11%** |

### 4.2 Results

This section presents the results of our evaluation for $K = 1$ and $K = 3$.

**Evaluation of Pathway Predictions for K = 1.** The results for $K = 1$ are summarised in the left-hand side of Table 1. For the *OV* experiment, the frequency-based models *TM* and *PTM* statistically significantly outperform the distance-based baseline *PDM*.[4] More importantly, *PTM* attains statistically significantly better results than *TM* with respect to all measures, which means that using personalised transition frequency counts is beneficial. For the *PV* experiment, all models perform at a relatively constant level with the progression of a visit, with *PTM* achieving the highest average *CA*(1) of 56% (averaged over 1000 equally-spaced visit fractions). Further, *TM* and *PTM* perform statistically significantly better than *PDM* for 66% and 82% of a visit respectively, and *PTM* performs consistently at least as well as *TM* (statistically significantly better than *TM* for 21% of a visit, while *TM* never outperforms *PTM*). These results indicate that other visitors' movements are better predictors of a visitor's next exhibit than the spatial layout of the museum. Additionally, personalisation aids prediction, as personalised *PTM* outperforms non-personalised *TM* for the *OV* experiment, and for some portion of a visit for the *PV* experiment.

**Evaluation of Pathway Predictions for K = 3.** The right-hand side of Table 1 summarises the results for $K = 3$. For the *OV* experiment (as for $K = 1$), the frequency-based models *TM* and *PTM* statistically significantly outperform the distance-based baseline *PDM*. Further, personalised *PTM* performs statistically significantly better than non-personalised *TM* with respect to all measures. For the *PV* experiment (as for $K = 1$), all models perform at a relatively constant level with the progression of a visit (*PTM* attains an average *CA*(3) of 54%). In addition, *TM* and *PTM* perform statistically significantly better than *PDM* for 63% and 77% of a visit respectively, and *PTM* attains a statistically significantly higher classification accuracy *CA*(3) than *TM* for 31% of a visit (*TM* never outperforms *PTM*). These results are consistent with those for $K = 1$. Comparing the *Sequence* 3 variants of our models with the *Sequence* 1 variants for the *OV* experiment, precision for *Sequence* 3 is lower while recall is higher. This is because at each stage of a visit, predicting the next three exhibits leads to a larger accumulated set than predicting only the next exhibit (higher recall). However, the predictions for three exhibits are less likely to be correct than those for one exhibit (lower precision). For the *PV* experiment, *PTM* outperforms *TM* for a slightly longer portion of a visit for $K = 3$ compared to $K = 1$ (31% vs. 21%).

[4] The statistical tests performed are one-tailed paired t-tests (significance level $\alpha = 0.05$).

## 5 Conclusions and Future Work

This paper proposed a frequency-based *Personalised Transition Model (PTM)* for predicting a user's pathway through a physical space. Specifically, our model estimates a target user's transition probabilities between discrete locations utilising personalised transition frequency counts, which in turn are estimated from the movements of other like-minded users by means of a nearest-neighbour collaborative approach (using a pathway-based similarity measure). We evaluated *PTM* by predicting a museum visitor's next $K = 1$ and $K = 3$ exhibits, and showed that in our scenario (1) *PTM* and *TM* (a non-personalised frequency-based baseline) outperform a distance-based baseline, which means that other people's movements are better predictors of a visitor's pathway than the spatial layout of the museum; and (2) personalisation aids prediction, as *PTM* outperforms *TM* for the overall measures (recall, precision and F-score) and for at least some portion of a visit for the realistic *Progressive Visit* experiment.

Overall, *PTM* yields only a modest (yet statistically significant) improvement over *TM*. The small size of this improvement may be due to the sparsity problem, i. e., the small size of our dataset, which contains only a few visitors with very similar pathways. In the future, we intend to apply *PTM* to larger datasets to investigate this further. We also plan to investigate models for predicting longer location sequences.

**Acknowledgements.** This research was supported in part by grant DP0770931 from the Australian Research Council. The authors thank Museum Victoria for its assistance; and David Abramson and his team for their help with the computer cluster.

## References

1. Bohnert, F., Zukerman, I.: Non-intrusive personalisation of the museum experience. In: Houben, G.-J., McCalla, G., Pianesi, F., Zancanaro, M. (eds.) UMAP 2009. LNCS, vol. 5535, pp. 197–209. Springer, Heidelberg (2009)
2. Bohnert, F., Zukerman, I., Berkovsky, S., Baldwin, T., Sonenberg, L.: Using interest and transition models to predict visitor locations in museums. AI Communications 21(2-3), 195–202 (2008)
3. Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik, T., Not, E., Rocchi, C.: Adaptive, intelligent presentation of information for the museum visitor in PEACH. User Modeling and User-Adapted Interaction 18(3), 257–304 (2007)
4. Wang, Y., Aroyo, L., Stash, N., Sambeek, R., Schuurmans, Y., Schreiber, G., Gorgels, P.: Cultivating personalized museum tours online and on-site. Interdisciplinary Science Reviews 34(2), 141–156 (2009)
5. Cantino, A.S., Roberts, D.L., Isbell, C.L.: Autonomous nondeterministic tour guides: Improving quality of experience with TTD-MDPs. In: Proc. of the 6th Intl. Joint Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS-07), pp. 91–93 (2007)
6. Han, S.J., Cho, S.B.: Predicting user's movement with a combination of self-organizing map and Markov model. In: Kollias, S.D., Stafylopatis, A., Duch, W., Oja, E. (eds.) ICANN 2006. LNCS, vol. 4132, pp. 884–893. Springer, Heidelberg (2006)
7. Krumm, J.: A Markov model for driver turn prediction. In: Proc. of the Society of Automotive Engineers (SAE) 2008 World Congress (2008) Paper 2008-01-0195
8. Krumm, J., Horvitz, E.: Predestination: Inferring destinations from partial trajectories. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 243–260. Springer, Heidelberg (2006)