Assessing the Impact of Measurement Uncertainty on User Models in Spatial Domains

Daniel F. Schmidt, Ingrid Zukerman, and David W. Albrecht

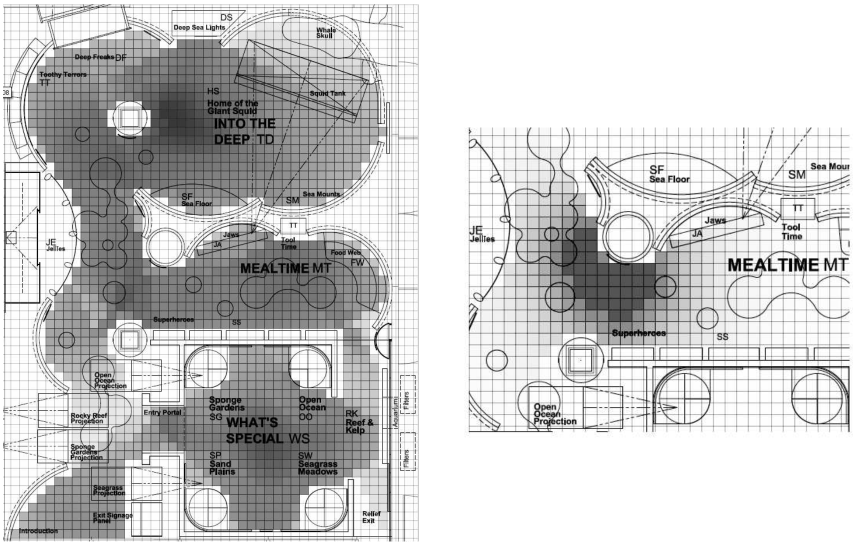
Faculty of Information Technology
Monash University, Clayton 3800, Australia
{daniel.schmidt,ingrid.zukerman,david.albrecht}@infotech.monash.edu.au

Abstract. This paper examines the problem of uncertainty due to instrumentation in user modeling systems within spatial domains. We consider the uncertainty of inferring a user's trajectory within a physical space combined with the uncertainty due to inaccuracies in measuring a user's position. A framework for modeling both types of uncertainties is presented, and applied to a real-world case study from the museum domain. Our results show that this framework may be used to investigate the effects of layout in a gallery, and to explore the degradation in the predictive performance of user models due to measurement error. This information in turn may be used to guide the curation of the space, and the selection of sensing technologies prior to instrumenting the space.

1 Introduction

Advances in mobile computing and sensing technology have enabled the instrumentation of physical spaces in order to track the movements of people and model their behaviour [1, 2]. Typically, these systems are implemented by equipping the space or the users with sensing technology, and applying machine learning or probabilistic techniques to build user models from logged sensor input [3, 4]. In principle, this approach appears to be sound. However, in practice it may be error prone, and hence expensive, as the selected sensor technology or configuration may turn out to be inadequate for the task. Further, compared to virtual spaces, physical spaces pose additional challenges to user modeling, owing to the inaccuracies inherent in sensory observations.

In this paper, we propose a framework for investigating the impact of different sensing technologies on the predictive performance of user models *prior* to deploying a particular technology. To this effect, we simulate sensor logs of users, and compare the predictive performance of a user model derived from these logs with that of a user model derived from perfect observations. Our framework was implemented in the context of the Marine Life Exhibition at Melbourne Museum (Figure 1(a)), where the derived models were used to predict exhibits viewed by museum visitors (the perfect observations were obtained by manually recording the exhibits actually viewed by visitors [5]). These predictions will eventually be used by a recommendation system that suggests exhibits of interest.



(a) Entropy mapping of the exhibition: darker colour indicates higher entropy.

(b) Probability of viewing exhibit 'Eat or be Eaten' from a square: darker colour indicates higher probability.

Fig. 1. The Marine Life Exhibit at Melbourne Museum

Our approach requires the following models: (1) a predictive user model of exhibits to be viewed, which also provides an upper bound of performance; (2) a spatial viewing model representing positions from which each exhibit can be seen; and (3) models of sensor characteristics for different types of sensors. Our predictive model is built from logs obtained by manually tracking the exhibits viewed by visitors [5]. To link this information to logs that can be obtained from sensors, we first need to infer a plausible viewing position for each exhibit; we employ the spatial viewing model to make this inference. Owing to sensor inaccuracies, a visitor's position recorded by a sensor may differ from his/her actual position. The nature and magnitude of this difference depends on the type of sensor (and even on the specific sensor). Models of sensor behaviour are needed to incorporate such distortions into the inferred position of a visitor.[1]

This paper is organized as follows. Sections 2-4 describe our three models. Section 5 describes the integration of the predictive user model and the spatial viewing model to generate synthetic pathways through the exhibition and predict viewed exhibits from positional information. The results of our evaluation are presented in Section 6, followed by concluding remarks.

[1] In principle, we could manually record a person's position and viewed exhibit directly. However, our experience shows that recording two separate information items at the same time places an excessive burden on human trackers, making their logs more error prone. More importantly, one of our objectives is to produce useful insights from a relatively small amount of easily recorded information.

2 Predictive User Model

There is a range of statistical models used in collaborative systems for predicting users' interests from observed behaviour, e.g., [6–8]. These systems have focused on making predictions in the virtual rather than the physical space. The body of work pertaining to building predictive models from sensory observations in physical spaces is more reduced, e.g., [3–5], with [5] being the main proponent of such models for the museum domain.

Our approach requires a user model that exhibits good predictive performance, and can be easily sampled from to generate synthetic visitors (Section 5.1). In this paper, we have adopted Bohnert *et al.*'s *Transition Model* to represent visitors' movements between exhibits in a museum [5]. This model is a stationary 1-stage Markov model, where $\mathrm{P}_{i,j}$ approximates the probability of moving from exhibit i to exhibit j ($i, j = 1, \ldots, M$ and M is the number of exhibits). The Transition Model has a reasonable predictive performance on a homogeneous exhibition, such as Marine Life, where visitors' behaviour is mainly determined by the layout of the exhibition [5] (a hybrid model combining interest with transitions outperforms a pure Transition Model, but it includes a non-parametric component, which makes sampling more difficult).

The main issue in fitting the Transition Model is estimating the transition probabilities from the available traces. The *sparse data problem* (also known as the 'small n, large p' problem) occurs when there is a small number of data points n compared to the number of parameters p to be estimated (in our case $p = M^2$, $M = 22$ exhibits, and $n = 317$ total exhibits viewed by 44 visitors). As a result of this problem, many transitions between exhibits have zero observed counts. Hence, estimating transition probabilities using a method such as Maximum Likelihood will lead to zero transition probabilities for these transitions (even when there is no physical reason for this to happen).

To overcome the sparse data problem, we employ a Bayesian approach, where our prior distribution over the possible transition probabilities $(\mathrm{P}_{i,1}, \ldots, \mathrm{P}_{i,M})$ from a particular exhibit i is given by a Dirichlet distribution, $\mathrm{Dir}(\alpha_i, \ldots, \alpha_i)$ (i.e., all the parameters have been set to α_i). The posterior distribution of these transition probabilities is given by another Dirichlet distribution, $\mathrm{Dir}(n_{i,1} + \alpha_i, \ldots, n_{i,M} + \alpha_i)$, where $n_{i,j}$ is the number of times a user was observed to have moved from exhibit i to j. To estimate the probabilities $\mathrm{P}_{i,j}$ it is common to employ the mean *a posteriori* estimates

$$\hat{\mathrm{P}}_{i,j} = \frac{n_{i,j} + \alpha_i}{N_i + M\alpha_i} \tag{1}$$

where $N_i = \sum_{k=1}^{M} n_{i,k}$ is the total number of times visitors viewed exhibit i, and α_i can be interpreted as the number of *a priori* observed counts per exhibit.[2]

However, as Hausser and Strimmer [9] point out, there is no general agreement regarding the value of α_i. Moreover, they demonstrate that for a small n and large p, in terms of Mean Square Error, a better estimate of $\mathrm{P}_{i,j}$ is obtained by choosing the α_i for an exhibit i to be

[2] α_i is often assigned a single value, in which case it is called a 'flattening constant'.

$$\alpha_i = \frac{N_i \left(1 - \sum_{k=1}^{M} \hat{\theta}_{i,k}^2\right)}{M \left(\sum_{k=1}^{M} \hat{\theta}_{i,k}^2 + (N_i - 1) \sum_{k=1}^{M} \left(1/M - \hat{\theta}_{i,k}\right)^2 - 1\right)} \tag{2}$$

where $\hat{\theta}_{i,j} = n_{i,j}/N_i$ denotes the Maximum Likelihood estimate of the transition probability from exhibit i to exhibit j.

3 Spatial Exhibit Viewing Model

One of the most interesting and difficult aspects of instrumenting a space such as a museum is inferring abstract concepts, such as interest or intention, from measured coordinates. In the museum domain, the time a visitor spent viewing an exhibit is treated as proportional to his/her interest, and thus provides a form of implicit rating, which can then be used by a recommender system. Therefore, in order to infer interest from measurements, one must be able to infer which exhibit is being viewed by a visitor when standing in a particular place.

Our approach consists of building a probabilistic model of the viewing areas for each exhibit in the physical space. To facilitate this, we divided the physical space into a grid. The dimensions of the grid were chosen to balance level of detail with computational expense (a fine-grained grid provides a lot of detail, but its integration into a viewing model is computationally expensive). The actual grid size we chose is $61 \times 47 = 2,867$ squares, where a square is about 30cm $\times$ 30cm.

At each square we placed a multinomial distribution which represents the probability that a visitor standing at that square is observing each exhibit. Figure 1(b) illustrates this distribution for the 'Eat or be Eaten' exhibit in the Marine Life Exhibition — the darker squares indicate a higher probability of viewing the exhibit from there. We now need to specify the probability of viewing each exhibit from each square. This was done as follows. We observed visitors' behaviour in the Marine Life exhibition, and for each exhibit, marked out areas on the grid where people stood most often to view the exhibit. These high probability areas (the darkest in Figure 1(b)) were assigned an unnormalized probability of 1. The probability of the remaining areas was determined by making $G(i; x, y)$ – the unnormalized probability of viewing exhibit i from square (x, y) – proportional to the distance between square (x, y) and the closest high-probability square (x_h, y_h) as follows

$$G(i; x, y) \propto \exp\left(-\frac{(x - x_h)^2 + (y - y_h)^2}{\lambda_i}\right)$$

where λ_i is chosen to control the rate of decay. This parameter, which may differ for each exhibit, reflects how large the viewing areas are in the physical space, and must be chosen by the space modeler. Based on our observations, we chose the same λ_i $(= 3)$ for all the exhibits $(i = 1, \ldots, M)$, as the gallery space was quite homogeneous.

Clearly, squares from which it is not physically possible to view an exhibit should have a zero probability associated with them. This is handled by simply marking the squares that correspond to walls, and setting $G(i; x, y)$ to zero if a straight line cannot be drawn from square (x, y) to any of the exhibit squares.

The final probability of viewing exhibit i from square (x, y) is estimated by normalizing over all the exhibits

$$\mathrm{P}(i|x, y) = \frac{G(i; x, y)}{\sum_{j=1}^{M} G(j; x, y)} \tag{3}$$

An interesting use of our viewing model is for assessing the clutter in a gallery. For each square in the gallery, we have an M-state multinomial distribution over the M exhibits (i.e., a list of probabilities that a visitor standing at that square is observing each exhibit). We represent the clutter at each square by the entropy of this multinomial. Specifically, the entropy $H(\mathbf{P})$ of an M-nomial with probabilities $\mathbf{P} = (\mathrm{P}_1, \ldots, \mathrm{P}_M)$ is given by $H(\mathbf{P}) = -\sum_{j=1}^{M} \mathrm{P}_j \log \mathrm{P}_j$. $H(\mathbf{P})$ is maximized when all exhibits are equally likely to be viewed, and is minimized when one $\mathrm{P}_j = 1$ and the rest are zero (i.e., there is no uncertainty). Figure 1(a) illustrates the entropy of each viewing square in the Marine Life Exhibition; dark shading indicates high entropy, and light shading indicates low entropy.

4 Sensor Models

The final component of the user modeling system is the sensor model. This component allows us to simulate real-life sensing technologies that are required to deploy a predictive model in a physical space. Ideally, a sensor model should be simple enough to easily integrate into a user modeling system, and also abstract enough to be able to represent a wide range of real-life sensing possibilities. A suitably abstract sensor model would allow different types of sensing technologies to be simulated by changing several parameters. Our basic model is that the measured coordinates (x', y') are a realization of a random variable whose probability distribution depends on the true location (x, y) and the type of sensor technology deployed. This distribution should be chosen to represent the behaviour of some real-life sensor technology. Below we propose models for indoor GPS, RFID tags and accelerometers. Our evaluation is based on our indoor GPS model (Section 6).

Indoor GPS or localization technologies. We adopt a simple model whereby the (x, y) coordinates are distorted by additive Gaussian noise. Under this regime, the *measured* coordinates are found by sampling from a bivariate normal distribution $N((x, y), \mathbf{C})$ with mean (x, y) and covariance $\mathbf{C}$. Usually one can assume that the accuracy is the same in all directions, and so we can make the simpler model choice $\mathbf{C} = \sigma^2 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, and σ is a constant that is chosen to reflect the expected accuracy of the device. For example, if the GPS is nominally accurate to within ν meters, and the (x, y) coordinates are measured in meters, then $\sigma = \nu/2$ would be a suitable value, as such a choice places the bulk of the probability mass within the circle defined by $x^2 + y^2 = \nu^2$.

RFID-tag arrays. An array of RFID tags positioned through the physical space can be modeled in a similar fashion to an indoor GPS. Given an array of active RFID tags, the physical space is divided into a set of (possibly overlapping) cells that are covered by the RFID tags — each cell potentially spanning several squares. When a user (wearing a passive RFID tag) moves into a cell, and the active RFID is activated, the system is aware of the user's approximate position. The uncertainty of the user's exact (x, y) position within a cell may be modeled by treating the measured (x', y') as a Gaussian distribution $N((a_k, b_k), \mathbf{C}_k)$, where (a_k, b_k) are the coordinates of the center of the cell covered by active RFID tag k, and the covariance matrix $\mathbf{C}_k$ is chosen to approximate the area of the cell. A more refined model of a user's position may be devised by treating the different cells as discrete states in a Markov model, and noting that a user's likely (x, y) position on entering a cell depends on the previous cell s/he was in (and thus, the direction from which s/he moved into the new cell). However, the Gaussian model will be insufficient to represent this extra information.

Accelerometer based sensing. The behaviour of accelerometer-based technology may be modeled as a state-space evolution of (x, y) coordinates with suitable Gaussian process noise over acceleration (rather than position, as for the previous devices). In order to model the behaviour of accelerometer-based technology, we need to simulate trajectories of users' (x, y) coordinates through the physical space. Such trajectories should include a sequence of points in the path between two consecutively visited exhibits, and the time required to traverse this path. The path may be approximated using a shortest path algorithm, and the traversal time may be approximated using average speeds of visitors and the length of the path. This information enables the calculation of acceleration vectors $(\ddot{x}, \ddot{y})_t$ at each time t along the trajectory. Thus, given a starting position $(x, y)_0$, the *measured* positions of the user (x', y') evolve over time according to the state-space equations

$$(x', y')_{t+\delta} = (x', y')_t + (\dot{x}', \dot{y}')_t \delta \tag{4}$$

$$(\dot{x}', \dot{y}')_{t+\delta} = (\dot{x}', \dot{y}')_t + (\ddot{x}', \ddot{y}')_t \delta + \boldsymbol{\varepsilon}_t \tag{5}$$

where $\boldsymbol{\varepsilon}_t$ is distributed as $N((0, 0), \mathbf{C})$, and $\mathbf{C}$ is a suitable covariance matrix representing the noise due to imperfect measurement of acceleration.

A major problem with acceleration-based tracking is *measurement drift*. That is, this technology produces relative positions (in contrast with absolute positions generated by GPS and RFID tags). Hence, the noise distorting the acceleration measurements will cause the estimated positions to increasingly drift away from the truth; the longer the sequence of acceleration measurements, the bigger the expected drift. This problem may be alleviated by deploying several absolute positioning devices (e.g., RFID tags) around the space, and using them to reset a user's position when s/he moves past them.[3]

[3] If a Kalman filter [10] is used to estimate a user's current position from the state-space model (Equations 4 and 5), the effect of resetting a user's position to within the

5 Integrating the User Model with the Viewing Model

The Transition Model presented in Section 2 is based on precise knowledge of the last exhibit viewed by a visitor. However, if information on the visitor's behaviour is being automatically gathered by instruments, then all that is available is a sequence of (possibly distorted) (x, y) coordinates. Assuming that there exists some criterion for detecting that a visitor is stationary (and hence viewing an exhibit), we can decompose the complete (x, y) sequence into a sub-sequence of *stationary* (x, y) coordinates (at present, we do not model 'hovering' around an exhibit). From these, we must attempt to infer which exhibit the visitor is viewing, and then employ our user model to predict which exhibit the visitor will view next on the basis of this information.

Recall that our manually gathered data consists of a sequence of viewed exhibits (rather than (x, y) coordinates). Hence, in order to make predictions from (x, y) coordinates, we must first generate positional pathways from information regarding viewed exhibits. We generate synthetic pathways, driven by the predictive user model, rather than pathways tailored to the 44 observed users. This is done to ensure that any deterioration in predictive performance can be attributed to the use of positional coordinates (instead of precise exhibits) and to sensing distortion due to instrumentation error. Synthetic users were also generated by [11] for plan-based activities in the virtual space. However, their objective was to generate new, plausible users, while ours is to filter out prediction errors made by the user model.

5.1 Generating User Pathways

We generate realistic synthetic trajectories for the Marine Life Exhibition as follows. We first use the Transition Model (Section 2) to generate a tour (ordered list) of viewed exhibits, and then apply our spatial viewing model (Section 3) to transform this list of exhibits to a sequence of plausible spatial coordinates within the physical space (these coordinates are subsequently distorted by measurement error).

Generating a tour. The Transition Model proposed in Section 2 assumes that the primary driving force behind a tour is the layout of the gallery, and that the probability that a visitor moves to an exhibit depends entirely on the last exhibit viewed. As mentioned above, this model yields reasonable predictions for our dataset. Also, Markov models are easy to sample from, thus facilitating the generation of synthetic tours. Our tours begin at a (fictitious) 'start' position, with the first viewed exhibit being drawn from the possible transitions from this position; the next exhibit is drawn from the transition probabilities of the first exhibit, and so on, until a (fictitious) 'end' exhibit is drawn.

accuracy provided by an RFID tag may be naturally incorporated into the Kalman filtering process by setting the covariance matrix of the Kalman-filter state estimate to the covariance matrix of the RFID-tag noise model.

The generated tours depend on the estimates of the transition probabilities obtained in Section 2. Due to the small amount of data used to estimate the multinomial distributions that comprise the Transition Model, these estimates have a large variance. This variance is not taken into account when point estimates, such as those in Equation 1, are employed to define the multinomial distributions from which samples are drawn. To overcome this problem, we use the complete Bayesian predictive distribution to generate tours, as follows. For each exhibit i in a tour, we first sample $\phi_1, \ldots, \phi_M$ from the Dirichlet distribution $\mathrm{Dir}\,(n_{i,1} + \alpha_i, \ldots, n_{i,M} + \alpha_i)$ for the exhibit ($n_{i,j}$ and α_i are given in Section 2), and then sample the next exhibit from the multinomial distribution $\mathrm{Multi}(\phi_1, \ldots, \phi_M)$. For small samples sizes, if one was to use a point estimate to generate tours, the resulting tours would contain less variability than warranted by the data. The full predictive distribution takes this overdispersion into account, yielding tours with higher variability.

Generating User Coordinates. Once a tour of exhibits has been generated, we need to place the visitors in physical (x, y) coordinates within the gallery in a plausible fashion. We employ Bayes' theorem to produce the probability of a visitor being at square (x, y) conditioned on the fact that s/he has been viewing exhibit i, where $x = 1, \ldots, 61$ and $y = 1, \ldots, 47$ (Section 3). This yields

$$\mathrm{P}(x, y|i) = \frac{\mathrm{P}(i|x, y)\pi(x, y)}{\sum_x \sum_y \mathrm{P}(i|x, y)\pi(x, y)}$$

where $\mathrm{P}(i|(x, y)$ is obtained from Equation 3.

It remains to specify a prior distribution $\pi(\cdot)$ over the possible (x, y) positions where a visitor may be standing. Assuming a simple prior of ignorance, whereby every square is equally likely to be occupied by a visitor, we obtain

$$\mathrm{P}(x, y|i) = \frac{\mathrm{P}(i|x, y)}{\sum_x \sum_y \mathrm{P}(i|x, y)} \tag{6}$$

Now, given that a synthetic visitor is viewing exhibit i, we just need to sample from a multinomial distribution representing all the squares in the space to determine a square occupied by the visitor.

5.2 Predicting Exhibits from Positional Information

When a visitor stands in a particular location, there is some uncertainty regarding which exhibit s/he is viewing. The more exhibits are in close proximity (i.e., the more cluttered is an exhibit area), the higher the uncertainty. We consider two approaches for inferring viewed exhibits from positional information in light of this uncertainty: *Argmax* and *Weighted*.

– ***Argmax*** selects the most probable exhibit given the coordinates (x, y) of the user, i.e.,
$$j_{\max} = \underset{j \in \{1, \ldots, M\}}{\arg\max} \; \{\mathrm{P}(j|x, y)\} \tag{7}$$

The Transition Model is then used to estimate the probability of the next exhibit i assuming that the current exhibit being viewed is $j_{\max}$

$$\hat{\mathrm{P}}(i|x,y) = \mathrm{P}_{j_{\max},i} \tag{8}$$

– ***Weighted*** employs the Transition Model to estimate the probability of going to the next exhibit i from each other exhibit in the gallery, and calculates a weighted average of these probabilities on the basis of the probability of viewing each exhibit from coordinates (x, y).

$$\hat{\mathrm{P}}(i|x,y) = \sum_{j=1}^{M} \{\, \mathrm{P}(j|x,y) \times \mathrm{P}_{j,i} \,\} \tag{9}$$

We expect the differences in the performance of Argmax and Weighted to be greatest when the (x, y) coordinates are in areas of high uncertainty, i.e., areas with many exhibits. When only one exhibit is feasibly viewable from a particular (x, y) coordinate, the two predictors are expected to coincide (Section 6).

6 Evaluation

We first review the data collection process, followed by a description of our experiments and the results we obtained.

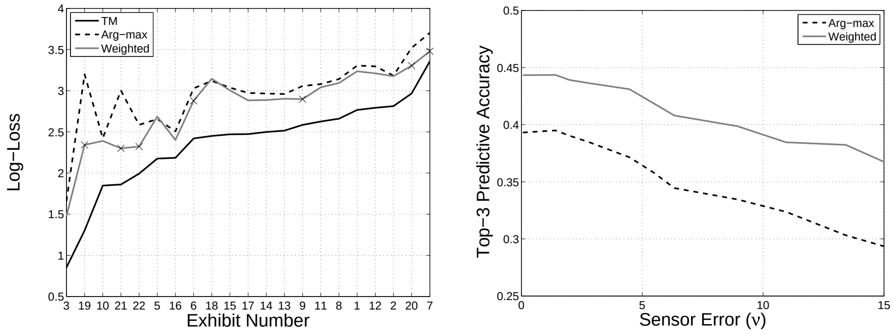
6.1 Data collection

As mentioned above, our framework was evaluated on data obtained from the Marine Life Exhibition at Melbourne Museum (Figure 1). The dataset, which was gathered manually, consists of tour traces from 44 visitors (Section 2). These traces contain an ordered list of the exhibits viewed by each visitor, and the time spent at each exhibit (which is not used in our models, but is necessary for assessing interest [5]). There are $M = 22$ exhibits in the Marine Life Exhibition, and on average, a visitor viewed 7.2 exhibits. We augmented the exhibit list with fictitious 'start' and 'end' exhibits in order to naturally incorporate an initial and final event into our predictive model.

The data for the viewing model were obtained separately from the user modeling data. This was done by observing the movements of visitors to the Marine Life Exhibition as they viewed the exhibits, and manually annotating a grid-divided map of the gallery to record their positions (Section 3).

6.2 Experimental setup

We conducted two experiments as follows. First we evaluated the performance of our two position-based prediction models, Argmax and Weighted (Section 5.2), compared with the performance obtained by the Transition Model alone, i.e., from direct observations of the exhibits viewed. We then introduced distortions modeled by our indoor GPS sensor model (Section 4) into the position-based predictive models in order to examine the effect of sensor inaccuracy on predictive performance. For both experiments we generated 1000 synthetic tours as described in Section 5.1.



(a) Log-loss at each exhibit, plotted in decreasing order of predictive performance of the Transition Model.

(b) Degradation in predictive performance for indoor GPS for the Weighted and Argmax models.

Fig. 2. Results for the Marine Life Exhibition at Melbourne Museum

6.3 Results

Figure 2(a) shows the results obtained by the three predictive models, Argmax, Weighted and Transition Model, in terms of the average log-loss of the predictions made at each of the 22 exhibits (i.e., the negative-log of the probability with which the exhibit actually viewed next was predicted). This average, which is calculated over all the synthetic visitors, summarizes how well the predictive models perform at each exhibit. The curve is plotted in order of decreasing predictive performance of the Transition Model, and the crosses mark exhibits for which the difference between the Argmax and Weighted models is statistically significant at the 0.05 level. For the position-based models, this plot was produced on the basis of the (x, y) coordinates where each synthetic visitor 'stood' to view each exhibit in his/her tour. The predictions made by the Transition Model were obtained directly from viewed exhibits. These predictions represent an upper bound on predictive performance (lower bound on log-loss).

In general, the Weighted method outperforms the Argmax method. When Weighted does better than Argmax, as for Exhibit 19 ('Tool Time') and 21 ('Deep Freaks'), the difference is quite substantial. In contrast, when Argmax outperforms Weighted, as for exhibit 18 ('Sea Floor'), the difference is rather marginal. It is worth noting that the viewing areas for both 'Tool Time' and 'Deep Freaks' have a significant amount of overlap with the viewing areas of other exhibits, while 'Sea Floor' is quite separate from its neighbour 'Sea Mounts'. Thus, these models behave according to the expectations set out in Section 5.2.

The effect of instrumentation accuracy on predictive performance was tested with respect to indoor GPS — the instrumentation option being considered at present by Melbourne Museum. We calculated the average predictive accuracy of our models under various levels of sensor noise (the average was computed over all the exhibits in all the tours). The predictive accuracy of a model was calculated by scoring 1 if one of the top-3 most probable exhibits was viewed next, and 0 otherwise. We chose top-3 (rather than top-1) because top-1 ignores the fact that top probabilities are often quite similar in our scenario. Figure 2(b) shows the degradation in the predictive performance of the Weighted and Argmax models as a function of increasing sensor error ν (Section 4). The true prediction error baseline, which is obtained when the viewed exhibit is known, is 0.54 on average (0.32 for top-1 accuracy),[4] compared with 0.44 for Weighted and 0.39 for Argmax when $\nu = 0$. This drop in performance as one changes from precise observations to positional observations may be largely attributed to the clutter in the gallery (recall that error due to the predictive model has been filtered out, since this model is also used to generate the synthetic pathways). Our results show that performance degrades slowly as sensor error increases. For instance, an error of $\nu = 5$ squares (1.5 meters) results in only approximately 10% drop in performance. This indicates that a fairly inaccurate sensor technology or a fairly coarse instrumentation of the museum space may be suitable, which can significantly lower instrumentation costs. At the same time, more accurate predictive user models may be necessary to improve the baseline predictive performance, possibly in combination with a reduction in the clutter of certain exhibit areas.

7 Conclusion and Future Work

We have offered a framework for investigating the impact of sensing technologies on the predictive performance of user models in physical spaces. Our framework combines a predictive user model with a spatial viewing model to produce pathways of synthetic users from a relatively small dataset. It then incorporates simulated sensing distortions from different types of instruments. This framework was applied to a small, real-life dataset obtained from Melbourne Museum. Our results show that the Weighted position-based predictive model outperforms the Argmax model, and that the Weighted model can attain tolerable predictive performance, even in the presence of a substantial sensory distortion.

There are several interesting avenues for further investigation. Firstly, we propose to implement models of the other positioning devices mentioned in Section 4, viz RFID tags and accelerometers. In addition, in order to improve the realism of our models we intend to do the following.

– Devise a more accurate spatial viewing model by considering particular restrictions of museums. For example, our model could reduce the probabilities of squares that are too close to walls or exhibits, and take into account size of exhibits (bigger exhibits are more likely to be viewed from farther away than smaller exhibits). The association of suitable attributes with exhibits will in turn enable the application of machine learning techniques to learn models of viewing areas for new exhibits.

[4] The predictive performance of the baseline model is lower than that in [5] due to our sampling approach, which generates tours with higher variability.

– Combine a tour generated from a predictive user model and the (x, y) coordinates generated from the spatial model into a dynamic trajectory through the physical space (rather than just stops at particular exhibits). This requires the derivation of a path between exhibits (e.g., by using a shortest-path algorithm), and a suitable stopping criterion to determine when a visitor has paused to interact with an exhibit or is 'hovering' around the exhibit. This criterion could be based on factors such as the direction and velocity of a visitor's approach to a particular square.

Acknowledgements. This research was supported in part by a grant from the Monash Research Fund and grant DP0770931 from the Australian Research Council. The authors thank Carolyn Meehan and the staff of Melbourne Museum for their assistance in data collection and museum mapping, and Fabian Bohnert for his help with the modeling task.

References

1. Hightower, J., Borriello, G.: Location systems for ubiquitous computing. IEEE Computer **34**(8) (2001) 57–66
2. Carmichael, D.J., Kay, J., Kummerfeld, B.: Consistent modelling of users, devices and sensors in a ubiquitous computing environment. User Modeling and User-Adapted Interaction **15**(3-4) (2005) 197–234
3. Horvitz, E., Apacible, J., Sarin, R., Liao, L.: Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service. In: UAI05 – Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, Edinburgh, Scotland (2005) 275–280
4. Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H., Hahnel, D.: Inferring activities from interactions with objects. IEEE Pervasive Computing **3**(4) (2004) 50–57
5. Bohnert, F., Zukerman, I., Berkovsky, S., Baldwin, T., Sonenberg, L.: Using interest and transition models to predict visitor locations in museums. AI Communications – Special Issue on Recommender Systems **21**(2-3) (2008) 195–202
6. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: GroupLens: An open architecture for collaborative filtering of Netnews. In: CSCW-94 – Proc. of the 1994 ACM Conf. on Computer Supported Cooperative Work. (1994) 175–186
7. Herlocker, J.L., Konstan, J.A., Borchers, A., Riedl, J.: An algorithmic framework for performing collaborative filtering. In: SIGIR-99 – Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. (1999) 230–237
8. Bell, R., Koren, Y., Volinsky, C.: Chasing $1,000,000: How we won the Netflix progress prize. ASA Statistical and Computing Graphics Newsletter **18**(2) (2007) 4–12
9. Hausser, J., Strimmer, K.: Entropy inference and the James-Stein estimator. URL:http://arxiv.org/abs/0811.3579 (November 2008)
10. Gelb, A.: Applied Optimal Estimation. MIT Press (1974)
11. Blaylock, N., Allen, J.: Generating artificial corpora for plan recognition. In: UM'05 – Proceedings of the International Conference on User Modeling, Edinburgh, Scotland (2005) 179–188