A Computer-Supported Methodology for Recording and Visualising Visitor Behaviour in Museums

Fabian Bohnert and Ingrid Zukerman

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

Abstract. Developing user modelling and personalisation techniques for museums requires datasets about visitor behaviour in the physical museum space (i. e., visitor pathways). Advancing traditional paper-based tracking techniques, this demonstration presents a computer-supported methodology for (1) observationally recording visitors' time-annotated pathways in physical museums, and (2) post-collection processing and visualising the gathered data. Our methodology was used to collect and post-process a dataset of more than 170 pathways of visitors to Melbourne Museum (Melbourne, Australia). In this paper, we report on our data collection and the lessons learnt from it.

1 Introduction

The *Kubadji* project (`http://www.kubadji.org/`) is developing user modelling and language technologies to support the creation of personalised mobile technology for museums. Key elements under investigation are techniques and models for (1) inferring a visitor's interests and activities from non-intrusive observations of his/her behaviour in the physical museum space, (2) recommending exhibits of interest, and (3) personalising the content delivered for these exhibits via mobile technology. The *GECKO* project focuses on the first of the three objectives, and is investigating user modelling and personalisation techniques which rely on non-intrusive observations of users' behaviour in physical spaces [1]. The development of such techniques for museums requires datasets about visitor behaviour in the physical museum space (i. e., visitor pathways). Although clearly inappropriate for model deployment, datasets that are suitable for developing these techniques may be obtained by observationally tracking museum visitors.

In this demonstration, we present a computer-supported methodology for recording museum visitors' time-annotated pathways, where fully-automated tracking, e. g., [2], is not yet readily available. At the core of our methodology are software tools for on-site tracking of visitors, and for post-collection processing, visualisation and analysis of the gathered data – called GECKO*tracker* and GECKO*visualiser* respectively. The first part of this paper is devoted to describing these tools. The second part of the paper reports on the findings and lessons learnt from data collection conducted at Melbourne Museum (`http://museumvictoria.com.au/melbournemuseum/`) in Melbourne, Australia from April to June 2008, where our methodology was employed to collect and post-process a dataset of more than 170 museum visits.

The paper is organised as follows. Section 2 describes our tools GECKO*tracker* and GECKO*visualiser*, before we discuss the data collection at Melbourne Museum and feedback regarding our tracking methodology (Section 3). In Section 4, we examine areas of application for our methodology, followed by our conclusions in Section 5.

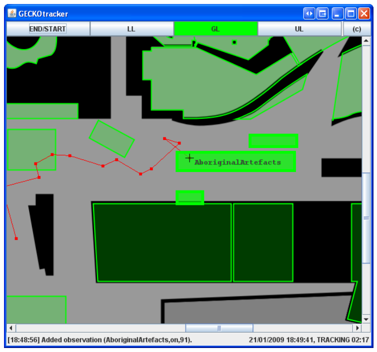

Fig. 1. GECKO*tracker*– Screenshot

Fig. 2. Visitor tracking at Melbourne Museum

2 Software Tools

Traditional manual tracking methodologies used in the museum domain include using printed site maps and a stopwatch to record visitors' pathways and the time spent at various exhibits [3]. However, depending on the required level of detail and frequency of events, such logging techniques can overwhelm a tracker, potentially causing tracking errors. Additionally, these techniques require a substantial transcription effort to digitise the data. This motivated us to to develop a computer-supported methodology for recording museum visitors' time-annotated pathways. Hence, in the framework of the *GECKO* project, we developed two Java-based tools for manual tracking and visualisation of collected datasets, GECKO*tracker* and GECKO*visualiser* respectively.

GECKO*tracker*. GECKO*tracker*'s main component is a clickable interface showing a digitised site map of the physical museum space (Figure 1). Additionally, the interface comprises buttons to start/end a tracking session, and to switch between different maps (e. g., encoding different floor levels of a museum building). The software resides on portable computers carried by human trackers – one tracker follows one museum visitor at a time (Figure 2). When following a visitor, a tracker logs the visitor's position by clicking on the map, while the computer clock delivers the time. A viewing event is registered when the tracker clicks on an exhibit.With every click, the map re-centres at the clicked location, enabling the tracker to focus on the tracking task. Placing the mouse cursor over an exhibit reveals its textual identifier. Visitor pathways and viewing events are written to text files in human-readable format. To encode the site maps, we use an established standard for representing vector graphics: the *Scalable Vector Graphics (SVG)* file format. In this encoding, clickable exhibits are represented by shapes, and the exhibits' colours and identifiers (as displayed by GECKO*tracker*) are determined by the shapes' attributes. Third-party software, i. e., the open-source vector graphics editor Inkscape (`http://www.inkscape.org/`), was used to create the site maps.

GECKO*visualiser*. GECKO*visualiser* complements GECKO*tracker*, and is used for post-collection visualisation and analysis of the gathered data (Figure 3). It supports different views of the data in two linked formats: visualisation on the site map (left portion of the application window) and textual log (right portion of the application window). Different maps (e. g., for different floor levels of the physical museum) are accessed by selecting one of the tabs shown above the site map (in our case, 'LL', 'GL' and 'UL').

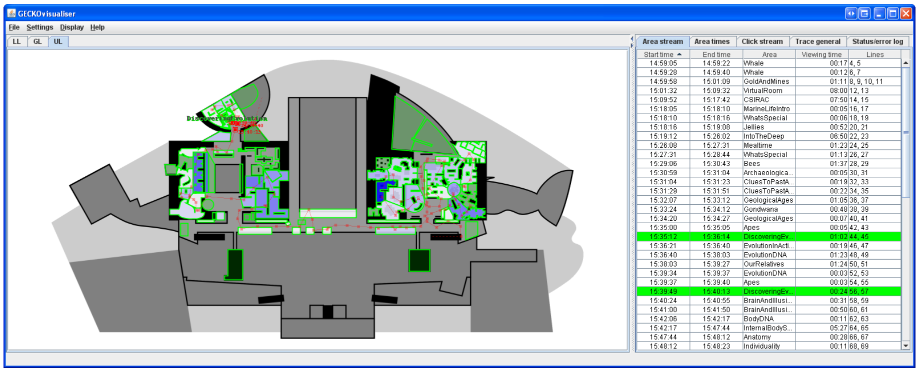

Fig. 3. GECKO*visualiser*– Screenshot

When opening a visit trajectory, GECKO*visualiser* processes the information to extract a visitor's pathway and the viewing durations associated with visited exhibit areas. A visual representation of this information is depicted on the site map on the left-hand side of the application window. The pathway is shown as a red line (clicks are shown as red squares), with floor level transitions indicated by dashed lines. The distribution of viewing times is visualised by colouring the exhibit areas. The level of blueness indicates the amount of time spent by a visitor – the darker the blue, the higher the associated (total) viewing duration. Unvisited areas are transparent. This visual representation enables an analyst to quickly gain an understanding of a visitor's viewing behaviour.

The tabs on the right-hand side of the application window provide access to various textual views of the visit trajectory. The screenshot in Figure 3 shows the 'Area stream' tab. Each row of this tab's table represents one viewing event, chronologically ordered for a visit. The tab 'Area times' lists the viewing times by exhibit area, accumulated over a visit (i. e., multiple viewing events at one area are merged into one viewing duration). The 'Click stream' tab provides a table view of the recorded click events (in visit order), and the 'Trace general' tab gives summary statistics of a visit trajectory. Finally, the tab 'Status/error log' is used for information such as file parsing errors.

For GECKO*visualiser*, the site map view and textual log tables are linked such that mouse movements over the site map highlight an exhibit area's viewing events in the table (colouring rows in green). In addition, click events next to the mouse cursor are highlighted by displaying their time stamps (and fading the other clicks). Mouse movements over a text table highlight the viewing events of the hovered exhibit area in the table (in green). Additionally, for viewing event tables (i. e., tabs 'Area stream' and 'Area times'), the corresponding exhibit area is highlighted on the site map. For the 'Click stream' tab, mouse movements over the table highlight the corresponding click event on the site map. Linking the site map and textual logs enables a data analyst to easily identify certain stages of a visit, and to replay the progression of a visit.

3 Case Study Melbourne Museum

Using GECKO*tracker* and GECKO*visualiser*, we recorded and post-processed the pathways of more than 170 visitors to Melbourne Museum. This section reports on our experiences and on the observations made during this case study.

Site map creation. The exhibited collection of Melbourne Museum comprises a few thousand exhibits. As we did not require data recorded at this detailed level, we asked museum staff to help us group the exhibits into exhibit areas that are semantically coherent and spatially confined (often indicated by named themes in the museum). This yielded 126 exhibit areas. We then used paper-based floor plans of the physical museum space to create a digital SVG file-based representation.

Pilot phase. Prior to the main data collection, we piloted our methodology with the help of two trackers (we tracked 19 visitors). This enabled the early identification and correction of some minor issues with GECKO*tracker* and the site maps before the main data collection. More importantly, the pilot phase yielded valuable insights regarding areas of the museum where visitors are difficult to follow. We also gained insights about procedures to resolve problems with laptop computers running low in battery after several hours of tracking. For example, based on our experiences during the pilot phase, we decided to exchange laptops after tracking only one visitor (irrespective of how long their visit turned out to be), and instructed our trackers to take battery-saving measures.

Briefing session. Before the data collection, we briefed our trackers on the usage of GECKO*tracker*, the layout of the museum, and its digital representation on the site map. Additionally, we defined a viewing event as a participant looking at or engaging with an exhibit (either when walking or standing still), and/or a participant taking a photograph of an exhibit. We showed footage of an actual museum visit to minimise ambiguities with respect to this definition. The briefing included a walk-through of the museum, during which our trackers took turns in test-tracking their colleagues. Although time-intensive, the walk-through turned out to be the most valuable part of the briefing, as it allowed the trackers to experience the methodology under supervision.

Observational study. The main data collection took place between April and June 2008. We only tracked first-time adult visitors travelling on their own, to ensure that neither prior knowledge about the museum nor other visitors' interests influenced their decisions about which exhibits to view. We employed 14 trackers in total, comprising university students and museum staff. So as not to disturb the general audience unnecessarily, we scheduled a maximum of three trackers per day. With respect to technical equipment, we used four small laptop computers with mouse pad (EEE-PCs) and two tablet computers with stylus pen. This allowed us to track six visitors in a day without recharging the equipment.

Post-collection data processing. After the data collection, the visitor pathways were post-processed using GECKO*visualiser* to visualise the recorded visit trajectories. For instance, we corrected mistakes that our trackers made while tracking, such as accidental clicks on exhibit areas. This was possible as we had asked our trackers to manually record such mistakes. Additionally, we removed mis-clicks reflecting tracking events that could not possibly have occurred, e. g., visitor transitions from one end of the museum to the other and back within a few seconds, or transitions outside museum walls and back. We also removed incomplete visitor pathways, e. g., due to a laptop computer running out of battery, or a visitor leaving unexpectedly.

Dataset. The resulting dataset comprises 158 complete visitor pathways in the form of time-annotated sequences of visited exhibit areas, with a total visit length of 291:22:37

hours, and a total viewing time of 240:00:28 hours. The dataset also contains demographic information about the visitors, obtained by means of post-visit interviews conducted by our trackers. In total, we obtained 8327 viewing durations at the 126 exhibit areas, yielding an average of 52.7 exhibit areas per visitor (41.8% of the exhibit areas). Hence, on average, 58.2% of the exhibit areas were not viewed by a visitor.

Feedback analysis. After completing the observational study, we asked our trackers and other museum staff to give their feedback regarding our tracking methodology. Below we summarise their comments, and discuss feedback we received from participating visitors.

- **Feedback regarding our tracking methodology.** The feedback we received from trackers and museum staff regarding GECKO*tracker* indicates that they value our software. The trackers particularly like the software's ease of operation, stating that it is "intuitive" and "easy to use". A valuable suggestion to improve GECKO*tracker* is to implement an undo function, such that mis-clicks can be revised on-site. The trackers also suggested showing the recorded pathway to enable personal control of one's tracking and to ease orientation (we had decided not to display the pathway during the data collection). The trackers who had used a paper-based methodology before described our methodology as more efficient and more accurate. Indeed, our trackers were generally confident that they tracked accurately. They also said that the site maps encode sufficient information for the trackers to correctly identify exhibit areas, and most trackers thought the site maps are intuitive. On average, the trackers stated that they had familiarised themselves with GECKO*tracker* within two tracking sessions.
- **Feedback regarding participants.** Our trackers indicated that they were required to approach between 3-5 visitors before a visitor agreed to being tracked. They also reported that keeping a distance of 5-10 meters from participants is most practical when tracking, with the actual distance depending on the spatial arrangement of the exhibits. This distance enables accurate tracking, while allowing participants to move freely. Most trackers stated that the visitors they followed were aware of their presence only initially. They also agreed that in general, their presence did not influence the participants' visiting behaviour. These comments concur with the participants' feedback on this matter.
- **Feedback from participants.** Feedback from participants regarding our tracking methodology indicates that most visitors did not feel disturbed by a tracker following them through the museum. In fact, some participants stated that quite early into their visit, they forgot that they were being tracked (despite being approached at the start of the visit to obtain their approval).

4 Discussion

We used the methodology presented in this paper to collect and post-process a dataset of pathways of visitors to Melbourne Museum. This dataset will be used in the *Kubadji* and *GECKO* projects as a basis for developing techniques for modelling museum visitors from non-intrusive observations of their movements through the space, and for delivering personalised recommendations and content about exhibits to visit.

Input we received in discussions with museum staff points to additional application areas for our software that are directly relevant to museums and art galleries. Museum staff, such as curators and exhibition designers, can use our software to understand visitor flow patterns and entry points into exhibitions. The tools could also be employed to identify attractive exhibits on the one hand, and exhibits that receive little visitor traffic on the other hand. Our software is currently limited to recording and visualising temporal data (i. e., time-annotated visitor pathways). Although viewing time is correlated with interest [4], it might not be the best measure for inferring a visitor's intention, appreciation and learning. For observational studies where such measurements are necessary, our methodology should be combined with techniques focusing on such measures, such as interviews or questionnaires.

5 Conclusions and Future Work

In this paper, we offered a computer-supported observational tracking methodology for recording time-annotated visitor pathways in physical museums. We presented two software tools: GECKO*tracker* for on-site tracking of visitors, and GECKO*visualiser* for post-collection processing, visualisation and analysis of the gathered data. We also reported on the findings from data collection conducted at Melbourne Museum, where our methodology was used to collect and post-process more than 170 museum visits. Overall, our methodology was received positively by museum staff, trackers and participating visitors.

In the future, we propose to extend our software by incorporating undo and editing facilities, so that a tracker can undo mis-clicks (when using GECKO*tracker*), and a data analyst can edit a pathway directly on the site map or in a data table (when using GECKO*visualiser*). We also plan to endow GECKO*visualiser* with direct capabilities for visualisation of summary statistics derived from the dataset, such as showing visitor pathway clusters and flow patterns on the site map.

Acknowledgements. This research was supported in part by grant DP0770931 from the Australian Research Council. The authors thank Shlomo Berkovsky, Timothy Baldwin, Liz Sonenberg, Karl Grieser, and Sandrine Balbo for their valuable comments on this research. Thanks also go to Carolyn Meehan and her team from Museum Victoria for fruitful discussions and their support of this work. Finally, we thank Jessica Mills from JPIX Photography for her photographic assistance.

References

1. 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) (2008) 195–202
2. Hightower, J., Borriello, G.: Location systems for ubiquitous computing. IEEE Computer **34**(8) (2001) 57–66
3. Diamond, J.: Practical Evaluation Guide – Tools for Museums and Other Informal Educational Settings. AltaMira Press (1999)
4. Parsons, J., Ralph, P., Gallager, K.: Using viewing time to infer user preference in recommender systems. In: Proc. of the AAAI Workshop on Semantic Web Personalization (SWP-04), in conjunction with AAAI-04. (2004) 52–64