

# Visualizing, Clustering, and Predicting the Behavior of Museum Visitors<sup>1</sup>

Claudio Martella<sup>2a</sup>, Armando Miraglia<sup>a</sup>, Jeana Frost<sup>a</sup>, Marco Cattani<sup>b</sup>, Maarten van Steen<sup>c</sup>

<sup>a</sup>*VU University Amsterdam, Amsterdam, The Netherlands*

<sup>b</sup>*Delft University of Technology, Delft, The Netherlands*

<sup>c</sup>*University of Twente, Enschede, The Netherlands*

---

## Abstract

Fine-arts museums design exhibitions to educate, inform and entertain visitors. Existing work leverages technology to engage, guide and interact with the visitors, neglecting the need of museum staff to understand the response of the visitors. Surveys and expensive observational studies are currently the only available data source to evaluate visitor behavior, with limits of scale and bias. In this paper, we explore the use of data provided by low-cost mobile and fixed proximity sensors to understand the behavior of museum visitors. We present visualizations of visitor behavior, and apply both clustering and prediction techniques to the collected data to show that group behavior can be identified and leveraged to support the work of museum staff.

*Keywords:* proximity sensing, mobile sensors, museum visitor analysis, hierarchical clustering, visualization, recommendation, prediction, matrix factorization.

---

## 1. Introduction

Museums provide the public with a variety of services including learning, entertainment and social interaction, and they are increasing in popularity [2]. Yet, there are few accessible methods for museum staff to gather actionable information about the visitor experience, learn from past experiences, evaluate shows for funders, and improve future work. In the museum community at large, including science, technology and historical museums, professionals advocate the use of evaluation techniques throughout the design process [3, 2]. Within the art world, evaluations can help optimize the layout of ongoing shows, improve traveling exhibitions, inform future design choices [4], and strengthen requests for funding [5]. However, conducting evaluation studies using traditional research methods is expensive, often forcing management to cancel such studies due to budget constraints [6]. Low-cost sensor networks designed to meet the needs of fine-arts museums could simplify evaluation and ensure its integration within routine design processes.

---

<sup>1</sup>Note that this paper is an extended version of the work entitled “Leveraging Proximity Sensing to Mine the Behavior of Museum Visitors” [1].

<sup>2</sup>Corresponding author. Email: c.martella@vu.nl. Telephone: +31205987748. Postal address: Dept. Computer Science, De Boelelaan 1081A, 1081HV Amsterdam, The Netherlands

### *1.1. Motivation*

We conducted a case study at the CoBrA Museum of Modern Art (CoBrA, for brevity)<sup>3</sup>, where over 180 visitors volunteered to participate in the experiment by wearing one of our proximity sensors. Before the experiment, we conducted a series of semi-structured interviews with the curator, the artistic director, and the head of business development to better understand their processes and needs. In addition to the staff of CoBrA, we also interviewed an independent exhibition designer who works regularly at CoBrA, and an administrator from the van Gogh museum who commissioned a large-scale observational study of that museum and has applied the findings to the re-organization of the permanent exhibition. Interview notes were coded and themes were identified. The analysis revealed three themes.

**Inform design of future shows.** Currently, museum staff designs exhibitions based on theories and experience about what pieces are popular and how visitors will move through the gallery spaces. For example, the artistic director at CoBrA believes that wall placement may impact visitors movement through a show more than the displayed artworks, but expects that visitors also deviate from the ordering that curators envision. Both creative staff and those in charge of business development at CoBrA expressed a need for objective information to test such assertions and improve intuitive understanding about the gallery. The artistic director stated “If we learn more about that flow, [it] will have an impact on how we design, how we display the works of art.” Within a larger institution, the van Gogh Museum, parts of the galleries were redesigned based on the results of an observational study. The study had identified that the previous placement of certain famous artworks were causing congestion in parts of the gallery space. Moreover, less popular pieces were ignored when positioned close to more popular ones. Finally, the position of some divider walls confused some visitors about what path they were expected to follow in the museum, resulting in some parts of the exhibition being skipped completely. As we show in this paper, similar observations — in fact corresponding visualizations — can be obtained with our sensor-based system, at a much lower cost.

**Report to funders and potential partners.** For the last ten years, the head of business development at the CoBrA has received requests from funders to demonstrate the value of shows. He stressed the need for “quantifiable information” about the experience of visitors. Museums track some metrics: ticket sales, museum shop sales and some periodic surveys, but these do not indicate how visitors interact with different artworks or level of satisfaction once inside of the exhibition. Specifically, the head of business development would like to know how much time visitors spend viewing particular pieces and how they move through the show. He wants to share these metrics externally in funding requests and inform the design of marketing campaigns. Finally, he found that, for example, identifying most popular artworks during the first weeks of a new exhibition could help the design of the marketing campaign, perhaps based on those pieces.

---

<sup>3</sup><http://www.cobra-museum.nl/>

**Increase participation.** In the museum world, the staff of CoBrA reports a concern that visitors can remain passive and fall into the “museum shuffle”, that is, visitors walk from piece to piece mindlessly, spending a few seconds at each piece before moving on. Museum staff wants to find ways to know when this occurs such that they can intervene and change behavior. The ultimate goal is to avoid this passive mode and enhance engagement with the content to increase the exhibitions “impact” beyond the visit itself. The interviewees’ main objective for a museum visit is for visitors to learn something. To do so, they want to “get in touch” with visitors using different methods of understanding the visitor experience. The artistic director acknowledged current methods were not providing the right type of information: “we know them in statistics [...] but we do not know about the quality of their experience”.

The results of these interviews were used for the design of our data collection and data analyses.

## *1.2. Contributions*

In this paper, we propose that measuring visitor behavior with inexpensive sensors, and visualizing and mining the resulting data, can aid the museum staff in the design and evaluation of fine-art exhibitions. We argue that measuring how visitors distribute their time across artworks is essential to our goal of capturing their behavior. In particular, we refer to which pieces visitors observe during their visit, for how long, and in what sequence. To this end, we employ an infrastructure comprising off-the-shelf proximity sensors, which we use to measure which artworks visitors face over time [1]. We present a number of visualizations that display the behavior of the visitors, from how they distributed their time across exhibits and rooms, to the most frequent paths followed. These visualizations show that certain rooms were more popular than others, and that in those less popular rooms visitors followed a less structured path, perhaps due to confusion. Moreover, we show how patterns of group behavior can be identified by the application of hierarchical clustering to the data. For example, we show that 10% of the volunteers visited the exhibition in the wrong order, starting from the end. Furthermore, we demonstrate the power of the identified patterns by leveraging them to predict visitor behavior by means of matrix factorization techniques. We do this by applying recommender-systems techniques traditionally used to predict user-item ratings.

The remainder of this paper is organized as follows. First, we provide an overview of the experiment and data-filtering pipeline. More details about the design, implementation and evaluation of the filtering pipeline is presented in previous work [1]. Then, we present three types of visualizations of visitor behavior. We continue by presenting the results of the clustering and prediction of the behavioral data. Finally, we conclude with a discussion about a small focus group interview we conducted with the CoBrA museum staff regarding our results, and the opportunities for future work.

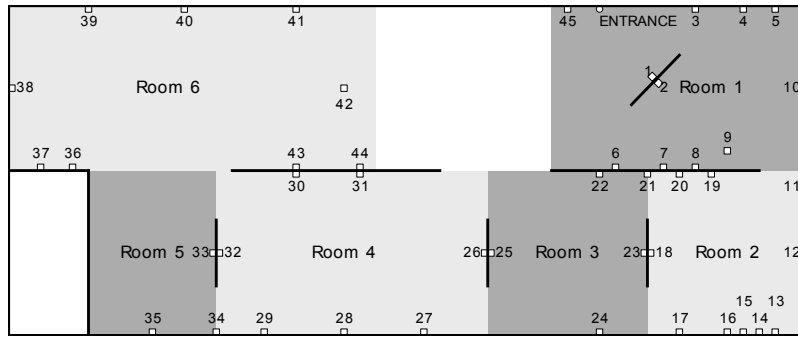


Figure 1: Planimetry of the “The Hidden Picture” exhibition held at the Cobra Museum of Modern Art. The Figure shows the position of the entrance, the rooms and the artworks. It also defines ID numbers for the rooms and the artworks.

## 2. Overview

We conducted a 5-day experiment spread across 2 weekends at CoBrA. Our data collection focused on the temporary exhibition entitled “The Hidden Picture”, a curated sample of the corporate collection of the ING Bank. The exhibition was displayed in the dedicated open space on the top floor of the museum. The space is configurable, and divider walls were used to separate the space into 6 “open rooms” dedicated to different themes. The overall space was about 100 meters long and 25 meters wide, with a ceiling reaching about 5 meters, while divider walls were some 3.5 meters high.

Rooms 1 and 2 focused on figurative art, rooms 3 and 4 mostly on abstract art, room 6 on pieces inspired by nature, for a total of 60 pieces. The pieces varied in size, style and medium, including photos, paintings, sculptures, videos, and an installation with a cage hosting a live chameleon. None of the pieces were highly famous, and were hence appealing the visitors based on immediate reaction rather than on prior knowledge. Of the 60 pieces, we instrumented 45 exhibits with our sensing infrastructure.

### 2.1. Data collection architecture

We designed a system based on inexpensive radio-based proximity sensors. Our sensing solution is compliant to the Zigbee standard and it can be implemented for example through Bluetooth low energy (BLE) beaconing, available in modern smartphones. To give us freedom to investigate our solution, instead we deployed ad-hoc devices running a duty-cycled MAC protocol [7] that allows us to run our system for weeks with a single battery charge.

The sensing infrastructure comprises *mobile* devices and *anchor* points (or simply anchors). Mobile devices are sensor nodes worn by the visitors. They are attached to a lanyard worn around a visitor’s neck. Due to the shielding effect of the visitor’s body, the radio communication range is steered to the front with a controlled angle of around 90 degrees and some 2-3 meters of distance. Anchors are sensor nodes positioned at the base of each exhibit. We installed anchors inside of enclosure aluminum boxes designed to shape the

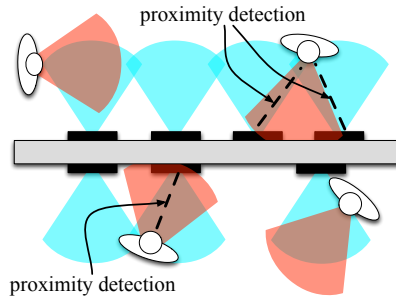


Figure 2: Detections between mobile devices and anchor points installed at exhibits. A detection occurs when an anchor point is in the range of a mobile point. We steer the detection range for face-to-face proximity detection to measure when visitors face an exhibit.

communication range to approximately 60 degrees and 2-3 meters of distance. With this setup, mobile devices and anchors can communicate only when the visitor is facing an exhibit.

Every second, anchors transmit through the radio a unique anchor identifier (*AID*) that is received and timestamped by mobile devices within range. We consider the reception of an *AID* by a mobile device a *proximity detection*. Note that our sensors do not measure radio signal strength (i.e., *RSSI*). While it does not enable us to measure distance between points, it allows a cheaper and more energy-efficient solution (moreover, computing distance with *RSSI* is still a challenging problem as it provides noisy information that largely varies depending on environment conditions). Every second, mobile devices transmit the list of detections received during the previous second together with their unique mobile-point identifier (*MID*) at a longer range of approximately 100 meters, which are received by one or more *sinks*.

*Sinks* are computers that receive mobile-device transmissions through the same type of sensor node used for anchors and mobile devices, and store the timestamped lists of detections in a central repository. *Sinks* are installed in various areas of the exhibition space to ensure full coverage and some degree of overlap. Note that due to the overlap of the areas covered by the *sinks*, mobile devices transmit their messages together with a randomly generated number that we use together with timestamps to remove duplicate detections from the database.

When a mobile device is handed out anonymously to a visitor, the visitor is assigned a unique user identifier (*UID*) that is associated to the corresponding *MID*. Each visitor check-in and check-out times are stored together with the *UID-MID* mapping. Our raw-data database comprises this mapping and the list of timestamped detections collected by *sinks*.

## 2.2. Data-filtering pipeline

Raw sensor data was noisy and incomplete, with many missing and wrong detections. A number of factors can produce noise in the data including density of people in the space, humidity, air temperature, objects and surfaces nearby [8] potentially resulting in detections being lost or wrongly added. Moreover,

sometimes mobile points can detect anchor points that are further than the expected range, for example on the other side of a wall, or missing detections for a substantial amount of time. Using raw proximity detections without filtering would produce inaccurate measurements, for instance assigning face-to-face proximity to wrong artworks or missing correct face-to-face proximity completely.

To overcome these effects, we devised a filtering pipeline that comprises a so-called particle filter [9] and a clustering technique used to filter bursty and noisy data [8]. While the details of the pipeline and its evaluation are presented in previous work [1], here we provide an overview of its main functionalities.

Our goal, which we define as *positioning*, differs from the known challenge of localization. Whereas the goal of localization is to compute the absolute position of an individual in space, the goal of positioning is to map the position of an individual *at* one (or no) point-of-interest, in our case *in front* of an artwork. In this sense, we consider face-to-face proximity as an instance of relative positioning. We utilize and modify particle filters to fit our problem of positioning. Through the particle filter, for each individual we obtain a series of pairs  $(id, t)$  that tell us that the individual was positioned facing artwork  $id$  at time  $t$ .

Particle filters have been successfully applied to the problem of indoor localization with noisy sensors to estimate absolute positions of individuals in rooms. Usually, radio-based communication is used to measure the distance between a mobile sensor, for example worn by an individual, and a number of fixed landmark sensors positioned in known locations, for example at the four edges of a room. Distance from landmarks can be computed by measuring packet round-trip time or radio-signal strength. A so-called particle is characterized by a 2D or 3D coordinate and sometimes an angle of movement. Each particle represents an estimate of the current position and direction of movement of the individual. Initially, a number of particles (in our case 1000) are assigned coordinates, i.e., locations in the room, chosen uniformly at random. Periodically, for example every second, the following three steps are executed based on sensor measurements and particles data: (i) estimation, (ii) re-sampling, and (iii) movement.

In a first step, the likelihood of each particle prediction is assessed by computing the error between the particles' distance from the landmarks and the individual's distance from the landmarks as measured by the mobile sensor. In a second step particles are re-sampled based on their likelihood, positioning particles with lower likelihood close to particles with higher likelihood. After the re-sampling phase, in a third step particles are moved either according to the angle of movement, if present, or at random at a distance reachable through the assumed speed of movement of an individual (e.g., walking speed, 1m/s in our case). Impossible moves, for example through walls, are rejected. When a sensor reading could not be obtained, particles are just moved according to the movement step.

At each time, the position of an individual can be predicted by computing the weighted average of all the particles positions (where the weights are the particles likelihoods). We can also compute the prediction confidence as the deviation of individual particles predictions from the computed average. When sensor readings are missing for a number of seconds, either due to errors/noise or due to the individual not being in

front of any artwork, particles quickly spread across the space due to the third step where particles move at random.

Compared to localization, the main difference in applying particle filters to positioning in our setup is that we use *many* landmarks that *do not* provide measurements of distance and have a defined cone-shaped *limited range* of detection (see Figure 2). Note that while many artworks are involved, an individual is expected to be within range of a limited number of anchor points at each time due to their limited detection range. The core of our adaptation lies in the way we compute particle likelihood at each detection of a mobile point by one or multiple anchor points. Intuitively, we define the likelihood of a particle prediction to be higher the more the particle is positioned frontal to the detected anchor point and close to a distance of 1 meter. Precisely, for each anchor point we define a likelihood function that is a function of the distance of the particle from the anchor point and its angle with respect to the perpendicular originating from the anchor point, e.g., a painting canvas. For this function we define a multivariate Gaussian kernel that defines maximum likelihood at 1 meter of distance from the artwork, and at 0 degrees angle (right in front) from the artwork. The likelihood is defined as 0 for distances beyond 3 meters and angles larger than plus or minus 30 degrees. When a mobile point detects multiple anchor points at the same time, we define the likelihood of a particle as the sum of the likelihood of each anchor point.

Nonetheless, we can sometimes incur in situations where an individual is positioned only momentarily, for example for one second, at an artwork close-by or at no artwork at all (while the individual was positioned correctly for the previous seconds). This can happen due to wrong or missed detections respectively. Hence, the resulting positioning can present incomplete bursts, where an individual is positioned correctly in front of an artwork for a certain amount of time but with few gaps of few seconds in between, and spurious isolated wrong positionings, where an individual is suddenly positioned only for a few seconds at a different wrong artwork.

To overcome this problem, we applied a *density-based filtering* technique [8] that ignores the spurious positionings and reconstructs the missing ones. Note that instead of applying the density-based filter to the raw detections, here we applied the filter to the positionings. In fact, applying the filter to the raw detections would result in the reconstruction of some detections of nearby anchor points, making the work of the particle filter more difficult and less accurate. The end result of the pipeline is that for each moment in time, precisely for each second of the visit, we can position each participating individual either in front of an artwork, or in front of none if she was far from all artworks.

Finally, we extract through a *majority voting* filter the sequence of AIDs at which the visitor was positioned. The filter scans visitor-positioning data to detect transitions from exhibit to exhibit, and disambiguating situations where a visitor may appear facing two exhibits at the same time. During this step we also reconstruct the *path* followed by the visitor, that is the sequence of stops at exhibits followed by the visitor.

Once raw data has been processed by our pipeline, we obtain for each visitor a vector  $r_v$  of  $N$  elements, each representing the number of seconds spent at each of the  $N$  anchors/exhibits, and a sequence  $s_v$  of AIDs to represent the path followed to visit the exhibits. If an exhibit was never visited by the volunteer, the corresponding element of  $r_v$  will contain a value of 0, and the corresponding AID will be missing in  $s_v$ .

### 3. Background and related work

Existing research on museums and exhibitions focuses on the use of technology to support interactivity, for instance between visitors and with exhibitions, rather than supporting museum staff in curatorial work. The application of technology within museums can be broad [10, 11]. One topic of interest is collaboration and co-design [12, 13], in particular with attention to interaction [14, 15] and how different media can influence it [16, 17]. A second body of work focuses on how technology [18], and electronic guides [19, 20] can support learning. Finally, another topic of research is the impact of visiting in groups, as a family [21] or with children [22]. However, few projects have investigated the use of technology to understand how visitors behave in the museum [23]. How people negotiate their way through museums and galleries can have considerable implications for how they relate to and interpret exhibition content [24]. Observational studies have been conducted [25], but they have limitations concerning scale.

Recent advances in sensor-based systems to study social behavior, like those studied in Social Signal Processing [26], Participatory Sensing [27], and Opportunistic User Context Recognition [28] show how automated sensing, and in particular mobile sensing, can be an opportunity to pass beyond these limitations. Yet, few sensor-based systems have been used to study the behavior of visitors to fine-art exhibitions, in relation to their movements. Early attempts made use of indoor localization systems based on Bluetooth data collected from mobile phones, to trace the movement of visitors between rooms [29]. This data can be used, for example, to support multimedia guides [30, 31], but it captures only which room (or part of it) an individual is visiting. More recently, a device comprising positioning and physiological sensors has been used, complemented with entrance and exit surveys, to study the cognitive reaction and social behavior of a number of individuals in an exhibition. In that case, researchers used the data to classify experience into three categories: “the contemplative”, “the enthusiast”, and “the social experience” [32].

A similar device to measure position and spatial orientation (i.e., through a compass) of individuals has been used to study the behavior of visitor pairs in a museum. The study presents a system to classify pairs early in the visit into one of six classes, to provide socially-aware services to the pairs, for example to increase their engagement with the exhibition [33]. Also in this case, the granularity of their data does not allow them to discern which pieces visitors actually face. Similarly, data coming from a multimedia museum guide was used to predict visitor profile type among “ant”, “fish”, “butterfly”, and “grasshopper”, in an attempt to personalize visitors’ guide and avoid information overload [34]. Finally, similar work has

also been done in the context of virtual environments to trace user behavior [35], where full localization information is used to visualize movements of users in a virtual museum. More recently, the visualization of metrics such as popularity, attraction, holding power and flows has been explored [36, 37]. These works focus on visualization of the data without investigating deeper into data-mining techniques and data-filtering, and focus on simple exhibit-to-exhibit transitions without tackling longer sequences as paths (as in this paper).

A simple way of monitoring a crowd consists in precisely tracking the absolute position of each individual. While this technique is a feasible option in outdoor scenarios, where the GPS system can be exploited, for indoor conditions accurate localization is still an open problem. Among the wide literature of radio-based localization techniques, only few [38, 39, 40] are accurate enough to be employed in museum scenarios. Unfortunately, none of these techniques perform consistently throughout the museum areas. It is known, in fact, that the localization error increases significantly at the edges of rooms and in hallways [41], conditions that are often present in museum topologies. For our tracking application, this phenomenon can be even more problematic, since even small estimation errors could lead to visitors being associated to the wrong exhibit, positioned in the wrong room or placed on the wrong floor.

A different approach for quantifying visitor engagement is to measure activity on social media and museum Web pages. These approaches apply a “blackbox” approach to the museum experience, and consider typical social media and Web 2.0 actions such as likes, shares, comments, etc. as proxies to engagement and satisfaction measurement [42], or to implement models similar to loyalty and affiliate programs used by other industries [43]. While these methods enable the monitoring of the long-term engagement of visitors, they capture coarse information that does not allow insights into the outcome of more specific curatorial decisions that involve, for example, exhibition layout.

#### **4. Visualizing visitor behavior**

In this section we present a number of visualizations that we have developed to show insights of visitor (group) behavior present in the proximity data. While our work is complementary to the existing related work regarding attraction power [36], it introduces a richer understanding and presentation of paths, moving from existing work that pictures only exhibit-to-exhibit transitions to full sequences of exhibits and their frequency (i.e., paths).

##### *4.1. Artworks and rooms popularity*

Based on the findings of our assessment study, we used the time spent in front of an art piece as a proxy for visitor appreciation. We computed the aggregate time each piece was viewed and presented this total in a histogram organized by room and medium. The color of the bar indicates the medium whether it is a painting, a video, or a sculpture. The horizontal line in correspondence with each room shows the average

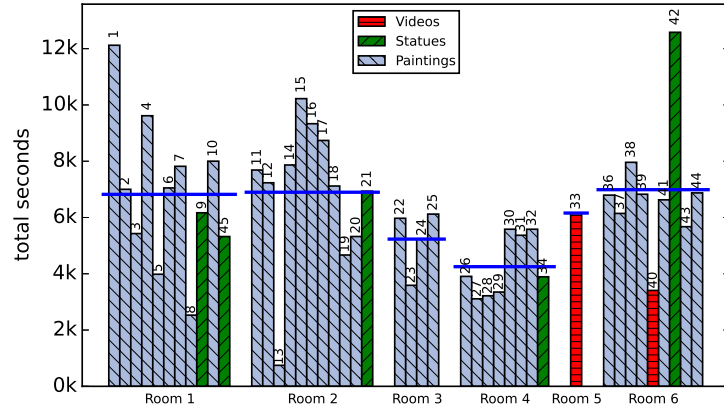


Figure 3: Histogram showing the popularity of each artwork organized per room. Each bar represents the total number of seconds all the visitors have spent in front of each artwork. We also show the average time spent in each room, represented by the blue horizontal bar.

value across the bars (see Figure 3). The visualization shows that artwork 42 (i.e., a piece that included a live chameleon) was most popular followed by 1, the artwork placed right at the entrance of the exhibition (it was also used as the poster for the exhibition around the city). Medium did not seem to predict popularity (for instance sculptures did not, on average, receive more or less attention than paintings). It is also interesting to note that the video documentary labeled as 33 received much more attention than the artistic loop video, of comparable length, labeled as 40.

Looking at room-level popularity, one can notice two things. First, most of the time was spent in the first two rooms and in the last one, though visitors also allocated comparable time to room 5 due to the video documentary. Rooms 3 and 4 however received substantially less attention. These rooms were later in the exhibition and contained more abstract works. Second, the variance of time spent in these two rooms is relatively small, suggesting that no piece in these rooms attracted a great deal of attention. Rather, visitors seemed to walk sequentially through these rooms, without stopping at any piece in particular (in an example of the “museum shuffle”). This hypothesis is confirmed by the analytics we present next.

#### 4.2. Visitors common movement paths

Once we know when a visitor is facing a particular piece and for how long, we can construct their path within the museum. Given we do not have absolute positioning information, we consider a path as the sequence of artworks that a person has faced during her visit, ordered by time. For our goal, this representation summarizes the path taken by the visitor without the noise contained in long sequences of absolute coordinates, and it is well suitable, as we will show, for data-mining tasks. When we project this sequence onto the actual layout of the space to visualize the path, we choose the shortest trajectory that connects the artworks and does not pass through walls. Hence, in our visualizations, a displayed path

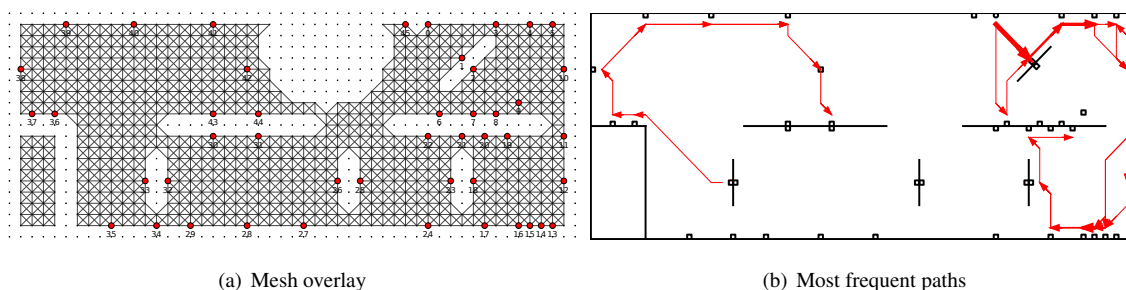


Figure 4: (a) The mesh used as layout to compute paths between paintings. The planimetry is split in cells, each one mapping 1 square meter of the layout space. Red dots labeled with a number represent artworks. Walls are represented as disconnected dots in the mesh. The circular gap in the upper-central part of the mesh is due to a non-walkable atrium. (b) Most frequent paths followed by the participants. The thickness of each line is proportional to the frequency of the represented subpaths.

represents a logical trajectory in space. Also, for readability, we display paths by computing them over a mesh, constructed by the discretization of the layout space with 1 square meters cells. Figure 4(a) shows the computed mesh. In this way, we minimize clutter and present information in an orderly fashion.

To visualize routes, we constructed a visualization of the most popular paths. Based on interview data with a member of the van Gogh Museum staff, visualizations of all visitors together prove confusing. Given enough visitors, the image is easily filled with lines and all informational value is lost. Moreover, group behavior appears only indirectly, emerging from the overlap of the individual paths. Instead, we decided to take a direct approach that would minimize the effort of the user to recognize group behavior. The result is a visualization of the layout with a number of lines that represent common paths, whose thickness is proportional to their frequencies across all visitors. We proceeded as follows.

First, we computed which paths and subpaths were common among all the individuals. In particular, we computed the longest common subsequences (LCS) between all paths, and their frequencies across all paths. For example, suppose that two visitors at some point have faced artworks 1, 2, and 3, in this order. In our statistics, we only considered subpath 1-2-3, and not subpaths 1-2 and 2-3, as they are included in LCS 1-2-3. Note that two visitors can have multiple LCSs of different lengths, for example accounting for common subpaths in different rooms.

Second, we merged the 10 most frequent LCSs into a single visualization. Since LCSs can overlap, computing the thickness of each LCS requires to account for the frequencies of each overlapping subsequence. For example, consider two LCSs, namely 1-2-3 and 2-3-4 both with associated frequency 1. Merging the aforementioned LCSs requires to increase the frequency of the overlapping subpath 2-3. In this example, the result is a final path 1-2-3-4 where subpaths 1-2 and 3-4 have an associated frequency of 1 while 2-3 has an associated frequency of 2. Hence, the line would be thicker in correspondence to the line connecting 2 to 3.

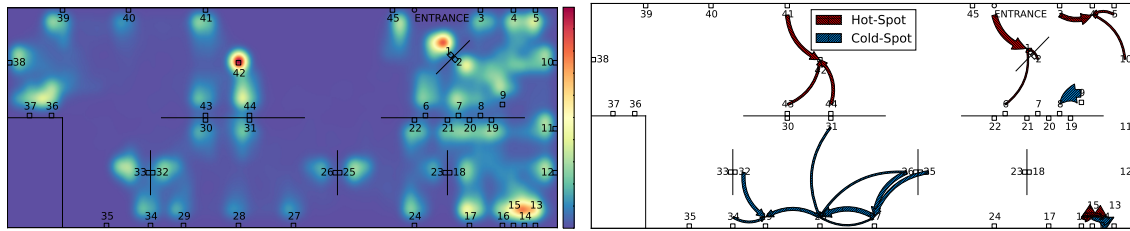
Figure 4(b) shows the resulting visualization. One can notice a number of aspects. First, visitors tended

to turn left after artwork 1, and approached it as first piece or right after artwork 6. While of scenic effect, the initial wall might have confused the visitors who did not know what was “expected” from them. Second, visitors tended to walk along the external edges of the space. In an independent observational study, this same pattern was observed in the Van Gogh Museum. Third, some artworks are not reached by the line. It is possible that some of those paintings were often skipped by the visitors, perhaps because recognized and ignored while walking the dominant external path, hence appearing less frequently in the paths. Another more likely explanation is that due to their position outside of the dominant line, they were reached by the visitors at different points of the path, hence generating less frequent LCSs. We verified this claim by visualizing the common paths generated by choosing the 20 most frequent LCSs.

It is interesting to notice how room 3 and 4 do not include the line at all. As mentioned earlier, according to our measurements these rooms contained less popular paintings, and were sometimes completely skipped. Also, the layout of the room requires a decision by the visitor in front of each divider wall, again whether to visit the right-side or the left-side of the room first. Similar behavior was observed in another study [44], when museum visitors entered spaces without an obvious point of orientation their paths scattered in different directions. It is likely that visitors took very different decisions, if they decided to visit the room, and visited the pieces in very different order, perhaps skipping some. Again, this would generate less frequent LCSs that would not make it to the top 10.

#### *4.3. Common positions of visitors*

In this study we track proximity to art pieces rather than absolute location. Although we lack this information, we can leverage the coordinates provided by the particle filter, and the fact that often a mobile point detects multiple anchor points at the same time, to compute a reasonable approximation. While through the enclosures we steer the detection ranges towards a face-to-face measurement, there are still areas of overlap between the ranges where a mobile point detects multiple anchor points at the same time. This is particularly true for pieces positioned very close to each other. In occasions when a mobile point detects multiple anchor points the particle filter will predict a position that is in, or close to, those areas of overlap. That is because indeed it is likely that a person was effectively positioned not precisely in front of an artwork. Moreover, over a certain window of time, it is more likely that the mobile device will detect the closer anchor more frequently than the further anchor, and this information can be exploited by the filter as well. In this sense, particle filters applied locally to a number of close-by anchor points act as a localized localization system, though we still lack distance measurements for a precise localization. In other words, particle filters compute a sort of multi-lateration between close-by anchor points, for example, to triangulate the position of the visitors in a specific area of the museum. Finally, we can also sometimes estimate positions further from anchors detection range. For example, if we measure a visitor leaving an exhibit and appearing at another exhibit further away after some short time, and this time is coherent with walking speed and the distance



(a) Visitors positions heatmap

(b) Hot and cold spots

Figure 5: (a) Heatmap showing the amount of time spent in areas close to artworks. Warmer colors are associated with areas where the participants have cumulatively spent most time. (b) The relationship between the most (red arrows, IDs 42, 1, 15, 4, and 16) and the least (blue arrows, IDs 13, 8, 27, 28, and 29) viewed artworks respectively and their neighboring pieces. The pictures show the neighboring sources covering 50% of the total paths.

between the two exhibits, then we can position the visitor around the line connecting the two exhibits during that time.

We computed the estimations of the absolute positions for each visitor, for those moments in the visits where the estimate confidence was over a threshold we chose empirically. With these estimates, we produced the heatmap presented in Figure 5(a). The areas of the heatmap are colored according to the total amount of time visitors have spent in that area. While it does not precisely contain the same information, we can notice that this visualization resembles a spatial representation of the histogram presented in Figure 3. The visualization confirms that most of the visitors time was spent in room 1, 2, and 6. Also, as artworks 13, 14, 15, and 16 were within a distance of 1 meter from each other, the smoothing phase of the heatmap generation aggregates their values and it is difficult to distinguish artwork-related information. This is acceptable, and in fact expected, as the visualization shows spatial estimation.

#### 4.4. Artworks attraction and exhibition design

In some exhibitions, curators and designers guide visitors through a logical sequence of artworks. Such a sequence could be intended to describe a story or to keep the visitor engaged and interested during her visit. However, visitors might follow a different path or be unexpectedly attracted by some pieces more than others. To provide an insight about how artworks attract visitors, for example highlights, we produced the visualization depicted in Figure 5(b). Such visualization is based on the following analytics: by selecting the five most (*hot-spots*) and the least (*cold-spots*) attractive artworks we show their relation with the pieces that commonly precede them during visits (i.e., paths). All the arrows represent the most common transitions from one piece to the targeted one and the thickness relates to the frequency of each transition for that piece. In other words, a thicker arrow indicates that many visitors have first visited the source piece and the target piece right after. Additionally, we limit the number of incoming arrows to the ones that cover 50% of the total frequency for each target. In this manner, we focus our attention to the source artwork from which the

target mostly attracts the participants.

Figure 5(b) shows information related to the capacity of hot- and cold-spots to attract visitors. On the one hand, piece 1 mostly attracts visitors from the entrance and less from the neighboring pieces. For artwork 42, instead, visitor transitions are *evenly* distributed among the most immediate pieces. This result suggests that visitors may approach 42 due its attractiveness more than due to a sequential scanning of the room. However, such effect could also be attributed to its central placement with respect to its neighbors. On the other hand, cold-spots have less evenly distributed arrows, with a preference for the previous pieces. If we look at the artworks in room 4, we observe that the attraction is catalyzed towards the left part of the room (while it appears so, it is not a display of a path). In fact, the thicker arrow consistently comes from the closest piece on the right. This suggests that those artworks were mostly viewed while passing through the room, further suggesting, together with their popularity measures, the lack of interest of the participants in these pieces.

## 5. Clustering visitor behavior

In this section we present the application of data-mining techniques applied to visitor data. We used the set of  $r_v$  vectors and  $s_v$  path sequences from all 182 volunteers as dataset for two clustering tasks: (i) identifying common paths chosen by visitors when visiting the museum, and (ii) identifying patterns of the distribution of visiting time across rooms and exhibits. For both tasks, we utilize *Hierarchical Agglomerative Clustering* (HAC) [45], a bottom-up clustering algorithm where items initially form individual clusters and are merged in a series of iterations based on a clustering criterion. We chose the Ward method [46] as a criterion for choosing clusters to merge at each step, which focuses on minimizing total within-cluster variance.

The input of the algorithm is the distance matrix between all items in a dataset. To identify common paths, we compare all  $s_v$  sequences with the Jaro-Winkler (Jaro) [47] similarity metric, which is used to compute string similarity. Jaro is a type of string-edit distance that considers explicitly the mismatch in the order in which elements appear in two sequences (an operation called transposition) and how far mismatching elements are in the sequences. Intuitively, inverting in  $s_v$  the order of two exhibits that are nearby in the sequence is less penalized than inverting two exhibits far from each other in the sequence (e.g.  $\text{jaro}(\text{"ABCDE"}, \text{"ABCED"}) > \text{jaro}(\text{"ABCDE"}, \text{"AECDB"})$ ). Precisely, we use  $1 - \text{jaro}(a, b)$  as Jaro computes similarity while HAC requires distance. For the task of identifying patterns of time distribution, we compute the Euclidean distance between all  $r_v$  vectors. Before computing distances, we pre-process  $r_v$  vectors as follows. First, we use a threshold such that  $r_v$  contains only elements larger than 15 seconds (that is we consider for each visitor only the exhibits where she spent more than 15 seconds), and then we scale and center each adapted  $r_v$ . In other words, we transform  $r_v$  vectors into vectors describing how visitors dis-

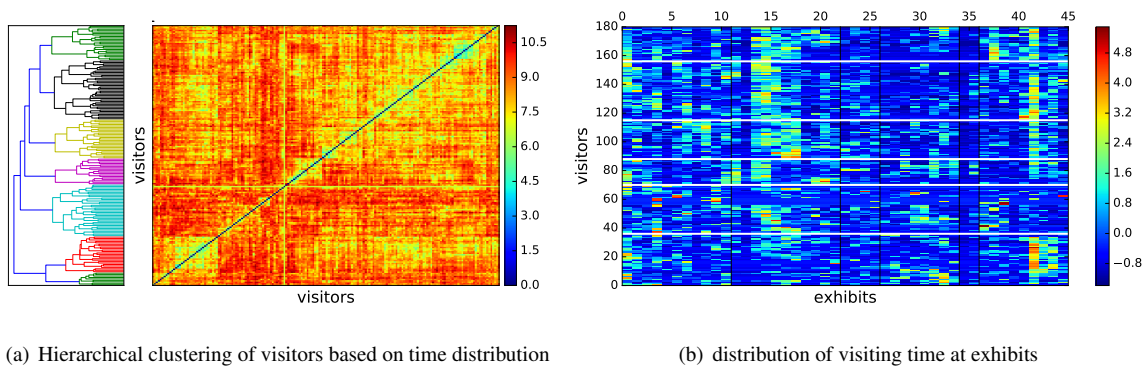


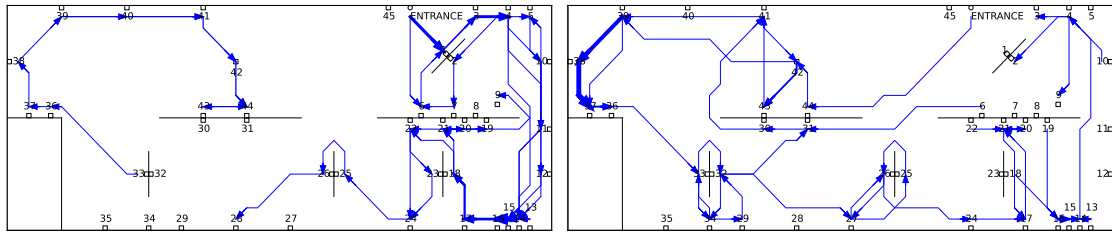
Figure 6: Hierarchical agglomerative clustering of visitor time distribution vectors. The clustering algorithm identifies one small group (1-38, bottom) that spent time mostly in room 6, and another major group (top) that is further clustered. Horizontal white lines show cluster divisions and black vertical lines room divisions.

tributed their time, among those exhibits where they spent more than 15 seconds. We then fed both datasets to the same clustering algorithm.

### 5.1. Time distribution across exhibits

In Figure 6(a) we show the distance matrix between the  $r_v$  vectors, which are organized according to the result of the agglomerative clustering displayed in the linkage dendograms, and in Figure 6(b) we show the set of pre-processed  $r_v$  vectors grouped by the result of the clustering algorithm (vertical black lines show room divisions and horizontal white lines show cluster division). The dendogram is a tree displaying the hierarchical clustering process, where each vertical line represents a merge between two clusters (the length of the horizontal line shows the distance of the merge). Because each cluster starts from a single row, one can notice that each merge starts by merging two rows together up to the root. The distance matrix shows the distance between each visitor. One can notice that distance between visitors in the same cluster is lower than between visitors between visitors belonging to different clusters. Because the visitors are organized according to the results of hierarchical clustering, the distance matrix can give an intuition of the boundaries of the clusters that are merged at each step.

The results show that the clustering algorithm identifies two major clusters. The first cluster includes the bottom visitors (1-38, light green and red in the dendogram) for their particular interest in 3 artworks in room 6 and little interest in room 2 and 3 except for exhibit 1. The second cluster includes all visitors between (65-182) who spent time in room 2 and in front of exhibit 42. This second cluster is further clustered in other clusters, for example, (i) (85-118, yellow) due to specific interest in the exhibits 7, 8, 9, (ii) (124-160, black) due to specific interest in the first three exhibits, (iii) (160-182, green) due to some specific interest in exhibits 38, 39, 40 and 41, as well as 19, 20, and 21. Finally, HAC puts all outliers in a cluster together (38-65, light blue), though some of them do show a pattern, as it is the case of 40-50, showing a common



(a) Paths extracted from the largest cluster

(b) Paths of volunteers visiting the exhibition in inverted order

Figure 7: Visualizations of two groups of common paths identified through hierarchical agglomerative clustering of visitor paths, one (a) representing the common trend among the largest group of visitors, and one (b) representing the 10% of visitors that visited the exhibition starting from the end.

interest in exhibits 2, 3, 4, as well as 14, 15, and 38 and 39.

## 5.2. Clustering of paths

We applied the same technique to paths. In Figure 7 we show two of the clusters identified with HAC. We constructed the visualization as described previously, but focusing only on each chosen cluster, and this time selecting the most frequent 20 LCSs.

The first cluster includes the largest group of visitors and shows the most common behavior. HAC identified further groups within this major cluster, for example, splitting visitors turning right towards exhibit 6 from visitors turning left towards exhibit 3 after exhibit 1 at the entrance. The second cluster in Figure 7(b) shows a cluster with 10% of the visitors who visited the exhibition space in “inverted order”, starting from room 6. Perhaps these visitors did not understand what was expected from them by the curators. Similarly, through HAC we identified another group of visitors who decided to visit room 6 first, only to return right back to room 1 and continued from there, perhaps after realizing their mistake.

Furthermore, we can notice that both groups make more different choices of paths while visiting the first rooms, whereas the last rooms are characterized by one common path where visitors scan the room sequentially along the outer walls. This phenomenon was dubbed “museum shuffle” by the staff of CoBrA and was associated to a decrease in attention after around 30 minutes in the visit.

## 6. Predicting visitor behavior

The previous section has shown patterns of group behavior in the data. That is, a group of individuals made similar choices during their respective visits and hence distributed the time in a similar way across artworks. Note that when we say that groups of individuals made similar choices, we do not mean that those visitors were together at the museum. In fact, most of the times they were not, and in fact visited the museum independently at a different time. Instead, the underlying assumption is that what these visitors shared was similar tastes and habits, which guided them to make similar choices.

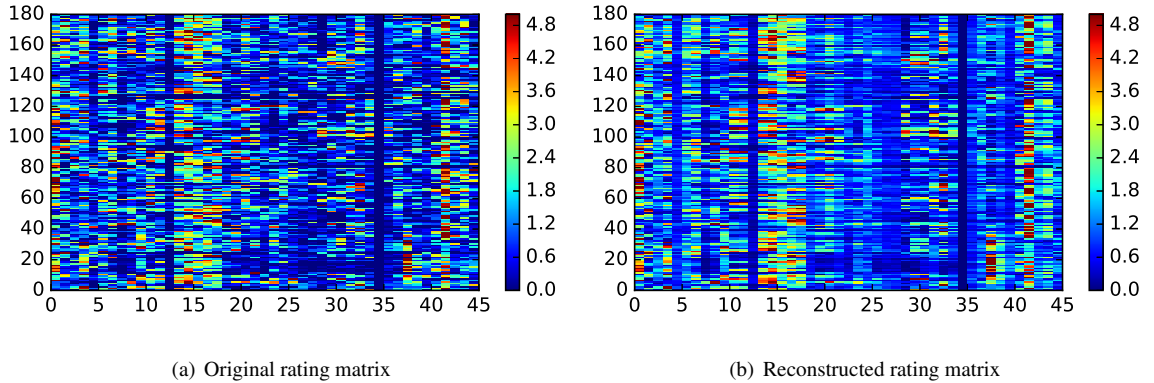


Figure 8: (a) The original rating matrix generated from the  $r_v$  vectors and (b) the reconstruction of such matrix from lower-rank matrices through Non-negative Matrix Factorization.

If such patterns exist, we should be able to exploit them to predict visitor behavior. If a visitor presents similar patterns of time distribution for some artworks as other visitors (i.e., she belongs to a cluster), we should be able to leverage what we know about those visitors’ visits to predict the behavior of the visitor at other artworks. In other words, we should be able to apply the “people who liked this item also liked these items” approach notably used in recommender systems, and in particular in collaborative-filtering approaches. These approaches are usually applied to predict user-item ratings based on other ratings.

To test this hypothesis, we model our data as follows. We generate a  $182 \times 45$  matrix  $R$  of ratings, where cell  $R[i, j]$  models the rating that visitor  $i$  has given to exhibit  $j$ . In other words, similar to the matrix introduced in Section 5, each row is constructed starting from a  $r_v$  vector. For each row, we compute the ratings based on the amount of time spent by the visitor facing the exhibits as follows. First, we standardize each cell by removing the median across the row and scaling according to the Interquartile Range (IQR). The IQR is the range between the 1st quartile (25th quantile) and the 3rd quartile (75th quantile). This is similar to removing the mean and scaling according to the standard deviation (i.e., the standard technique), but it is more robust to skewed datasets, as in our case (i.e., each visitor has many exhibits with low or zero values). Standardization of a dataset is a common requirement for many estimators. Then, we scale each cell in the row linearly to the continuous interval  $[0, 5]$ , such that the cells with the highest value are mapped to the rating 5 and the cells with the smallest value (often the value zero) are mapped to the rating 0. It is important to proceed with these two steps per-row, as visits have different lengths. The choice of the interval  $[0, 5]$  is arbitrary and was chosen to provide an intuitive understanding of the mapping and the scale of the errors. This procedure produces the matrix presented in Figure 8(a).

Once we have produced such matrix, we can use a state-of-the-art technique for ratings recommendation. We choose *Non-negative Matrix Factorization (NMF)* [48], as it is an established technique with an intuitive

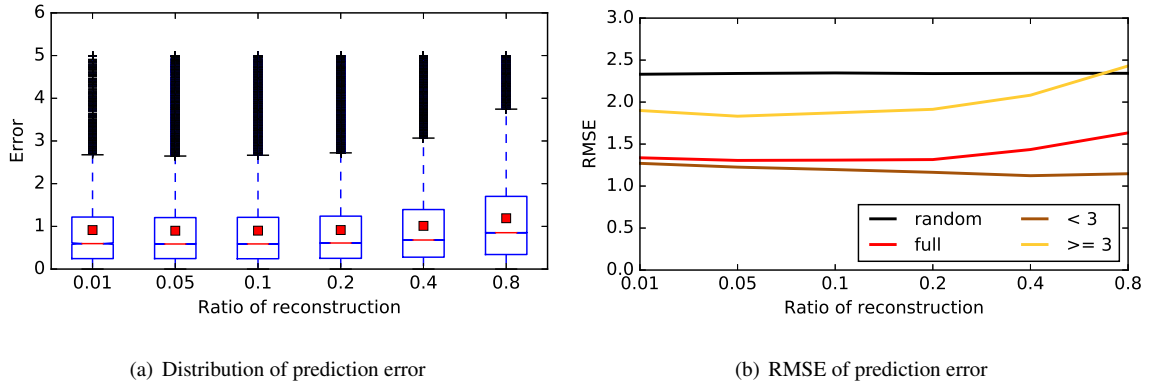


Figure 9: Prediction error when trying to predict visitor behavior through NMF for different reconstruction ratios. (a) Distribution of prediction error and (b) RMSE when choosing the test set with different criteria and against a random predictor baseline.

meaning<sup>4</sup>. The idea behind NMF is to compute two matrices whose product should produce the original matrix (i.e.,  $R$  in our case). NMF assumes the presence of some latent variables in  $R$  that can be identified and exploited to re-compute the original matrix. The two matrices usually have a lower rank, corresponding to the number of latent variables expected in the data (a reasonable choice is usually around the number of clusters expected in the data). When lower-rank matrices are used, NMF often produces values larger than zero in cells corresponding to un-rated items. These values can be used as predictions. We choose a rank of 15 though we had comparable results in the 10-30 range. For reference, Figure 8(b) shows the matrix as reconstructed by NMF using the full  $R$  matrix as input. One can notice that the patterns are still present in the reconstructed matrix, with an additional smoothening of the isolated values.

To establish the prediction capability of our model, we apply NMF to  $R$  as follows. We divide the ratings in  $R$  between a *training set* and a *test set* by choosing uniformly at random  $r$  ratings of the whole set, with  $r = 0.01, 0.05, 0.1, 0.2, 0.4, 0.8$  (hence we try to predict up to 80% of the original matrix using 20% of the data), and use these values as test set and the rest as training set. Then, we use the ratings in the training set to re-compute  $R$ , and compare the ratings in the test set between the values in the reconstructed matrix and in the original matrix. In practice, (i) we copy  $R$ , (ii) we set to zero the cells in the test set, (iii) we compute NMF of this matrix, and (iv) we compare the values in these cells to compute the prediction error, computed as the absolute difference between the two. Note that test set ratings can be initialized to the medium or mean of the corresponding row or column, but for simplicity we initialize them to zero (if we initialize the values to the mean of the corresponding column in the training set, we noted particularly an impact on larger values of  $r$ ). We repeat the test for each value of  $r$  100 times.

Figure 9(a) shows a box-and-whisker plot with the distribution of the prediction errors for each value of

<sup>4</sup>NMF gained general popularity in the field when a technique based on NMF won the Netflix prize.

$r$ . The box delimits the IQR, the band and the square in the box are the median and the mean respectively, the whiskers are set at the 1.5 of the IQR, and outer points are the outliers. We can see that the median prediction error is consistently beneath 1 for small values of  $r$ , increasing together with the error range only for large value of  $r$ . One thing to notice here is that outliers are also denser as  $r$  increases. These elements suggests that the error is accumulated around the higher ratings, while NMF is more accurate in predicting low ratings.

To validate this hypothesis, we compute the Root-mean-squared error (RMSE) aggregating all tests, selecting the test set in three different ways: (a) we select a test set across  $R$  as described previously, (b) we select a test set across  $R$  choosing only ratings smaller than 3, (c) we select a test set across  $R$  choosing only ratings larger than or equal to 3. In addition, we compute the error with a random predictor as a baseline. Figure 9(b) shows the RMSE for each value of  $r$ . One can notice that for small values of  $r$  NMF can achieve low RMSEs, that grow for  $r$  larger than 0.2. Moreover, we can also notice that errors are accumulated around large ratings more than by the smaller ones. These errors may be accumulated by those ratings that are outliers to the trends, as noticed previously by the smoothening effect of NMF in Figure 8(b). These results show that the data collected by the proximity sensors not only show group behavior, but that can also be used to predict the behavior of the visitors. Note that in terms of pure RMSE, we obtained lower errors when initializing the test set cells to the mean value of the column (i.e., around 1.6 for  $r = 0.01$ ), but minimizing the RMSE is out of scope of this paper.

## 7. Discussion

To understand the utility of the visualizations, we conducted an initial study with the art director, the curator, and director of business development of CoBrA. Due to the limited size of this study, its outcome cannot be considered and generalized as a main result of this work, yet it is valuable to bootstrap a discussion about the strengths and limitations of our approach, and the opportunities for future work. We followed a “think aloud” method [49] in which respondents interpreted the visualizations for us as they reviewed them. We discussed the visualizations two times, first with minimal explanation from us of what is depicted, and then again after we provided more explanation. In both cases, we asked staff members to explain the content, and share their interpretation. We ended with questions on which visualizations were most helpful and a discussion about the overall impact of the system.

The respondents were able to decipher most of the visualizations. At some points, they emphasized a different aspect of a visualization. The most pronounced example of differences in interpretation were for the histogram presented in Figure 3. The artistic director and the curator focused on which pieces were most and least popular, for example referring to the popularity of the authors, while the business developer looked at the distribution of time across rooms. In others, their interpretations coincided. After looking at artworks

popularity (Figures 3 and 5(a)), visitors paths (Figure 4(b)) and the transitions to hot spots (Figure 5(b)), staff came to the conclusion that the route was clearer in the rooms in which they spent more time and that visitors were more likely to move unpredictably in the rooms in which they spent less time.

All respondents used the visualizations to reason about visitor behavior in both the show depicted and future ones. Although the business developer found the bar chart of time distribution accessible, the artistic director said that bar charts were not her “language”, yet both respondents interpreted the visualization with ease. They used that visualization alone and in conjunction with the visualization of common paths to begin to question what they saw in the exhibition overall, why certain pieces and/or rooms commanded more attention. For example, noting that people spent less time and did not have a clear route in two rooms in the later half of the exhibition, they both wondered if later rooms commanded less attention because the work was less appealing or because of a dip in energy. The visualizations generated questions for them to test and they discussed possible solutions for the show under development including adding more benches later in the show and providing more guidance in those areas.

The respondents agreed on the value of some visualizations over others. They were most interested in paths through the museum (Figure 4(b)) and the heatmap (Figure 5(a)) showing time spent in different areas of the exhibition. They found a visualization of transitions to least popular pieces least relevant and transitions to most popular pieces somewhat relevant (Figure 5(b)). The curator was in particular surprised in noticing how about 10% of the visitors visited the exhibition starting from the end. Respondents requested the data to share with staff in discussion of future shows and made only minor suggestions for improvements (i.e., adding the images of the art pieces to the visualization).

The visualizations altered how they thought about several issues. The visualizations raised “awareness” of the relationship between “content and structure” of the show. The artistic director acknowledged that most of her thoughts about design focused on the content and that these visualizations emphasized for her the importance of finding the appropriate structure for the upcoming shows. For all respondents, viewing the visualizations also highlighted questions for which there is no correct answer, e.g., what is the goal of design? Is it to lead the visitor a certain way or simply provide guidance? Although for them, the answer differed from show to show, the artistic director concluded that she wants people to “linger” and she wants to “allow freedom within structure”.

While the visualizations show that certain artworks were more popular than others and that visitors consistently spent less time in certain rooms, the curator found that popularity is not necessarily the only metric to be trying to optimize for. The curator takes into account many different factors when she decides which artworks to exhibit, depending on the type of exhibition. Certain exhibitions often mix popular with less popular but historically or contextually related artworks, while so-called “blockbuster” exhibitions may indeed target more popular works of famous artists. Moreover, both the curator and the exhibition designer found that it is often expected that visitors focus only on certain pieces and may ignore some

rooms completely. Nonetheless, it is difficult to justify the consistent ignoring of certain rooms by the vast majority of the visitors.

In general, it is not obvious how to translate visitor behavioral data into a metric of success or satisfaction regarding the exhibition. While there is a correlation between the amount of time spent at an exhibition or in front of an exhibit and the reaction of the visitor, as shown also by our results in behavior prediction, the quality of that reaction needs to be further explored. The curator found that having behavior data available to the museum staff without a further understanding of such metrics could lead to misinterpretations about the outcome of the decisions of some members of the staff. The internal response of the visitor could be investigated with other types of sensors in conjunction to proximity data as, for example, physiological sensors [32] and accelerometers [50].

Finally, related to topic of understanding is the problem of how to intervene to improve or steer the response of the visitors, once issues have been identified. The strategies and heuristics of intervention remain a realm of the expert personal experiences skills of the museum staff, but further research is necessary, perhaps based on tools such as those presented in this paper, where differences in visitor behavior are measured as a result of curatorial changes (which would require longer (longitudinal) studies).

## **8. Conclusions**

In this paper, we apply sensing, data analytics and visualizations to the domain of fine art museum work. We examined the use of low cost sensors and a number of visualizations and data-mining techniques to understand the behavior of the visitors of a museum, towards a tool to support the collaborative and data-driven decision-making of museum staff. We conduct an experimental study for the technology in a real-world exhibition and an initial qualitative study to gather needs and evaluate visualizations of the data with museum staff. Results suggest the ability to use low cost sensors to capture visitor behavior metrics of interest to staff and the qualitative results indicate the immediate benefits of the use of these analytics in practice. Future work is needed to explore more conclusively the benefits of a visitor analytics tool for museum staff.

## **Acknowledgements**

This publication was supported by the Dutch national program COMMIT and by The Network Institute. We would like to thank the staff of CoBrA for all their time and support during this project.

## References

- [1] C. Martella, A. Miraglia, M. Cattani, M. van Steen, Leveraging proximity sensing to mine the behavior of museum visitors, in: *Pervasive Computing and Communications (PerCom)*, 2016 IEEE International Conference on, IEEE, 2016.
- [2] B. Lord, M. Piacente, *Manual of museum exhibitions* (2014).
- [3] M. Borun, When should you call a professional, <http://www.familylearningforum.org/evaluation/power-of-evaluation/when-professional.htm>, 2015. [Online; accessed 20-May-2015].
- [4] V. G. Museum, De volle zaal liever kwijt dan rijk, <http://www.slideshare.net/RuigrokNetPanel/mie-2013-defslide>, 2013. [Online; published 06-June-2013; accessed 22-May-2015].
- [5] P. Werner, *Museum, inc: inside the global art world 21* (2005).
- [6] G. E. Hein, Evaluation of museum programmes and exhibits, *Hooper-Greenhill* (1994) (1994) 306–312.
- [7] S. Voulgaris, M. Dobson, M. Van Steen, Decentralized network-level synchronization in mobile ad hoc networks, *ACM Transactions on Sensor Networks (TOSN)* 12 (2016) 5.
- [8] C. Martella, M. Dobson, A. van Halteren, M. van Steen, From proximity sensing to spatio-temporal social graphs, in: *Pervasive Computing and Communications (PerCom)*, 2014 IEEE International Conference on, IEEE, 2014.
- [9] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P.-J. Nordlund, Particle filters for positioning, navigation, and tracking, *Signal Processing, IEEE Transactions on* (2002).
- [10] R. E. Grinter, P. M. Aoki, M. H. Szymanski, J. D. Thornton, A. Woodruff, A. Hurst, Revisiting the visit:: understanding how technology can shape the museum visit, in: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, ACM, 2002.
- [11] B. Brown, M. Chalmers, *Tourism and mobile technology*, in: *ECSCW 2003*, Springer, 2003.
- [12] H. Fuks, H. Moura, D. Cardador, K. Vega, W. Ugulino, M. Barbato, Collaborative museums: an approach to co-design, in: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM, 2012.

- [13] L. Ciolfi, The collaborative work of heritage: open challenges for cscw, in: ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus, Springer, 2013.
- [14] M. Blöckner, S. Danti, J. Forrai, G. Broll, A. De Luca, Please touch the exhibits!: using nfc-based interaction for exploring a museum, in: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, 2009, p. 71.
- [15] N. Correia, T. Mota, R. Nóbrega, L. Silva, A. Almeida, A multi-touch tabletop for robust multimedia interaction in museums, in: ACM International Conference on Interactive Tabletops and Surfaces, ACM, 2010.
- [16] P. Luff, C. Heath, M. Norrie, B. Signer, P. Herdman, Only touching the surface: creating affinities between digital content and paper, in: Proceedings of the 2004 ACM conference on Computer supported cooperative work, ACM, 2004.
- [17] L. Bannon, S. Benford, J. Bowers, C. Heath, Hybrid design creates innovative museum experiences, *Communications of the ACM* 48 (2005).
- [18] M. H. Szymanski, P. M. Aoki, R. E. Grinter, A. Hurst, J. D. Thornton, A. Woodruff, Sotto voce: Facilitating social learning in a historic house, *Computer Supported Cooperative Work (CSCW)* 17 (2008).
- [19] P. Föckler, T. Zeidler, B. Brombach, E. Bruns, O. Bimber, Phoneguide: museum guidance supported by on-device object recognition on mobile phones, in: Proceedings of the 4th international conference on Mobile and ubiquitous multimedia, ACM, 2005.
- [20] R. Wakkary, M. Hatala, Situated play in a tangible interface and adaptive audio museum guide, *Personal and Ubiquitous Computing* 11 (2007).
- [21] T. Hope, Y. Nakamura, T. Takahashi, A. Nobayashi, S. Fukuoka, M. Hamasaki, T. Nishimura, Familial collaborations in a museum, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2009.
- [22] P. Tolmie, S. Benford, C. Greenhalgh, T. Rodden, S. Reeves, Supporting group interactions in museum visiting, in: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, ACM, 2014.
- [23] T. Moussouri, G. Roussos, Mobile sensing, byod and big data analytics: New technologies for audience research in museums (2014).

- [24] S. Macdonald, Interconnecting: museum visiting and exhibition design, *CoDesign* 3 (2007).
- [25] B. Serrell, The question of visitor styles, *Visitor studies: Theory, research, and practice* 7 (1996).
- [26] A. Vinciarelli, M. Pantic, H. Bourlard, Social signal processing: Survey of an emerging domain, *Image and Vision Computing* 27 (2009).
- [27] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava, Participatory sensing, *Center for Embedded Network Sensing* (2006).
- [28] S. A. Hoseini-Tabatabaei, A. Gluhak, R. Tafazolli, A survey on smartphone-based systems for opportunistic user context recognition, *ACM Computing Surveys (CSUR)* 45 (2013) 27.
- [29] Y. Yoshimura, F. Girardin, J. P. Carrascal, C. Ratti, J. Blat, New tools for studying visitor behaviours in museums: a case study at the louvre, *19th International Conference on Information Technology and Travel & Touris* (2012) 1–13.
- [30] G. Wilson, Multimedia tour programme at tate modern, in: *Museums and the Web*, volume 3, 2004.
- [31] E. Bruns, B. Brombach, T. Zeidler, O. Bimber, Enabling mobile phones to support large-scale museum guidance, *IEEE multimedia* (2007) 16–25.
- [32] V. Kirchberg, M. Tröndle, The museum experience: Mapping the experience of fine art, *Curator: The Museum Journal* 58 (2015).
- [33] E. Dim, T. Kuflik, Automatic detection of social behavior of museum visitor pairs, *ACM Transactions on Interactive Intelligent Systems (TiiS)* 4 (2014) 17.
- [34] T. Kuflik, Z. Boger, M. Zancanaro, Analysis and prediction of museum visitors behavioral pattern types, in: *Ubiquitous Display Environments*, Springer, 2012, pp. 161–176.
- [35] L. Chittaro, L. Ieronutti, A visual tool for tracing users' behavior in virtual environments, in: *Proceedings of the working conference on Advanced visual interfaces*, ACM, 2004, pp. 40–47.
- [36] J. Lanir, P. Bak, T. Kuflik, Visualizing proximity-based spatiotemporal behavior of museum visitors using tangram diagrams, in: *Computer Graphics Forum*, volume 33, Wiley Online Library, 2014, pp. 261–270.
- [37] R. Strohmaier, G. Sprung, A. Nischelwitzer, S. Schadenbauer, Using visitor-flow visualization to improve visitor experience in museums and exhibitions., in: *Museums and the Web (MW2015)*, 2015.
- [38] R. Reimann, A. Bestmann, M. Ernst, Locating technology for aal applications with direction finding and distance measurement by narrow bandwidth phase analysis, in: *Evaluating AAL Systems Through Competitive Benchmarking*, Springer, 2013.

- [39] C. Beder, M. Klepal, Fingerprinting based localisation revisited: A rigorous approach for comparing rssi measurements coping with missed access points and differing antenna attenuations, in: Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on, IEEE, 2012.
- [40] C.-L. Li, C. Laoudias, G. Larkou, Y.-K. Tsai, D. Zeinalipour-Yazti, C. G. Panayiotou, Indoor geolocation on multi-sensor smartphones, in: Proceeding of the 11th annual international conference on Mobile systems, applications, and services, ACM, 2013.
- [41] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, S. Sen, A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned, in: Proceedings of the 14th International Conference on Information Processing in Sensor Networks, ACM, 2015.
- [42] M. Analytics, Museum analytics, 2013. <http://www.museum-analytics.org>.
- [43] R. Stein, B. Wyman, Seeing the forest and the trees: How engagement analytics can help museums connect to audiences at scale., MW2014: Museums and the Web 2014 (2014).
- [44] M. Tröndle, S. Greenwood, K. Bitterli, K. van den Berg, The effects of curatorial arrangements, *Museum Management and Curatorship* 29 (2014) 140–173.
- [45] O. Maimon, L. Rokach, Data mining and knowledge discovery handbook, volume 2, Springer, 2005.
- [46] J. H. Ward Jr, Hierarchical grouping to optimize an objective function, *Journal of the American statistical association* 58 (1963).
- [47] W. E. Winkler, String comparator metrics and enhanced decision rules in the fellegi-sunter model of record linkage. (1990).
- [48] S. Zhang, W. Wang, J. Ford, F. Makedon, Learning from incomplete ratings using non-negative matrix factorization., in: *SDM*, volume 6, SIAM, 2006, pp. 548–552.
- [49] L. Cooke, Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach, *Professional Communication*, *IEEE Transactions on* 53 (2010).
- [50] C. Martella, E. Gedik, L. Cabrera-Quiros, G. Englebienne, H. Hung, How was it?: Exploiting smartphone sensing to measure implicit audience responses to live performances, in: *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference*, ACM, 2015, pp. 201–210.