Noninvasive Bluetooth Monitoring of Visitors' Length of Stay at the Louvre

The ubiquity of digital technologies is revolutionizing how researchers collect data about human behaviors. Here, the authors use anonymized longitudinal datasets collected from noninvasive Bluetooth sensors to analyze visitor behavior at the Louvre Museum.

Yuji Yoshimura
Senseable City Laboratory, MIT

Anne Krebs
Louvre Museum

Carlo Ratti
Senseable City Laboratory, MIT

Recent emerging technologies—along with their subsequent rapid diffusion into our daily lives—have caused a structural change in human behavior analysis. In particular, the ubiquitous presence of wired and wireless sensors in contemporary urban environments is producing a detailed empirical record of individual activities. Furthermore, in addition to the ubiquity of sensors, computationally advanced computer systems make it possible to accumulate large datasets of human behavior at high frequencies—sometimes even in real time.

However, despite the widespread use of such data collection technology, the analysis of visitor behavior in art museums has not advanced much over the last few decades. The traditional pencil-and-paper-based tracking method is still commonly used to time and track museum visits,[1] partly because many of the technologies that researchers started using a decade ago for human behavior data collection[2] don't work properly in a museum setting.[3] For example, video cameras installed to observe visitor behaviors have trouble distinguishing between individuals when the density of visitors is high,[4] and such cameras can't always track visitors as they move beyond the area of a particular exhibit.[1]

To address these issues, we employed a Bluetooth detection technique,[5,6] which has many advantages over other technologies. First of all, depending on the specification, Bluetooth can have a more finely grained detection scale than passive mobile phone tracking.[7] Second, Bluetooth detection successfully works inside buildings, where GPS connectivity is often limited. Finally, in contrast to RFID[8] or ultra-wideband,[9] previous participation isn't necessary to equip any devices or tags or to download the proper application in advance. Because prior participation or registration isn't needed, we can perform data collection for longer than just one day or a few days, letting us generate large-scale datasets of human behaviors.

All of these advantages make Bluetooth tracking the most adequate methodology for our research. It doesn't replace simulation-based analysis, which can estimate visitors' movement in the museum,[10] but because simulation often requires a simplification of human behaviors, Bluetooth tracking can complement the analysis by generating relevant datasets in

Published by the IEEE CS ■ 1536-1268/17/$33.00

a consistent way and letting us analyze real and large-scale empirical data.

Methodology: Bluetooth Tracking System

Bluetooth tracking systems for human behavior data collection can also complement more traditional social sciences' qualitative and quantitative methods. This is particularly useful for museum visitor studies, which have a long tradition of employing qualitative interviews, questionnaires, and ethnographic observations. (See Eilean Hopper-Greenhill's "Studying Visitors"[11] for a review of such studies.)

A Bluetooth tracking system works as follows: When a Bluetooth-activated mobile device enters the detectable area, a sensor receives the emitted signal from the mobile device until the signal disappears. Thus, the sensor registers the time at which the signal appears—referred to as the "check-in" time. When the signal disappears, the sensor records the "check-out" time. Then, the difference between each mobile device's check-in and check-out time can be calculated to define the length of stay at the node (note that "node" refers to the detectable area formed by each sensor). Similarly, by looking at the first check-in time and the last check-out time over all nodes, it's possible to calculate how long a visitor stays in the study area.

Such a series of check-in and check-out time data, registered by installed sensors, makes it possible to construct a human trajectory throughout the study area, including the travel time between nodes. In fact, because Bluetooth detection offers systematic observation through unobtrusive measures, a considerable amount of research has employed this methodology for human movement tracking.[5,6] More technical aspects of Bluetooth detection—including information about different Bluetooth antennas, the discovery time, and detection interferences—are summarized elsewhere.[4]

For our study, eight Bluetooth sensors were deployed throughout the Denon wing of the Louvre, covering key places to capture visitors' behavior. Figure 1 presents the approximate locations of the sensors (nodes E, D, V, B, S, G, C, P) and some of the most representative artworks or areas of the museum.

The museum's administrative policies (such as those related to artwork protection or areas with restricted access) and certain technical or spatial restrictions (such as the circulation conditions inside the galleries) largely determined the installation locations and sometimes prevented placement for optimal detection. As a result, some sensors are situated next to the relevant artworks, while others are located near but not right next to the artworks (see the detailed spatial relationship in Figure 1). The detection range of a sensor can be 20 meters long and 7 meters wide and can be customized for smaller scales if necessary. Although a sensor's range can fluctuate depending on its location, each one covers the targeted areas.

We performed the data collection at different periods using a different number of sensors over five months, from April through August 2010. All of this information was collected without invading visitors' privacy, because we applied a secure hash algorithm encryption to each sensor by converting each device's media access control ID into a unique identifier.

After data cleanup and processing, in which we adjusted the data to remove any inconsistencies, we selected 80,693 unique devices to be analyzed for this article. By comparing the number of detected mobile devices and ticket sales, we found that, on average, 8.2 percent of visitors had activated Bluetooth on their mobile phones.

Factors Related to Length of Stay

Here, we analyze four different factors related to visitors' length of stay in the museum. The first factor deals with visitor routes in relation with the length of stay in the museum. This analysis is largely based on our previous research. The second factor relates to entry times, which are used to assess the distribution of visitors' lengths of stay in the museum, depending on when they entered the museum. The third factor provides visitors' lengths of stay near each specific node, which correlates to a certain exhibit area, and the fourth factor is the relationship between the length of stay at a specific node and the number of visitors around the node (density).

The Path Taken

Figure 2a shows that the median length of stay is very similar across all amounts of unique visited nodes (for example, if someone visits nodes E-V-B-E, the number of unique visited nodes is 3—E, V, and B). The difference of visitors' lengths of stay between all unique visited nodes and the minimum one (visiting just two nodes) is less than one hour. In addition, the number of unique visited nodes seems to have a slightly negative slope with the lengths of stay until four unique nodes are visited. This means that individuals who visited three or four unique nodes tended to rush during their visits, as opposed to individuals who visited one or two unique nodes.

To compute the correlation, we used a nonparametric correlation analysis (Spearman's rank correlation coefficient), because the variables don't seem to follow a normal distribution. We also include a series of boxplots to better explain the relationship between variables. A very low correlation value ($\rho = 0.073$, $p < 2.2e\text{-}16$) suggests that there is no relationship between these two variables. That is, the number of unique nodes visited seems to be almost independent of the length of stay.

On the other hand, Figure 2b shows a different relationship between the lengths of stay and the number of visited nodes. The correlation coefficient ($\rho = 0.186$, p-value $< 2.2e\text{-}16$) suggests a weak association between the two variables. The total number of visited

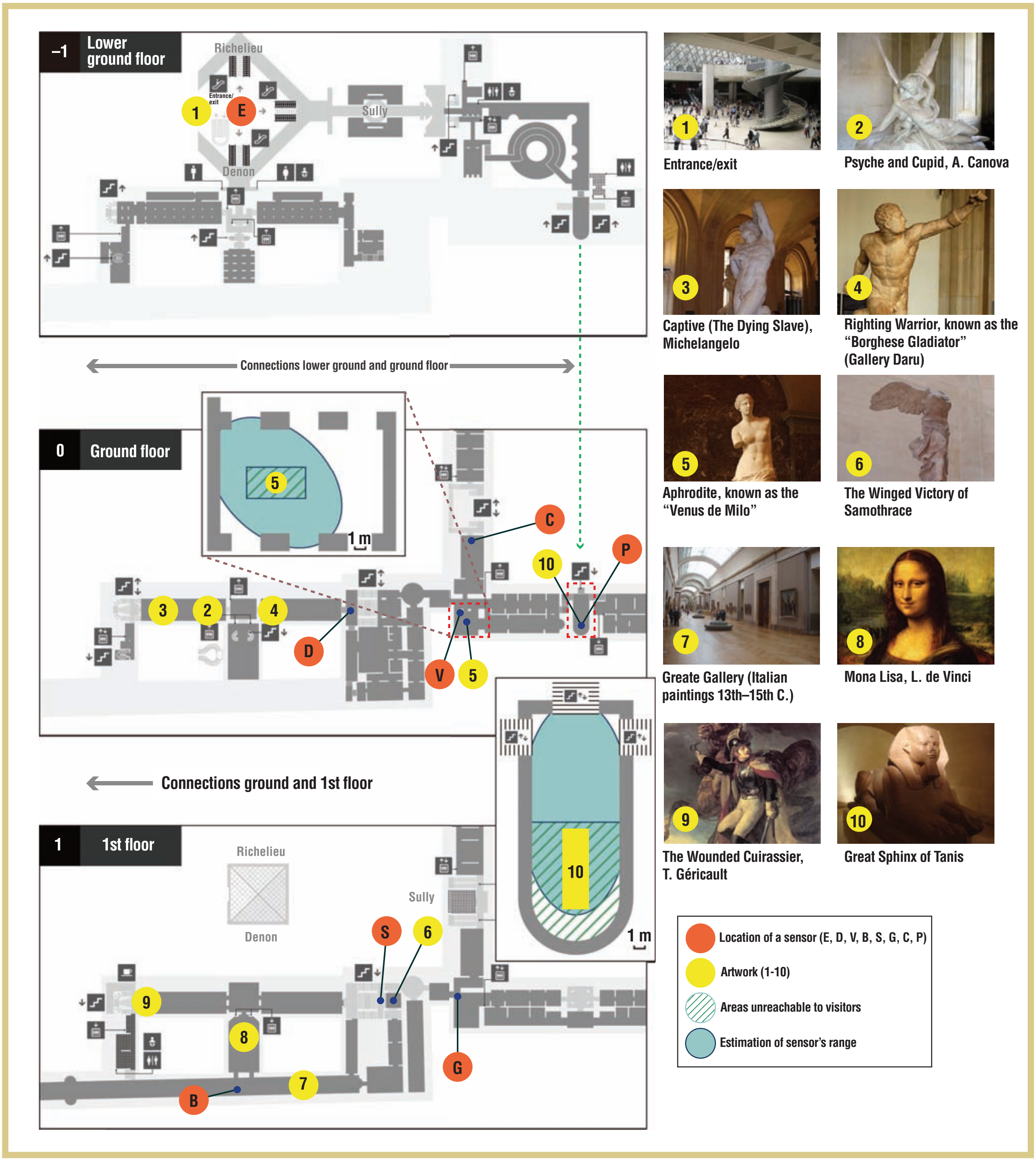

Figure 1. Approximate location of eight sensors (nodes E, D, V, B, S, G, C, P) with some of the most representative artworks of the Louvre collection departments nearby (1–10).

nodes slowly increases with the time spent in the museum, but not with the sharp rate of growth that one might expect.

Our previous research discovered that the behaviors of the *short-stay visitors* (the first 10 percent of the distribution in order of the shorter lengths of stay) and the longer-stay *visitors* (the last 10 percent of such distribution) are not significantly different in terms of the number

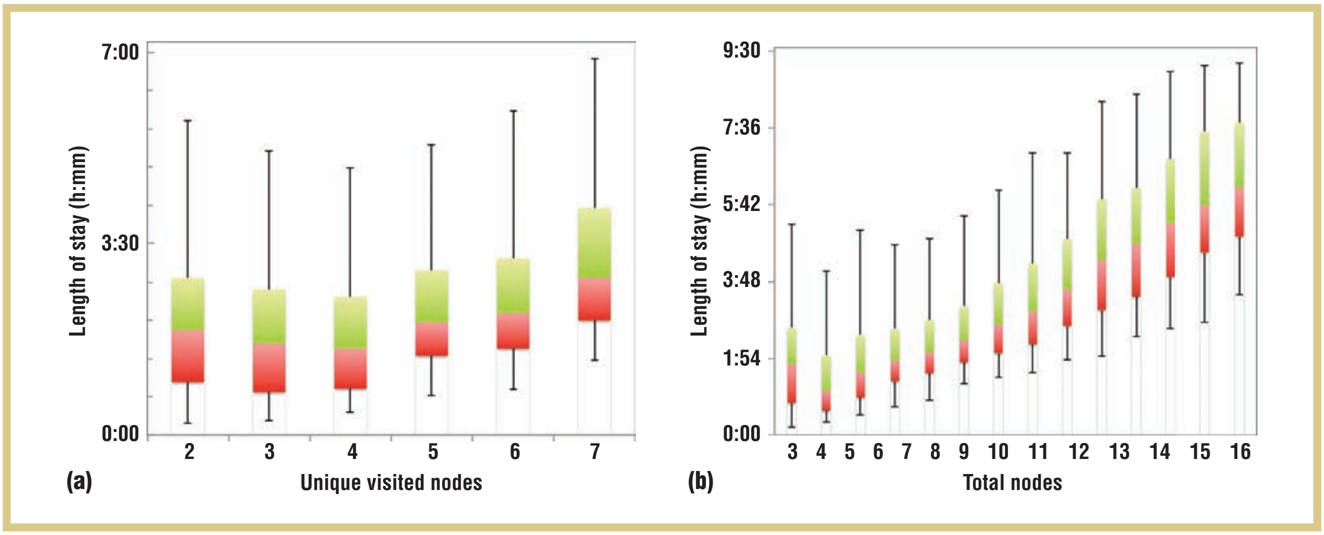

Figure 2. Considering visitor routes in relation to the length of stay: (a) The number of unique nodes visited versus the length of stay, and (b) the number of total visited nodes (shown <17) versus the length of stay. We took the 95th percentile as the upper extreme and the 5th percentile as the lower extreme.

of visited nodes and the visited order.[3] In addition, some routes of exploring the museum appear as frequent paths for both groups, even though the least frequent might be expected to be much more diverse in their choices given the extra time available. For example, E-D-S-B-D-V-C-E is the most frequently used path for both types of visitors who visited at least four nodes (see the green arrows in Figure 3). This path consists of entering from the Denon access gates and exiting from the Richelieu or Sully access gates and visiting most of the iconic artworks in the Louvre museum.

Intriguingly, it seems that short-stay visitors tend to choose this path more frequently than the longer-stay visitors, indicating that the former visitors explore all key locations possible within their limited staying times. They rush to visit the museum's "must see" artwork pieces, which are dispersed throughout the museum, rather than remain in a limited area. We speculate that both types' overlapping trajectories are causing the uneven distribution of visitors, resulting in some parts of museum becoming overcrowded.

Entry Times

We also examined a distribution of the average of visitors' lengths of stay in the museum classified by the time in which they visit. The goal was to determine whether visitors' entry times affected their lengths of stay at the museum. We divided the days of the week into two groups, separated by the two distinct closing times of the museum. The first group consists of Monday, Thursday, Saturday, and Sunday, when the doors close at 6 p.m., and the second group includes Wednesday and Friday, when the museum closes at 9:45 p.m.

As Figure 4a shows, when the museum closes at 6 p.m., the lengths of stay tend to decrease toward the closing hours of the museum. The earlier a visitor enters the museum, the longer that visitor tends to stay in the museum. Conversely, the results on Wednesday and Friday show a different tendency of visitors, as shown in Figure 4b. When the museum closes at 9:45 p.m., the lengths of stay at the museum decrease as time advances. However, the decrease in lengths of stay is slightly mitigated in the middle of the day. Just after the opening of the museum (from 10 to 11 a.m.), the length of stay is greatest, but in the late afternoon (from 5 to 6 p.m.), the lengths of stay increase slightly. So while some visitors maximize their utility (staying time) within the limited time the museum is open by visiting earlier, others try to take advantage of the longer hours and wait until the evening to visit.

All of these analyses and results indicate that the time visitors enter can be used to predict visitors' lengths of stay in the museum, but their lengths of stay in the museum don't correlate with the number of visited nodes over the course of their visit. Although the longer lengths of stay are correlated with a slightly larger number of visited nodes, the relationship is not significant.

The Length of Stay at Each Node

Our preliminary analysis showed that nodes E and S had much longer lengths of stay than the other nodes. The median lengths of stay at nodes E and S were 16:29 and 19:03 minutes, respectively, while the median length of stay at the other nodes was 3:14 minutes. Node E is at the desk for ticket sales, indicating (unsurprisingly) that visitors often wait in a long queue to purchase a ticket. Sensor S is located near the Winged Victory of Samothrace, which is one of the most famous and attractive

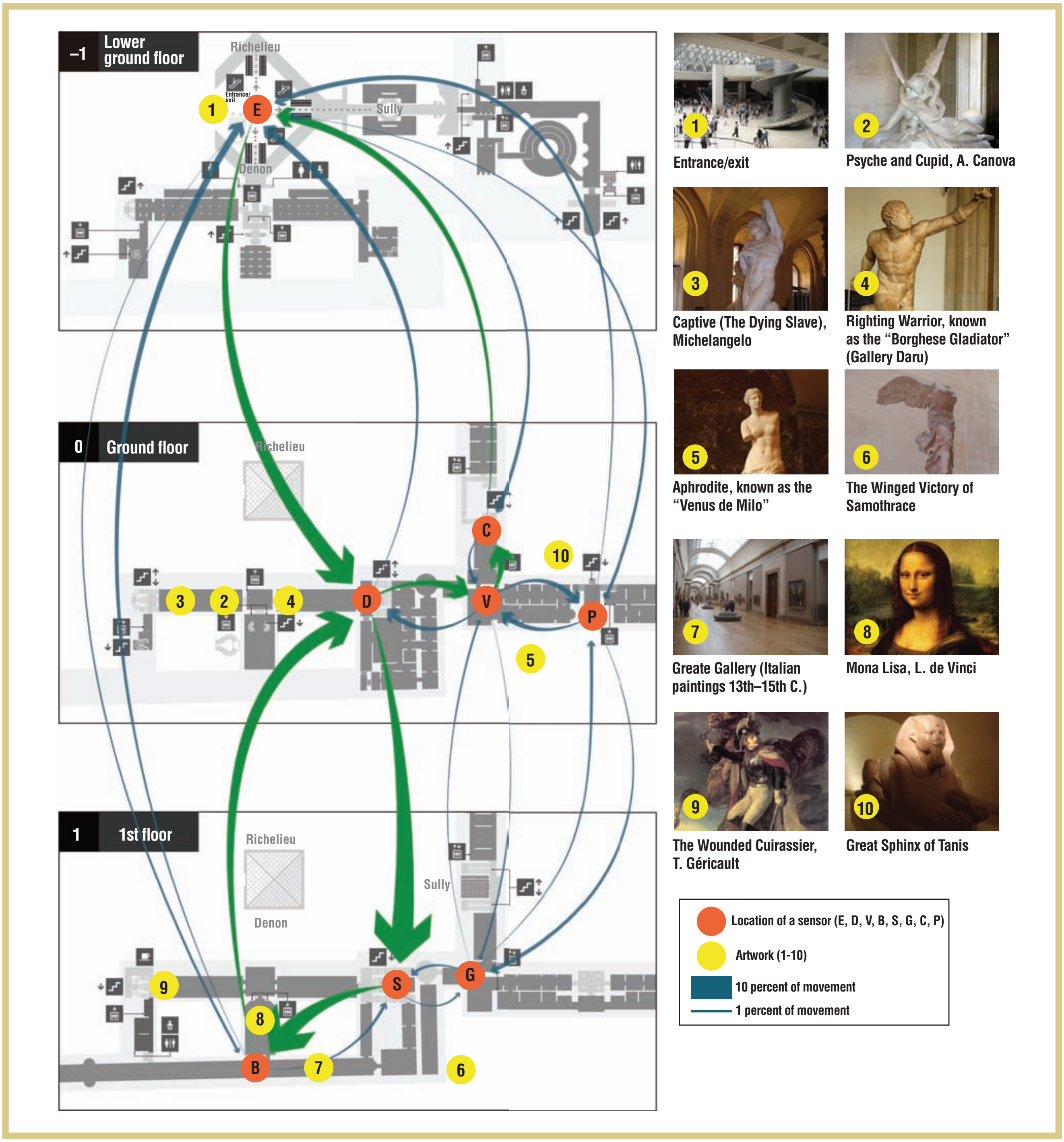

Figure 3. The arrows and their widths represent visitors' flow between nodes. The bold arrows each present 10 percent of visitors' flow over all movements, and the narrow arrows show 1 percent of movement. The green arrows show the most frequently used path (E-D-S-B-D-V-C-E) for those visiting at least four nodes.

exhibits in the museum. Also, there's a huge staircase in front of the exhibit (the *escalier Daru*), where many people sit down to rest during their visits. Those two factors make nodes E and S different from other nodes. We assume that the unique uses for these spaces result in the much longer lengths of stay, so we exclude these nodes from our following analysis.

Figure 5a presents the boxplot of visitors' lengths of stay at each node. We can observe that two groups exist: Nodes V, C, and B experience longer lengths of stay; nodes D,

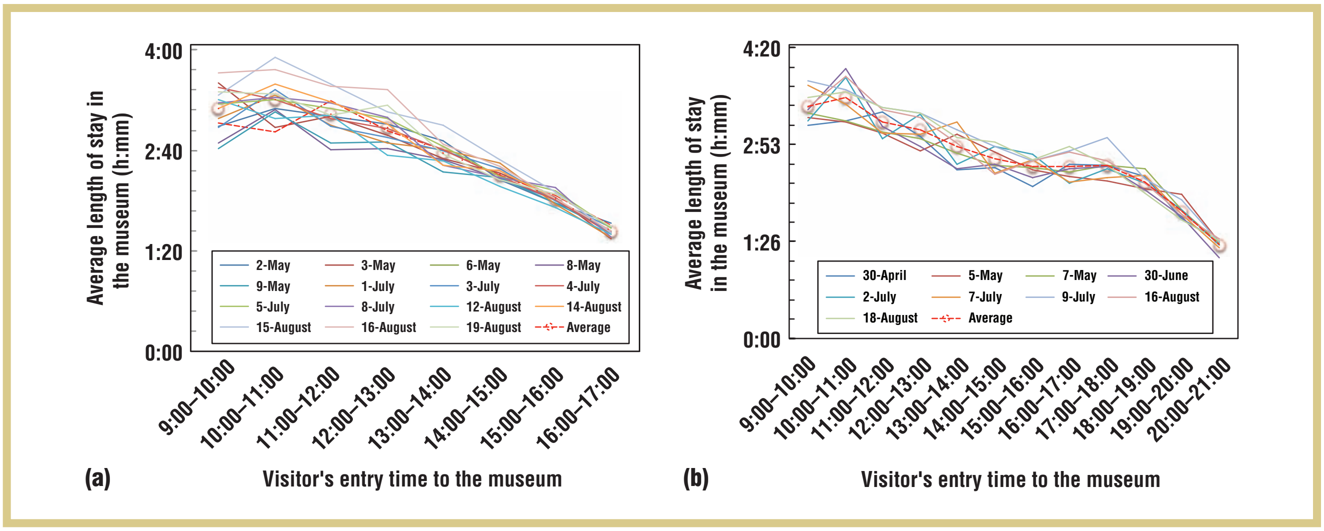

Figure 4. Visitors' length of stay versus their entry time. The distribution of the average stay times by visiting hours on (a) Monday, Thursday, Saturday, and Sunday, when the museum closes at 6 p.m., and (b) on Wednesday and Friday, when the museum closes at 9:45 p.m.

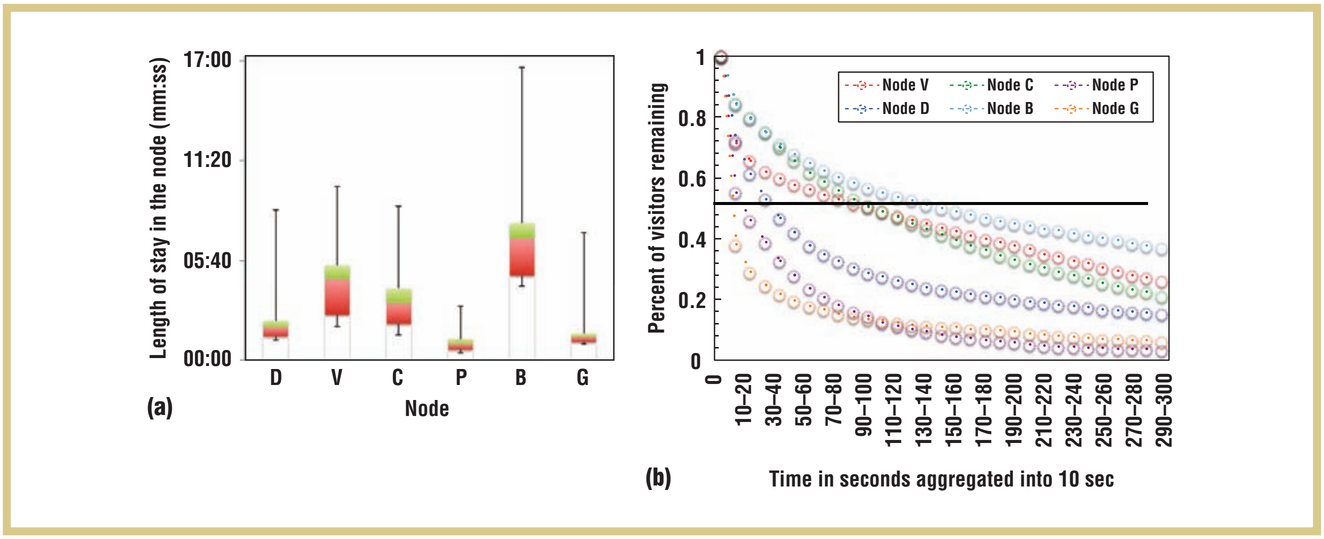

Figure 5. Analyzing the length of stay at each node. (a) The boxplot of the lengths of stay in each node. We took the 95th percentile as the upper extreme and the 5th percentile as the lower extreme. (b) The comparative visitor decay curves.

P, and G experience shorter stays. While the median length of stay for the former group is 157 seconds (σ = 304 seconds), the median for the latter is 49 seconds (σ = 203 seconds). Among these nodes, the range of visitors' lengths of stay at node P is much shorter than the range of other nodes; 90 percent of node P's visitors have lengths of stay between 6 seconds and 186 seconds.

Conversely, Figure 5b shows the comparative "visitor survival curve" (which is frequently used in visitor studies to analyze lengths of stay), when half of the visitors leave an exhibit or room.[12] We observe that the lengths of stay are largely varied among nodes: 20–30 seconds for node D, 90–100 seconds for nodes V and C, 0–10 seconds for node P, 130–140 seconds for node B, and 0–10 seconds for node G. However, most of the nodes experience lengths of stay within a few minutes. Again, we can observe that two kinds of nodes exist, which we can classify as shorter- and longer-stay nodes. Within the prior group, half of the visitors left nodes G and P within 10–20 seconds. In the case of node D, the stays last a little bit longer, but half of the visitors stayed for just 20–30 seconds.

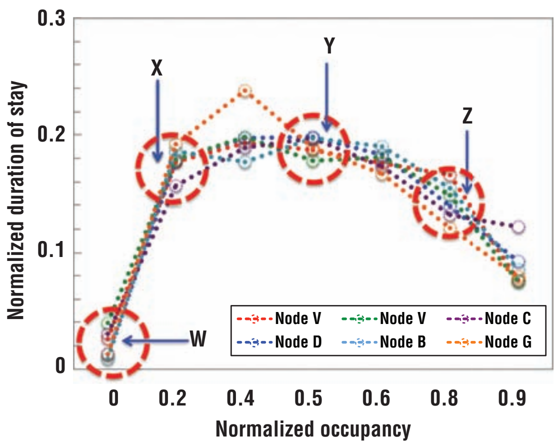

Figure 6. Distribution of the normalized occupancy versus the length of stay. Point X can be considered the equilibrium between an ideal length of stay for visitors when they are free to stay as long as they want.

TABLE 1
The threshold of the normalized occupancy near each node. Point X is the ideal length of stay for visitors who are free to stay as long as they want. Point Y is the maximum. After that, lengths of stay drop down as the occupancy level starts to exceed a certain threshold (point Z).

	Length of stay in seconds (and the normalized occupancy level)		
Node	**Point X**	**Point Y**	**Point Z**
D	183 sec. (0.235)	204 sec. (0.498)	188 sec. (0.636)
V	289 sec. (0.235)	315 sec. (0.368)	271 sec. (0.781)
C	225 sec. (0.221)	249 sec. (0.352)	187 sec. (0.753)
P	94 sec. (0.214)	118 sec. (0.5)	79 sec. (0.764)
B	353 sec. (0.238)	375 sec. (0.504)	293 sec. (0.772)
G	167 sec. (0.267)	208 sec. (0.351)	145 sec. (0.657)

The Relationship Lengths of Stay and Density

The perspective on lengths of stay, however, greatly changes when examining visitors' duration of stay in relationship to each node's degree of the occupancy. Figure 6 shows the relationship between each node's occupancy normalized by the maximum number of visitors in the area (x-axis) and the average duration of stay expressed in seconds (y-axis). As we can see, a clear tendency exists among the data. The average duration of stay first goes up with the room occupancy from point W to point X, and then stays around the maximum (point Y) on some occupancy level interval (point X to Z). After that, lengths of stay drop as the occupancy level starts to exceed a certain threshold (point Z).

The length of stay for point W is extremely short because it includes visitors who just pass by the area rather than stay. When the occupancy level increases from point W to X, the number of visitors who tend to stay longer also increases. Point X can be considered the equilibrium between an ideal length of stay for visitors when they are free to stay as long as they want. This is because visitors are free to look at the artwork between points W and X without any obstacles—in particular, the low density of other visitors. Conversely, from point X onward, the average length of stay remains almost flat until point Z, at which point visitors' lengths of stay start to decrease drastically. We speculate that this is because the high-density of other visitors can affect a visitor's comfort, resulting in a desire to escape the crowd.

The length of stay for point X varies depending on the node. Node B has the longest length of stay for point X (353 seconds), with a 0.238 normalized occupancy level among the other nodes (see Table 1). Conversely, node P represents the shortest length of stay for point X (94 seconds), with a 0.214 normalized occupancy level among the other nodes. Although the occupancy levels of those two nodes are similar, the former's length of stay is almost four times longer than the latter's. However, the normalized occupancy level for point X for node B is higher than that of node P (0.238 vs. 0.214, respectively). In addition, these nodes correspond to the maximum and minimum lengths of stay for point X, although the occupancy levels of both nodes are quite similar: Both are around 0.50. Furthermore, the length of stay at node P starts to decrease earlier than that of node B.

Regarding the relationship between nodes V and B, although point X for both shows a similar normalized occupancy level (0.235 vs. 0.238, respectively), the maximum length of stay (point Y) of node V has a much lower occupancy level than those of nodes P and B (0.368 for node V). Additionally, node V has the highest density for point Z (0.781) with the second longest length of stay (271 seconds), whereas the longest length of stay is at node B (293 seconds).

All of these facts indicate that node B better attracts and holds visitors than nodes V and P. Node B seems to inspire visitors to stay longer, even during a higher occupation density, while nodes V and P seem to cause visitors to stay for shorter durations when experiencing the same density. Also, the data shows a visitor's average length of stay and the node's occupancy level form a clear pattern. The crowd density around the node largely affects a visitor's length of stay either positively or negatively, but the effect can differ depending on the node.

We speculate that, up to certain occupancy limits, visitors are actually attracted by the crowd, although the attraction also depends on the nature of the galleries and the type of visitors. However, once the crowd's size reaches a certain level, visitors will try to avoid the node. This indicates that we can use point Z as a threshold to distinguish visitors' level of comfort, which enables us to manage the environment to avoid exceeding this density threshold. Also, we might want to consider some characteristic points, such as points X and Y, to enable the management of different types of environments or crowds inside the galleries.

Our methodology can fill in the gap in previous research on visitor behaviors in art museums. State-of-the-art technology enables us to record human behaviors in finer granular scale in space and time than the conventionally used methodologies.[2] In addition, our methodology enables us to analyze visitors' behaviors in a dynamic way that considers temporal factors, such as the time of day or seasonal effects. Using conventional methodologies, it's difficult to analyze and compare temporal factors because of the small samples collected during short periods of time or because of the reliance on indirect variables (such as a museum's spatial parameters[10]). Thus, our methodology and conventional methodologies are complementary rather than exclusive.

Although our proposed methodology generates valuable information for a more efficient crowd management tool, some issues remain. First, the sensor can only detect mobile devices that have their Bluetooth functions turned on. This indicates that the sample's representativeness might have a strong bias toward certain groups. Second, the sensor enables us to detect visitors' presence in the specific area, but such sensors can't specify whether visitors are actually looking at the artwork or if they are simply in the area. Third, although the qualitative method using the interviews and questionnaires can unveil visitors' psychographic factors, such as their expectations, experiences, and level of satisfaction,[11,13] Bluetooth detection techniques cannot disclose the visitors' motivations and inner thoughts in any way—this method merely identifies their presence and their precise lengths of stay. Finally, sensors can't collect sociodemographics (such as origin, age, gender, or profession) or other traditional behavioral variables.

Our findings can help with the management of visitor flow to reduce congestion at specific areas and around specific pieces of artwork. They can also be used to bring attention to less visited or less "attractive" artworks or rooms inside the museum by proposing interpretation tools or walking tours capable of increasing the value of such neglected spaces. Additionally, our data suggests that visitor behavior is based on some patterns, which make it possible to foresee their future movement in a dynamic way. Also, the data is useful for designing the spatial arrangement (perhaps improving the layout of exhibits, facilities, interpretation tools, and advertisements), depending on visitor activities and use of space. Finally, our findings indicate that efficient and effective congestion management of the museum can be realized by limiting the number of visitors that can enter, determined by the time of the day or year.

ACKNOWLEDGMENTS

We thank the Visitor Assistance and Security Department of the Louvre Museum for their support. We further thank Allianz, American Air Liquide, the Amsterdam Institute for Advanced Metropolitan Solutions, Ericsson, the Fraunhofer Institute, the Liberty Mutual Institute, Philips, the Kuwait-MIT Center for Natural Resources and the Environment, Singapore-MIT Alliance for Research and Technology (SMART), UBER, UniCredit, Volkswagen Electronics Research Laboratory, and all the members of the MIT Senseable City Lab Consortium for supporting this research.

REFERENCES

1. S.S. Yalowitz and K. Bronnenkant, "Timing and Tracking: Unlocking Visitor Behavior," *Visitor Studies*, vol. 12, no. 1, 2009, pp. 47–64.
2. S. Rasouli and H. Timmermans, *Mobile Technologies for Activity-Travel Data Collection and Analysis*, IGI Global, 2014.
3. Y. Yoshimura et al., "An Analysis of Visitors' Behavior in The Louvre Museum: A Study Using Bluetooth Data," *Environment and Planning B: Planning and Design*, vol. 41, no. 6, 2014, pp. 1113–1131.
4. N. Abedi, A. Bhaskar, and E. Chung, "Tracking Spatio-Temporal Movement of Human in Terms of Space Utilization Using Media-Access-Control Address Data," *Applied Geography*, vol. 51, 2014, pp. 72–81.
5. V. Kostakos et al., "Brief Encounters: Sensing, Modelling and Visualizing Urban Mobility and Copresence Networks," *ACM Trans. Computer Human Interaction*, vol. 17, no. 1, 2010, pp. 1–38.
6. M. Versichele et al., "The Use of Bluetooth for Analysing Spatiotemporal Dynamics of Human Movement at Mass Events: A Case Study of the Ghent Festivities," *Applied Geography*, vol. 32, no. 2, 2012, pp. 208–220.
7. C. Ratti et al., "Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis," *Environment and Planning B: Planning and Design*, vol. 33, no. 5, 2006, pp. 727–748.
8. T. Kanda et al., "Analysis of People Trajectories with Ubiquitous Sensors in a Science Museum," *Proc. 2007 IEEE Int'l Conf. Robotics and Automation* (ICRA), 2007, pp. 4846–4853.
9. M. Tröndle et al., "An Integrative and Comprehensive Methodology for Studying Aesthetic Experience in the Field: Merging Movement Tracking, Physiology,

and Psychological Data," *Environment and Behavior*, vol. 46, no. 1, 2014, pp. 102–135.

10. B. Hillier and K. Tzortzi, "Space Syntax: The Language of Museum Space," *A Companion to Museum Studies*, Blackwell Publishing, 2006, pp. 282–301.
11. E. Hooper-Greenhill, "Studying Visitors," *A Companion to Museum Studies*, Blackwell Publishing, 2006, pp. 362–376.
12. S. Bicknell, "Here to Help: Evaluation and Effectiveness," *Museums, Media, Message*, Routledge, 1995, pp. 281–293.
13. G. Hein, *Learning in the Museum*, Routledge, 1998.

the AUTHORS

Yuji Yoshimura is a researcher affiliated with the Massachusetts Institute of Technology's Senseable City Laboratory and a regular research partner of the Socio-economic Studies and Research Division of the Louvre Museum. He is also a founding partner of "laboratory urban DECODE," an international design office dedicated to urban innovation. His research interests include ubiquitous computing and the analysis of urban dynamics through sensor networks. Yoshimura has a PhD in computer sciences from the Universitat Pompeu Fabra. Contact him at yyoshi@mit.edu.

Anne Krebs heads the Studies and Research Division of the Louvre Museum, including audience studies and research programs addressing the social and economic role of museums, digital cultural engagement and the social determinants of individual giving to support art and culture. Krebs has a postgraduate diploma in humanities and social sciences from the Sorbonne University. Contact her at anne.krebs@louvre.fr.

Carlo Ratti is a professor of the Practice of Urban Technologies at the Massachusetts Institute of Technology, where he directs the Senseable City Laboratory. He is also a founding partner of the international design and innovation office Carlo Ratti Associati. His research interests include urban design, human-computer interfaces, electronic media, and the design of public spaces. Ratti earned his PhD in architecture from the University of Cambridge, UK, and moved to the Massachusetts Institute of Technology as a Fulbright Senior Scholar. Contact him at ratti@mit.edu.