Mobile Phones and Outdoor Advertising: Measurable Advertising

A system for measuring audiences of outdoor advertising in specific areas is based on the combination of mobile phone location estimations with Internet listings of social events.

Daniele Quercia
University of Cambridge

Giusy Di Lorenzo and Francesco Calabrese
IBM Dublin Research Laboratory

Carlo Ratti
Massachusetts Institute of Technology

Online advertising is the fastest-growing advertising medium, not least because it can track not only how long someone was on a webpage with an advertisement but also how many times someone clicked on that ad. By contrast, outdoor advertising has not reached its full potential because it cannot measure return on investment. People spend 27 percent of their time exposed to outdoor advertising, but such forms of advertising attracted only 5 percent of US media spending in 2008.[1] The problem is that, for measuring advertising effectiveness, media planners currently rely on gross traffic numbers or circulation counts from the Traffic Audit Bureau, which represent historical data that has never been audited. As expected, unvalidated measurements do not call for marketing dollars.

The industry would embrace outdoor advertising only if credible audience measurements were introduced. A more credible way to measure audiences for a billboard should include both the number of people in front of the billboard and the likelihood that those people might like a specific ad shown on the billboard. We propose a system that estimates the number of people based on the number of mobile phones near the billboard and infers people's preferences by combining location estimations from the mobile phones with listings of social events (such as a football game or music festival) that are freely available on the Internet.

Audience Measurement from Location Estimations of Mobile Phones

To properly measure a potential audience, the first step is to determine the number of people near a billboard. We do so in three steps.[2,3]

Step 1. *We collect location estimates of mobile phones.* From AirSage, we collected estimates of the locations of 1 million mobile phone users in the greater Boston area (20 percent of the entire population). The logs span 1.5 months in 2009 and are generated each time a mobile phone connects to the cellular network (that is, whenever the phone places or receives a call, sends or receives a text message, or is connected to the Internet).

Our dataset contained 130 million pairs of latitude and longitude estimates with corresponding time stamps. Mobile phone-derived location data has a greater uncertainty range than GPS data, with an average of 350 meters and median of 220 meters, as reported by AirSage based on internal and independent tests. More formally, a location estimate m_i is characterized by a position p_m that is expressed in latitude and longitude and is associated with a time stamp t_{m_i}. Mobility trajectories (such as those in Figure 1) are then formed by linking location estimates together to form temporal sequences of the form $\{m_1, m_2, \ldots, m_n\}$. Because location estimates

Published by the IEEE CS ■ 1536-1268/11/$26.00 © 2011 IEEE

are not regularly sampled but are only taken when the user connects to the cellular network, the data is a sparse representation of people's mobility. To examine the extent that this reflects people's mobility, we measured the interevent time (time between two consecutive network connections) for each user. Figure 2 shows the distribution of the first and third quartiles and the median of interevent time for the entire population. The arithmetic average of the medians is 84 minutes and the geometric average is 10 minutes. This means that we can use the data at hand for inferring people's mobility with a temporal resolution of approximately 1.5 hours. To make our analysis computationally tractable, we work on stratified samples of 80,000 users at a time.

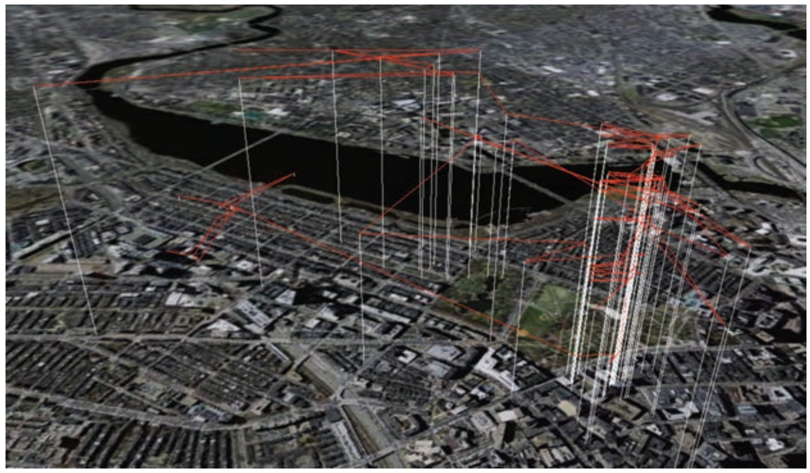


Figure 1. Mobility trajectories extracted from mobile phone data. A trajectory is a temporal sequence of a user's consecutive location estimations and is a reliable representation of the user's mobility with a spatial resolution of 350 meters and a temporal resolution of 1.5 hours.

Step 2. *We generate mobility trajectories from the location estimates.* From the previous step, we have a set of visited locations for each user. To make sense of those locations, we extract individuals' trips (trajectories) from them. Going, say, from home to work and then from work to the gym creates two different trajectories. We define a trajectory to be a set of consecutive locations visited by the user (which we call *stops*). We extract a user's trajectories by doing the following:

- *Identifying the user's trajectories.* If two subsequent location estimates are registered within two hours of each other, they are considered part of the same trajectory; if they are registered more than two hours apart, they form two separate trajectories. We took two hours from our temporal resolution of 1.5 hours (being the interevent time less than 1.5 hours [Figure 2]).
- *Removing noise from each trajectory.* We apply a low-pass filter on each trajectory's location estimates with a resampling rate of 10 minutes.[4,5]
- *Characterizing each trajectory as a set of consecutive stops.* We define a stop as the centroid of a set of consecutive location estimates that are registered within 500 meters. That is, a stop is the centroid of a set of location estimates $\{m_q, m_{q+1}, \ldots, m_z\}$ registered during the time interval $[t_{m_q}, t_{m_z}]$ for which $\max distance(p_{m_i}, p_{m_j}) < 500m \forall q \leq i, j \leq z$. A set of consecutive stops then forms a trajectory.

Compared to recent studies,[6] our geographical resolution is not limited to the area typically covered by a cellular network tower but is as little as 350 meters, as location estimations are accurately determined by a proprietary triangulation technique (that is, by AirSage's Wireless Signal Extraction Technology). Also, we consider many location estimations that are derived not only from calls but also from text messages and Internet surfing.

Step 3. *We finally count the number of mobile phone users in one area.* After this data processing, we can infer the number of mobile phone users in a certain area with a temporal resolution of 1.5 hours and a geographical resolution of 350 meters.

Ranking Social Events in Each Area of Residence

After determining the number of people in one area, we need to determine their preferences for social events. To this end, we determine the social events that the residents of an area (the people who make the area their most frequent stop between 10 p.m. and 7 a.m.) have likely attended. Then, in each area, we characterize residents' preferences based on our ranking of events.

Attendance at Social Events

We crawled the *Boston Globe* Calendar website (http://calendar.boston.com) to extract the social events that took place during our period of study. This website is a reputable and comprehensive list of more than 500 daily social events in greater Boston.[2] The events are organized in 14 categories: arts and crafts, business and tech, community, dance, education/campus, fairs and festivals, food and dining, music, other, performing arts, shopping, sports and outdoors, visual arts, and cinema.

To geographically map the events, we divided the 15-km^2 area of greater Boston in geographic cells of 500 × 500 meters. The choice of 500 meters

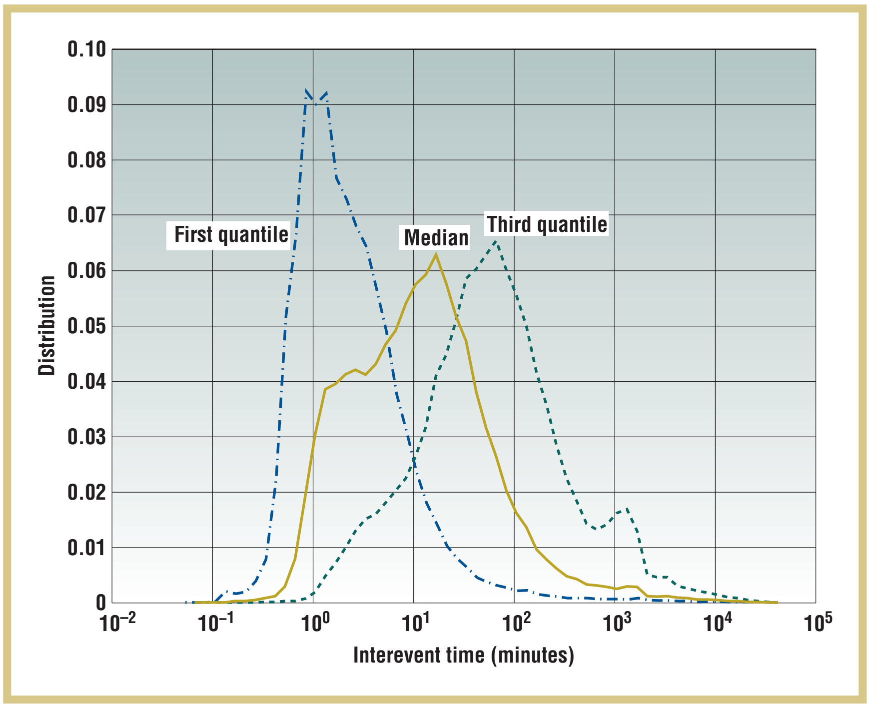


Figure 2. Distribution of interevent time (time between two consecutive cellular network connections) for our mobile phone users: median (solid line), first quartile (dash-dotted line), and third quartile (dashed line).

is conservative and is related to our average localization error of 350 meters. We then place the events and user stops in the corresponding cells. By intersecting user stops and social events at each cell at a specific point in time, we determine which users attended which social events. However, as one might expect, this approach may well produce false positives—that is, we might find that some people attended social events they had never been to. To reduce false positives, we undertake the following measures. First, we consider only people who live in a different location from the event location and stay for at least 70 percent of an event duration. We do not opt for full 100 percent overlap because this would require a person to make one call right before the event and another call right after it. Second, we find the largest set of events that satisfy the following five requirements:

- The event is highly attended, so the likelihood of being attended by our mobile users is significant.
- The event is geographically isolated compared to neighboring events. We impose a minimum radius of one kilometer between the event and any other large concurrent event.
- The event takes place in a well-defined geographic area of considerable dimension, thus minimizing the possibility of mistakenly considering people staying in places close to the event's venue (for example, in a nearby restaurant).
- The event is temporally isolated from any other large event.
- The event has a minimum duration of two hours. This makes it possible to distinguish between occasional stops and event attendance.

Those requirements separate events temporally and geographically, and make it possible to distinguish between people staying in a place and people going to a social event. Importantly, this conservative way of filtering drastically reduces our dataset, but it also ensures that we will work upon correct inferences with high probability. Concretely, to reduce false positives, we also reduce the number of social events from 500 a day to 53 for the entire period of study (some of which are shown in Figure 3), and the number of mobile users from 80,000 to 2,519. However, on the positive side, we can capture a large fraction of the population in a meaningful way. Our mobile phone users represent 20 percent of the entire population of greater Boston (as per latest US census) and are proportionally distributed across zip codes.

We can also profile event attendance consistently. We find that different events of the same type (for example, Shakespeare plays) return a fairly constant number of attendees. We also find that the rank of events by mobile phone attendance matches the rank by estimated head counts. The problem of estimating the actual number of attendees is still open because the ground truth often does not exist or, if it is available, is noisy (that is, based on head counts or aerial photography).

Ranking Events

After associating events to areas of residence, we now need to rank social events in each area. To do so, we produce a score for each event j in area i ($score_{i,j}$) in three ways.

Popular events in the area. The simplest way to produce a score is to assign the most popular events in the area. The score is proportional to the number of users who live in location i and have attended event j (denoted by $m_{i,j}$):

$$\widehat{score}_{i,j} = m_{i,j}. \tag{1}$$

For ranking purposes, what matters is not the score itself but how it compares to another score. The map in Figure 4a, which depicts the most popular event category in every location, shows the output of this approach. By visual inspection, we gather that this approach can identify only popular categories and ignores the remaining categories.

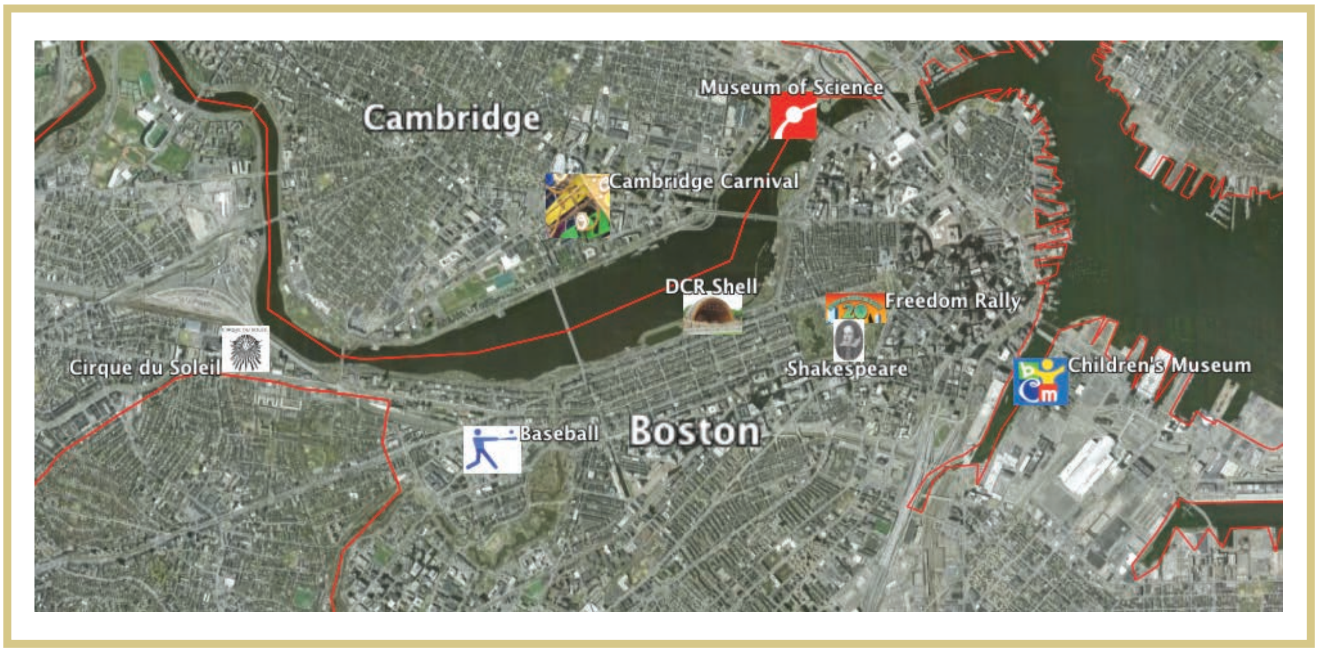


Figure 3. The locations and categories of social events in greater Boston considered in the study: Red Sox baseball games (sport), Cambridge Carnival (festival), Cirque du Soleil Alegria (performance arts), Friday flicks (cinema), Summer concerts (music), Friday nights (cinema), Shakespeare on the Boston Common (performance arts), Freedom Rally (festival); and Target Fridays (educational).

Term frequency-inverse document frequency. To fix the previous problem, we could identify events that are not necessarily popular in general but are popular in the area of residence. Term frequency-inverse document frequency (TF-IDF) is a widely used approach in the information retrieval literature that would do just that.[7] To paraphrase this approach in our context, we assign a higher score to social events that are highly popular in a particular location and that might not necessarily be popular in the remaining locations. The assumption is that the more unique an event is for a location, the more representative the event is for that location. Figure 5 shows the output of TF-IDF. The chart depicts, in each row, the popularity of the six event categories (cinema, family, music, and so on) in a specific location (zip code) with colored bars. Each bar is proportional to the number of residents who have attended events in the corresponding category. One clearly sees that each location has its own predominant category. For example, residents of the suburban area of Belmont (MA 02478) tend to attend family events, whereas residents of the central area of Boston Park Street (MA 02108) tend to attend music events.

TF-IDF is the product of two quantities, TF and IDF. Term frequency (TF) is how often event j has been attended by residents in location i (what we call $m_{i,j}$). We then normalize this count to prevent a bias toward locations whose residents have attended a disproportionate number of social events:

$$tf_{i,j} = \frac{m_{i,j}}{\sum_k m_{i,k}}.$$

However, to find less-attended events, we must be more discriminating. This is the motivation behind inverse document frequency. IDF aims to boost less frequent events. If r is the number of locations, and r_j is the number of locations from which event j is attended, we compute IDF as follows:

$$idf_j = \log\left(\frac{r}{r_j}\right).$$

If attendees at event j come from every location, $r = r_j$ and its IDF is log(1), which is 0.

The problem is that, in our case, IDF will likely be 0. We always find at least one resident in every location who has attended event j. Therefore, we modify the measure in a way similar to what Shane Ahren and his colleagues did.[8] We define the inverse frequency to be the inverse of the number of times event j has been attended:

$$idf_j^{'} = \log \frac{\sum_p \sum_q m_{p,q}}{\sum_q m_{q,j}}.$$

The more popular event j, the lower $idf_j^{'}$.

Finally, TF-IDF is the prediction that users who live in location i will attend event j ($\widehat{score}_{i,j}$):

$$\widehat{score}_{i,j} = tf_{i,j} \times idf_j^{'}. \qquad (2)$$

The idea is that a location has high $\widehat{score}_{i,j}$ for the events attended more often by its residents (high $tf_{i,j}$),

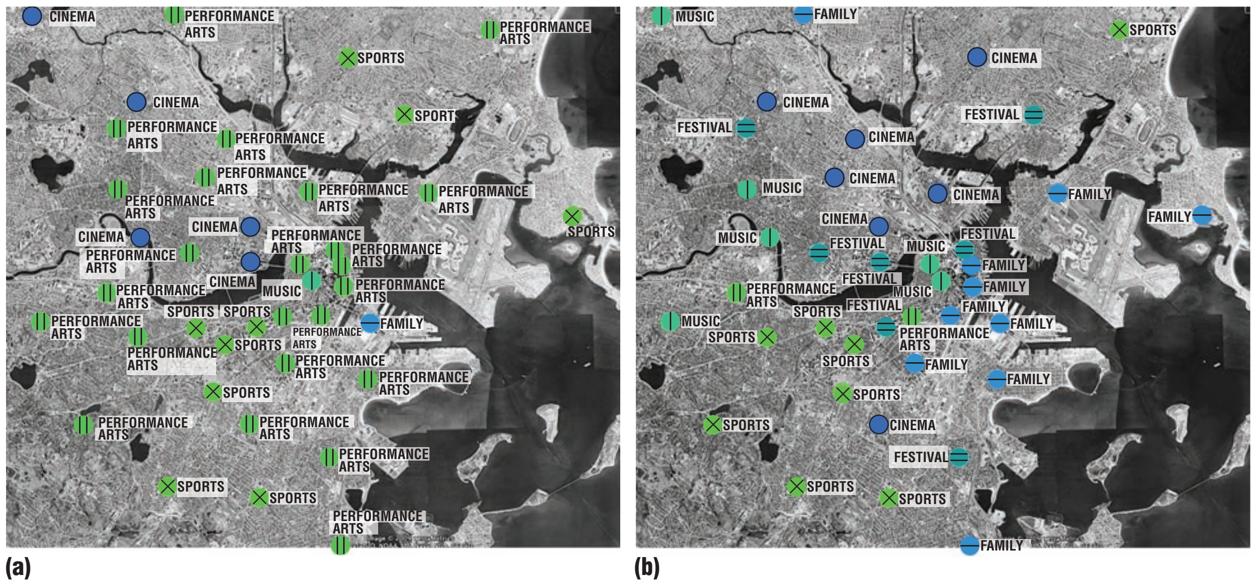


Figure 4. In each location (zip code), we show the event category that (a) is the most popular event in each area, and (b) has the highest term frequency-inverse document frequency (TF-IDF)—that is, the highest popularity in the area of residence.

discounting for those events that virtually everyone attends because they are useless as discriminators (events that have low $idf_j^{'}$). The map in Figure 4b shows the output of the TF-IDF approach. By comparing Figures 4a and 4b, we can see that TF-IDF does not identify only popular events, which suggests that it can find events that better represent the specificity of certain locations.

Eigendecomposition. An alternative way to rank events in an area is to process the scores produced by TF-IDF and determine areas whose residents tend to attend similar social events. To do this, we use an approach similar to Francesco Calabrese and his colleagues'.[9] We arrange TF-IDF scores in one (area × event) matrix and compute the covariance matrix from it. By eigendecomposing the covariance matrix, we obtain a set of eigenvectors. For each area, we determine the set of coefficients that best reconstruct the original TF-IDF scores from the eigenvectors. Those coefficients are a compact way of representing the preferences for social events in each area.

In our case, we identify four main clusters of coefficients (Figure 6a): cluster 1 (blue) reflects a mixture of different social events, cluster 2 (red) is dominated by music events, cluster 3 (magenta) by festivals, and cluster 4 (green) by family events. The clear predominance of certain types of event in most of the clusters suggests that the decomposition resulted in a reasonable representation of preferences and that it is easy to identify the predominant cluster in each area (Figure 6b). Using these clusters that reflect interests of residents in a certain area, we can identify areas with similar interests. This is useful for pricing advertising across areas—for example, similar pricing schemes should be applied in similar areas.

Existing Techniques for Audience Measurements

The cost per thousand impressions is the standard pricing model in advertising media.[10] According to this model, an ad's price depends on one of four criteria:

- the number of users who load the ad on their screens (online),
- the number of viewers exposed to the ad (television),
- the number of publication buyers (newspapers), or
- the number of people estimated to drive and walk by the ad's area (outdoor).

Counting the number of potential viewers of a piece of advertising requires audience measurements, usually validated by third parties. In the case of outdoor advertising, the Traffic Audit Bureau has developed a proprietary measure called Daily Effective Circulation (DEC), which is the estimated number of people who have the opportunity to see a billboard in one day. This is determined from the estimated number of vehicles that travel a specific road segment daily and is standardized using such factors as seasonality, illumination, and passengers per vehicle. Many media planners, however, dismiss DEC as a measure because of the common belief that reported traffic count does not reflect reality and, more worryingly, because the measure has never been validated.[1]

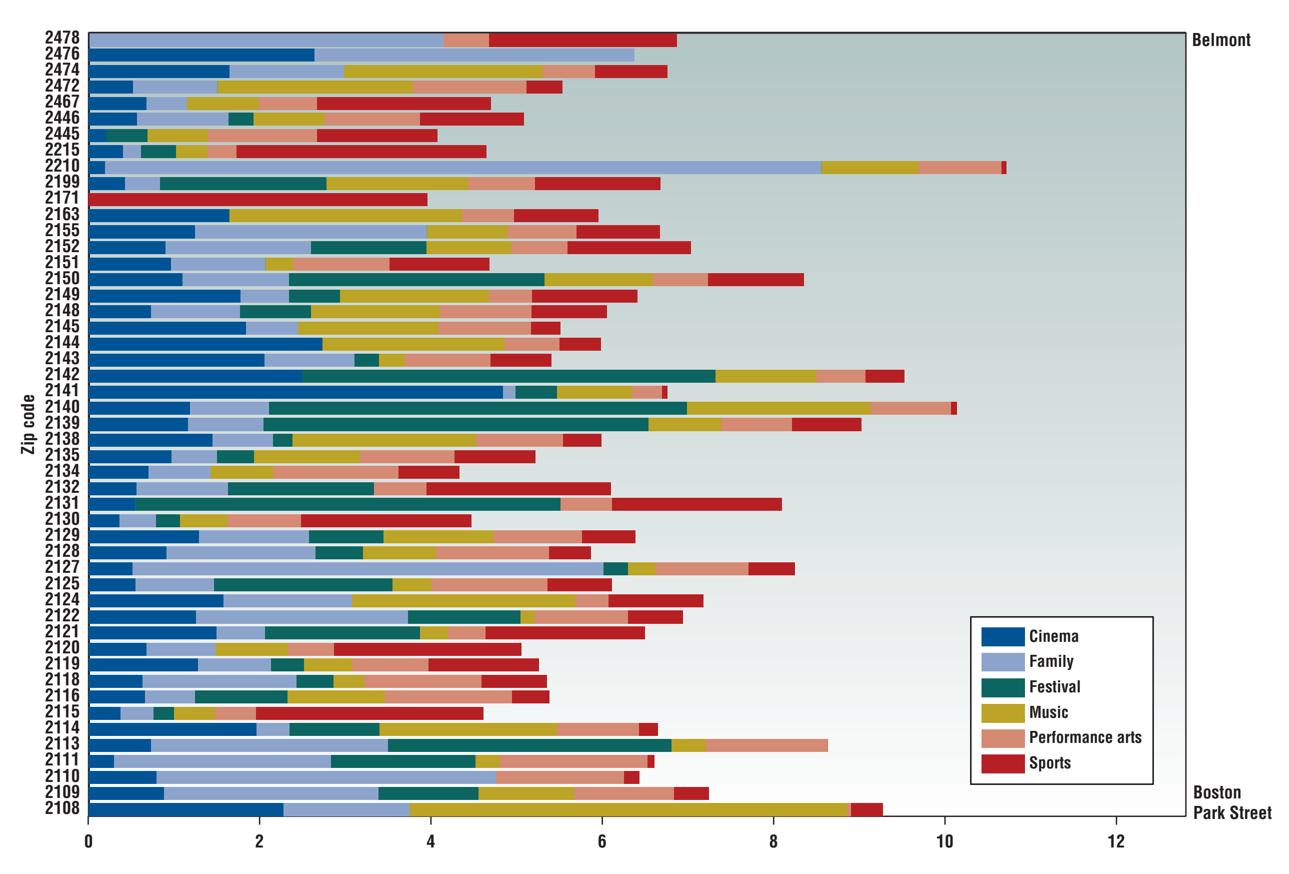


Figure 5. Popularity of event categories across locations. In each row, we consider the popularity in a specific zip code using bar charts. Each bar reflects the fraction of dwellers who have attended events in a certain category.

For this reason, the Traffic Audit Bureau has launched the "Eyes On" project, which aims to define a new audience measure. The idea behind this measure is that it should reflect the number of people who are likely to see an outdoor ad. So far the measure has been computed from demographic and ethnographic data as follows. The Bureau captures and processes high-definition videos upon which it then infers pedestrians' exposures to a specific outdoor ad (that is, it infers how many eyeballs are actually engaged with the ad). The final Eyes On rating integrates eye tracking, circulation, and travel survey data. The rating has been in development for the last five years, including several delayed launches, and its practical applicability is yet to be proven. The rating is destined to remain the same for a long time because updating it every year or so would be expensive.

In academic and industrial labs, researchers have focused on how to personalize advertising mostly on situated displays. Maria Karam and her colleagues built and evaluated the BluScreen architecture, in which individuals store data about their preferences and interests on their mobile phones and transmit the data to nearby situated displays using Bluetooth.[11] The displays can then tailor ads depending on the preferences of who is passing by. Chandra Narayanaswami and his colleagues built a similar system and drew preliminary conclusions about the opportunities and challenges of pervasive advertising.[12] Little work has measured the effectiveness of advertising on bigger displays such as electronic billboards.

Privacy Concerns

One of the concerns work such as ours raises is that of privacy. Although people claim to be concerned about privacy, their actions usually belie these claims. They easily share their pictures on Flickr, status updates on Twitter, and whereabouts on Foursquare with the public at large. However, apparently harmless decisions to share personal information can have unexpected long-term consequences. For example, Pleaserobme.com has been publishing the location of Foursquare's users who are somewhere other than their home. This website aims to make users of location-based services reflect upon whether they are

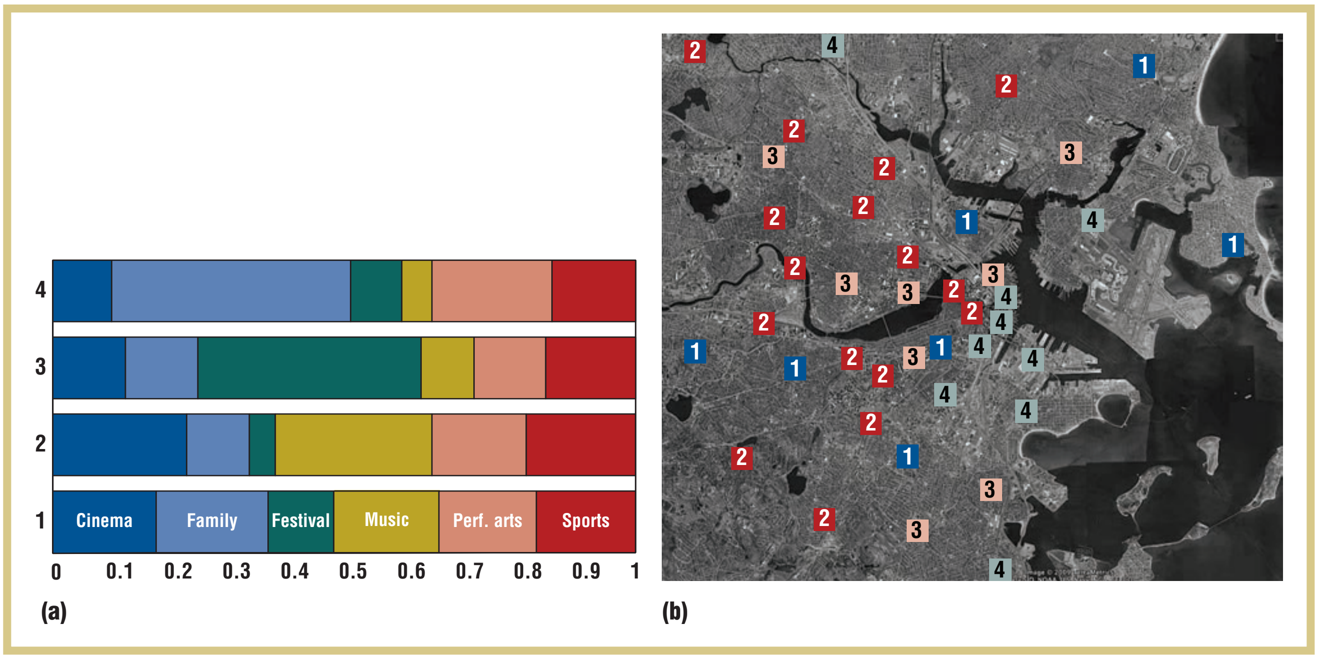


Figure 6. Clusters of preferences for social events generated by the eigendemposition: (a) the eigendecomposition identifies four clusters, and (b) predominant clusters in each area of residence.

giving away information a burglar would love to have and, more generally, whether they are over-sharing. Sharing decisions might be rational in the short term, but people underestimate what might happen to that information if strangers reuse it.

We recently proposed an obfuscation algorithm that can run directly on a mobile phone and lets privacy-conscious users of location-based services report, in addition to their actual locations, some erroneous (fake) locations.[13] A randomized response algorithm selects the erroneous locations in a way that makes it possible to accurately collect and process aggregated location data. Still, a tension exists between the marketer's need for personal information about individuals and the individual's right to privacy. The solution to this tension is to let individuals control ownership of their demographic and behavioral data and determine how and when the data will be used. It would be beneficial to let people signal whether they would like to be associated with the data they place on location-based services, and to be consulted about unusual uses. To this end, Roxana Geambasu and her colleagues recently proposed a system that, by integrating cryptographic techniques with distributed hash tables (DHTs), can make all copies of certain data become unreadable after a user-specified time, even if a person obtains a cached copy of the data.[14]

Measure Effectiveness

The evaluation of our approaches has been qualitative, largely because of lack of ground truth. To fix this problem, we have recently worked on methodologies that perform quantitative evaluations. This work has produced two interesting findings:[3]

- The most effective algorithm recommends events that are popular among residents of an area.
- The least effective algorithm recommends events that are geographically close to the area.

This last result has interesting implications for location-based services that emphasize recommending nearby events. However, in the specific case of outdoor campaigns, to measure the actual effectiveness of such campaigns, we could design measures that exploit the flexibility of new outdoor advertising technologies. For example, placing electronic billboards next to a "point of sale" makes it possible to track the effectiveness of individual outdoor campaigns by simply correlating ads shown at a particular time with point-of-sale data. It would be easy to determine whether specific advertising resulted in an increase in sales.

Our results suggest that mobile phone technologies can produce audience measurements that are more credible than current static measurements. We can reasonably expect that credible audience measurements will make it possible for outdoor advertising to reach its full potential in the future. More generally, we expect that this study will foster future research for three main reasons.

First, our findings are general in that they come from a representative population sample and complement previous work on social networking data. Critics might rightly point out that, in the future, location data will be voluntarily shared by mobile social networking users. However, those users will represent a specific part of the population for a long time, which leads to a self-selection bias. This bias is a major problem in many social sciences and originates from any situation in which individuals select themselves into a group. This causes a biased sample upon which any conclusions drawn might be wrong. By contrast, the market penetration of mobile phones suggests that each individual in the Western world has at least one mobile phone and, consequently, mobile phone users form a representative population sample.[15]

In addition, this study suggests that mining mobile phone data generates new business models. Mobile telecommunication operators could have a two-sided business model in which they would generate revenues not only from their final customers (mobile phone users) but also from upstream customers such as mobile social networking companies and advertising firms. AT&T, Sprint, and Orange have recently started to experiment with this model, and, as a result, they are sharing aggregate mobile data with various research communities.[16]

More importantly, this study goes beyond the mining exercise of recommending social events. The vision behind this research is to unearth human dynamics in the built environment. Our findings on event attendance in a geographical area could be applied to, for example, the field of crowd analysis, which is interested in modeling crowd behavior for predicting the use of space, determining accessibility, and planning emergency evacuations. In addition, sociologists, urban planners, and computer scientists have long been asserting how individuals cluster in geography based on personal interests, but the quantitative proof to strengthen these arguments has not always been available.[15] Being based on quantitative and large-scale human interactions, this work can add a new dimension to their observations and scholarship, with hopes that their work will continue to gain prominence as a rigorous science.

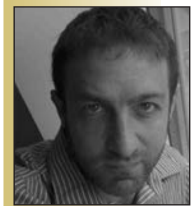
ACKNOWLEDGMENTS

We thank AirSage for making this study possible; the Engineering and Physical Sciences Research Council (EPSRC) Horizon project for its financial support; and Liang Liu and Francisco C. Pereira for their support.

REFERENCES

1. K. Li and A. Edgecliffe-Johnson, "Clear Channel Outdoor Fights Back," *Financial Times*, Mar. 2010.
2. F. Calabrese et al., "The Geography of Taste: Analyzing Cell-Phone Mobility in Social Events," *Pervasive Computing*, LNCS 6030, Springer, 2010, pp. 22–37.
3. D. Quercia et al., "Recommending Social Events from Mobile Phone Location Data," *Proc. IEEE Int'l Conf. Data Mining* (ICDM), IEEE CS Press, 2010, pp. 971–976.
4. F. Calabrese and C. Ratti, "Real Time Rome," *Networks and Communications Studies*, vol. 20, no. 3–4, 2006, pp. 247–258.
5. F. Calabrese et al., "A System for Real-Time Monitoring of Urban Mobility: A Case Study in Rome," *IEEE Trans. Intelligent Transportation Systems*, 2010, pp. 1–11.
6. M.C. Gonzalez, C.A. Hidalgo, and A.-L. Barabasi, "Understanding Individual Human Mobility Patterns," *Nature*, vol. 453, 5 June 2008, pp. 779–782.
7. R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison Wesley, 1999.
8. S. Ahern et al., "World Explorer: Visualizing Aggregate Data from Unstructured Text in Geo-Referenced Collections,"

the AUTHORS



Daniele Quercia is a Horizon Researcher in the Computer Laboratory at the University of Cambridge. His research interests include computational social science and social computing. Before joining the University of Cambridge, he was a postdoctoral associate at the Massachusetts Institute of Technology. Quercia has a PhD in computer science from University College London. He is a member of the Ordine degli Ingegneri di Torino and a senior member of Wolfson College in Cambridge. Contact him at daniele.quercia@cl.cam.ac.uk.

Giusy Di Lorenzo is a research scientist at the IBM Dublin Research Laboratory, Dublin, Ireland, and a research affiliate at the SENSEable City Laboratory at MIT. Her research interests include the analysis of urban dynamics and human mobility using data gathered from sensor networks. In particular, she is interested in developing context-aware applications to improve citizens' urban living experiences. Di Lorenzo has a PhD in computer and system engineering from the University of Naples Federico II. Contact her at giusydil@ie.ibm.com.

Francesco Calabrese is an advisory research staff member at the IBM Dublin Research Laboratory, Dublin, Ireland, and a research affiliate at the SENSEable City Laboratory at MIT. His research interests include ubiquitous computing; analysis of urban dynamics through sensor networks; and analysis and design of distributed, hybrid, embedded control systems, and computer numerically controlled machines. Calabrese has a PhD in computer and system engineering from the University of Naples Federico II. He's a member of IEEE and the IEEE Control Systems Society. Contact him at fcalabre@ie.ibm.com.

Carlo Ratti is an architect and engineer who practices architecture in Turin and teaches at MIT, where he directs the SENSEable City Laboratory. His research interests include urban design, human-computer interfaces, electronic media, and the design of public spaces. Ratti has a PhD in architecture from the University of Cambridge. He is a member of the Ordine degli Ingegneri di Torino, the Architects Registration Board (UK), and the Association des Anciens Elèves de l'École Nationale des Ponts et Chaussées. Contact him at ratti@media.mit.edu.

Proc. ACM Conf. Digital Libraries (JCDL), ACM Press, 2007, pp. 1–10.

9. F. Calabrese, J. Reades, and C. Ratti, "Eigenplaces: Segmenting Space Through Digital Signatures," *IEEE Pervasive Computing*, vol. 9, no. 1, 2010, pp. 78–84.
10. D.M. Chickering and D. Heckerman, "Targeted Advertising on the Web with Inventory Management," *Interfaces*, vol. 33, no. 5, 2003, pp. 71–77.
11. M. Karam, T. Payne, and E. David, "Evaluating Bluscreen: Usability for Intelligent Pervasive Displays," *Proc. Int'l Conf. Pervasive Computing and Applications* (ICPCA), IEEE CS Press, 2007, pp. 18–23.
12. C. Narayanaswami et al., "Pervasive Symbiotic Advertising," *Proc. ACM Hot-Mobile*, ACM Press, 2008, pp. 80–85.
13. D. Quercia et al., "SpotME If You Can: Randomized Responses for Location Obfuscation on Mobile Phones," *Proc. 31st Int'l Conf. Distributed Computing Systems* (ICDCS), IEEE CS Press, 2011; www.cl.cam.ac.uk/~dq209/publications/spotme.pdf.
14. R. Geambasu et al., "Vanish: Increasing Data Privacy with Self-Destructing Data," *Proc. 18th Usenix Security Symp.*, Usenix Assoc., 2009; http://vanish.cs.washington.edu/pubs/usenixsec09-geambasu.pdf.
15. N. Eagle, "Behavioral Inference Across Cultures: Using Telephones as a Cultural Lens," *IEEE Intelligent Systems*, vol. 23, no. 4, 2008, pp. 62–64.
16. T. Simonite, "Mobile Data: A Gold Mine for Telcos," *MIT Technology Rev.*, 27 May 2010; www.technologyreview.com/communications/25396/?a=f.