
senseable city lab:.::

Mobile Phones and Outdoor Advertising: Measurable Advertising

Abstract

Television and newspapers sit at the top of many agency marketing plans, while outdoor advertising stays at the bottom. The reason for this is that it's difficult to account for who views a billboard, so there is no way of consistently determining the effectiveness of outdoor advertising. As a result, agencies do not consider the medium and allocate their money elsewhere. To change this situation, one needs to create new credible audience measurements for the outdoor marketing industry. Here we propose a new way of performing audience measurements that combines mobile phone location estimations with information freely available on the Internet. We show that it is possible to estimate the number of people who drive or walk by a given area in Greater Boston from location estimations of a large fraction of mobile phone users in the region. We also infer the preferences for social events of the users by combing their location estimations with Internet listings of social events. This makes it possible to profile areas based on their residents' interests and dynamically change displayed advertising based on those assessments.

I. Introduction

Online advertising is the fastest growing advertising medium, not least because it is able not only to track how long someone was on a Web page that showed an advertisement (ad) but is also able to track how many times someone clicked on that ad. By contrast, outdoor advertising has not reached its full potential because of its inability to measure "return on investment". People spend 27 per cent of their time exposed to outdoor advertising, but such form of advertising attracted only 5 per cent of US media spending of 2008 [1]. As we shall see in Section IV, 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 one expects, unvalidated measurements do not call for marketing dollars.

The industry would embrace outdoor advertising only if credible audience measurements would be introduced. We suggest that one way of performing audience measurements that is more credible than existing ways is to estimate: 1) the number of people in front of a billboard; and 2) the likelihood that those people might like a specific ad shown on the billboard. We estimate the number of people based on location estimations coming from mobile phones near the billboard, and we infer people preferences by combining the location estimations with information freely available on the Internet. More specifically, we consider electronic billboards displaying ads about social events (e.g., football game, music festival) in Greater Boston and make two main contributions:

1. We estimate the number of people near a billboard area by inferring how many mobile phone users pass by the area (Section II).

2. We infer mobile phone users' preferences for social events by inferring which social events they have likely attended (Section III).

II. Audience Measurement from Location Estimations of Mobile Phones

The first step is to determine the number of people near a billboard. We do so in three steps (see [2], [3] for details):

1) *Collect location estimates of mobile phones.* From *Airsage Inc.*, we collected estimates of the locations of 1 million mobile phone users in the Greater Boston area (20% of the entire population). The logs span one and half months of the summer of 2009, and are generated every time a mobile phone connects to the cellular network (that is, whenever the phone places/receives a call, sends/receives a text message, or is on the Internet). Our dataset contained 130 million pairs of latitude and longitude estimates with corresponding timestamps. 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_i} that is expressed in latitude and longitude and is associated with a timestamp t_{m_i}. Mobility trajectories (such as those shown in Figure 1) are then formed by linking location estimates together to form temporal sequences of the form $\{m_1, m_2, \ldots, m_n\}$. Since location estimates 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 inter-event time (time between two consecutive network connections) for each user. Figure 3 shows the distribution of the first and third quartiles and the median of inter-event time for the whole population. The arithmetic average of the medians is 84 minutes and the geometric average is 10 minutes. This means that we are able to 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 sample 80,000 users.

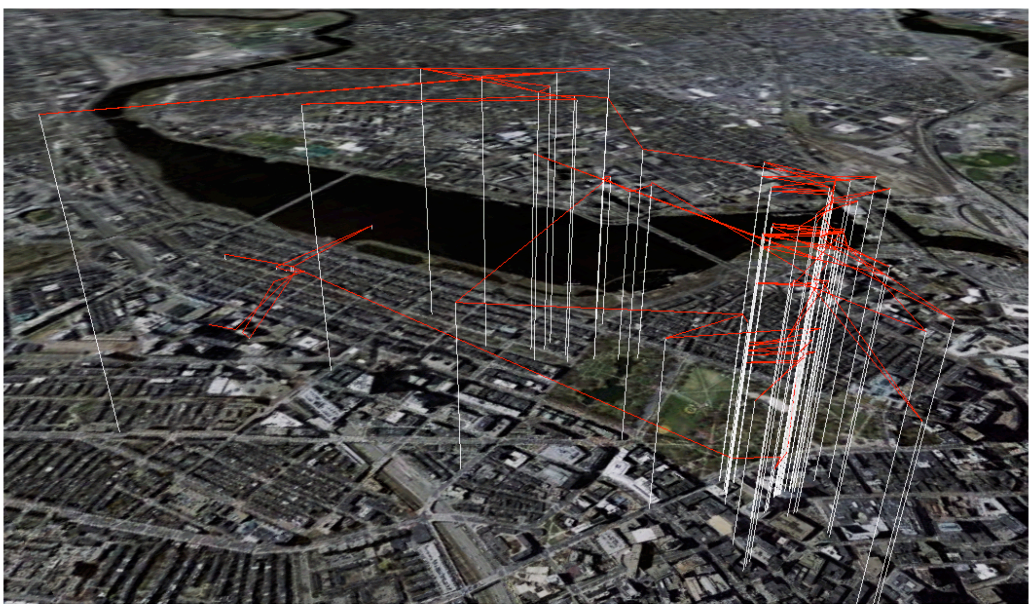


Fig. 1. Mobility trajectories extracted from mobile phone data.

2) *Generate mobility trajectories from location estimates.* The result of the previous step is that, for each user, we have a set of visited locations. To make sense of those locations, we extract individuals trips (trajectories) from them. Going 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:

- Identifying the user's trajectories: If two subsequent location estimates are registered within two hours from each other, they are considered part of the same trajectory; if they are registered more than two hours apart, then they form two separate trajectories. The choice of two hours comes from our temporal resolution of 1.5 hours (being the inter-event time less than 1.5 hours (Figure 3)).
- Removing noise from each trajectory: We apply a low-pass filter on each trajectory's location estimates with a resampling rate of 10 minutes ([6], [7]).
- Characterizing each trajectory as a set of consecutive stops: We define a stop to be the centroid

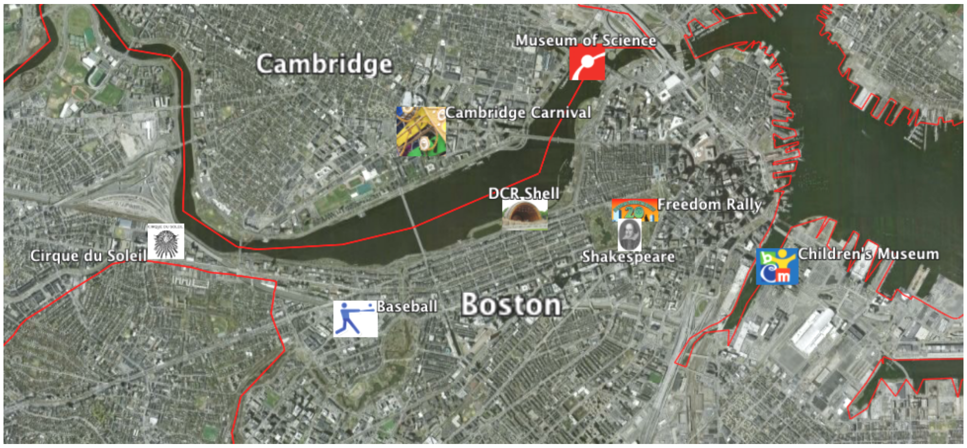


Fig. 2. On the map of Greater Boston, the social events under study are shown: Red Sox baseball games (in the category *Sport*); Cambridge Carnival (*Festival*); Cirque du Soleil Alegria (*Performance Arts*); Friday flicks (*Cinema*); Summer Concerts (*Music*); Friday nights (*Cinema*); Shakespeare on the Boston Common (*Performance Arts*); Freedom Rally (*Festival*); and Target Fridays (*Educational*).

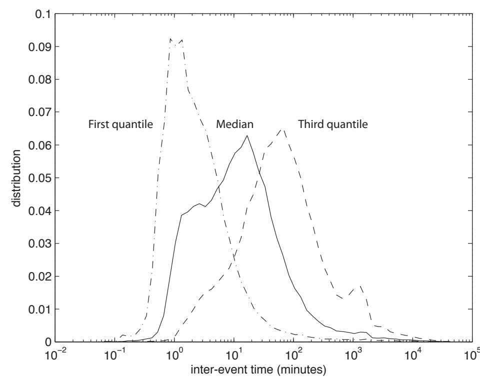


Fig. 3. The distribution of inter-event time (time between two consecutive cellular network connections) for our mobile phone users. Median (solid line), first quantile (dash-dotted line) and third quantile (dashed line).

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 the state of the art [8], our geographical resolution is not down to the area that is typically covered by a cellular tower but is as little as 350 meters, and that might be because of two reasons: a) We consider a larger number of location estimations - we consider not only those registered when making a call (as previously done) but also those registered when sending/receiving a text message or surfing the Internet; and b) In addition to cell tower IDs, we also use the triangulation values generated by the *Airsage Inc.*'s "Wireless Signal Extraction" technology.

3) *Count the number of mobile phone users in one area.* After this data processing, we are able to infer the number of mobile phone users in a certain area with a temporal resolution of 1.5 hour and geographical resolution of 350 meters.

III. Ranking social events in each area of residence

After determining the number of people in one area, one needs to determine those people's preferences for social events. To this end, in Section III-A, we determine the social events that have been likely attended by residents of an area (by people who make the area the most frequent stop at night between

10pm and 7am). Then, in each area, we characterize residents' preferences based on ranking events in three different ways (Section III-B).

A. Attendance at Social Events

We crawled the "Boston Globe Calendar" website[1] 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. The events are organized in 14 categories: Arts & Crafts, Business & Tech, Community, Dance, Education/Campus, Fairs & Festivals, Food & Dining, Music, Other, Performing Arts, Shopping, Sports & Outdoors, Visual Arts, and Cinema.

To make use of the events, we divided the $15km^2$ area of Greater Boston in geographic cells of 500 x $500m$. The choice of $500m$ is related to our average localization error, $350m$, as mentioned above. We then place the events and user stops in the corresponding cells. By intersecting stops and social events across cells, we are able to determine which social events have been attended by whom. However, as one might expect, we may produce false positives - some people may mistakenly be seen to attend social events they have never been to.

To decrease false positives, we checked the following assumptions: 1) the user stops in the same cell of the event location; and 2) we consider only people who live in a different location from the event location and stay for at least 70% of an event duration. We do not opt for full 100% overlap because this would require a person to make one call right before the event and another call right after it. More specifically, we find the largest set of events that satisfy five requirements. Each event:

1) Is highly attended so that the likelihood of being attended by our mobile users is significant.
2) Is geographically isolated compared to neighboring events. We impose a minimum radius of one kilometer between the event and any other large concurrent event.
3) Takes place in a well-defined geographic area of considerable dimension. This will minimize the possibility of mistakenly considering people staying in places close to the event's venue (e.g., in a nearby restaurant).
4) Is temporally isolated from any other large event.
5) 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 makes sure that we will work upon correct inferences with high probability. Concretely, to reduce false positives, we pay the price of reducing the number of social events from 500 a day to 53 for the whole period of study (some of which are shown in Figure 2), and the number of mobile users from 80,000 to 2,519. However, on the positive side, we are able to:

- Capture a large fraction of the population in a meaningful way. That is because our mobile phone users represent 20% of the entire population in Greater Boston (as per latest US census) and are proportionally distributed across zipcodes.
- Profile attendance at events in a consistent way. We find that different events of the same type (e.g., Shakespeare plays) return a fairly constant count of number of attendees. We also find that the rank of events by "mobile phone" attendance match the rank by estimated head counts. The problem of estimating the actual number of attendees is still open since the ground truth often does not exist or, if available, it is noisy (it is based on head counts or aerial photography).

Importantly, this conservative way of filtering drastically reduces our dataset, but it also makes sure that we will work upon correct inferences with high probability. Concretely, to reduce false positives, we pay the price of reducing the number of social events from 500 a day to 58 for the whole period of study, and the number of mobile users from 80,000 million to 2,519.

[1]http://calendar.boston.com/

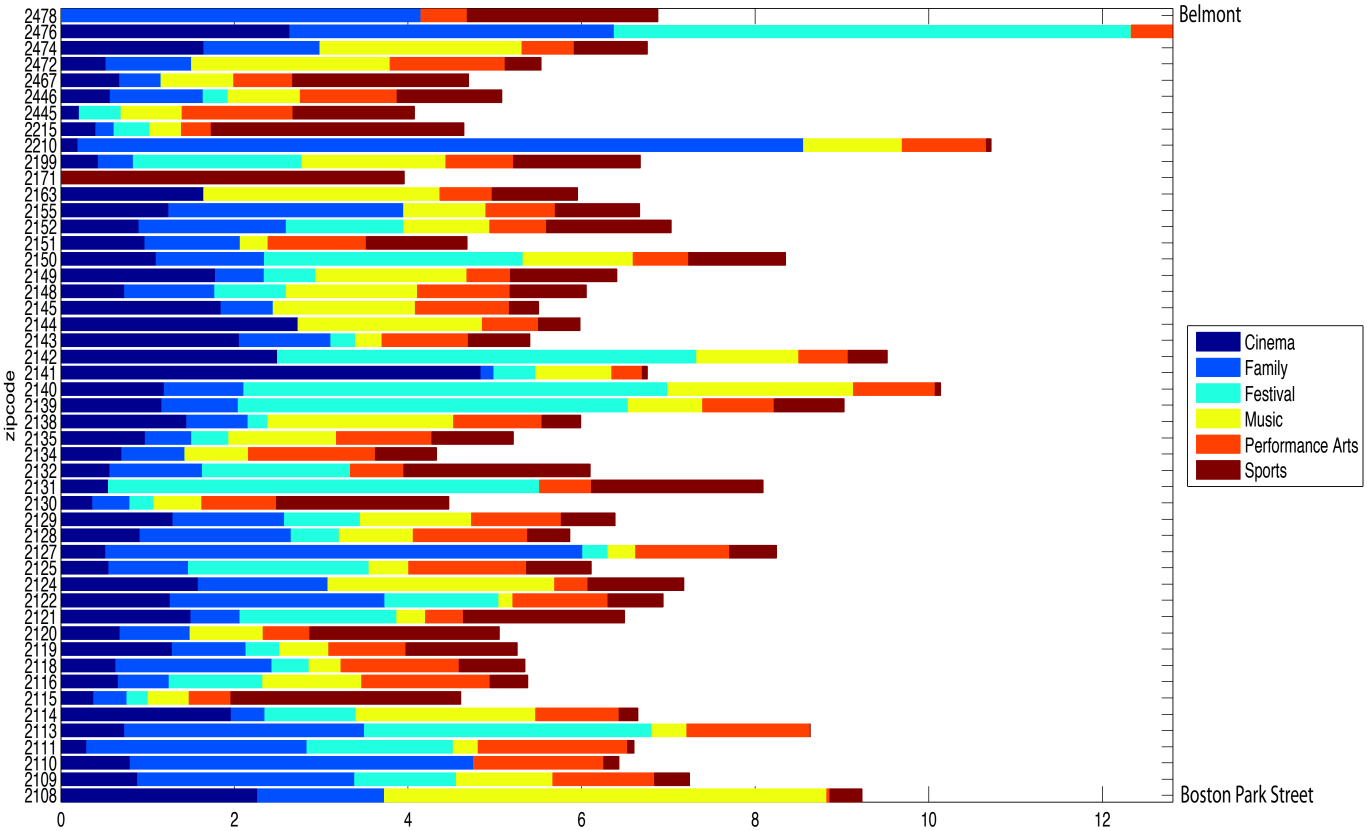


Fig. 4. Popularity of event categories across locations. In each row, we consider the popularity in a specific zipcode using bar charts. Each bar reflects the fraction of dwellers who have attended events of a certain category.

B. *Ranking Events*

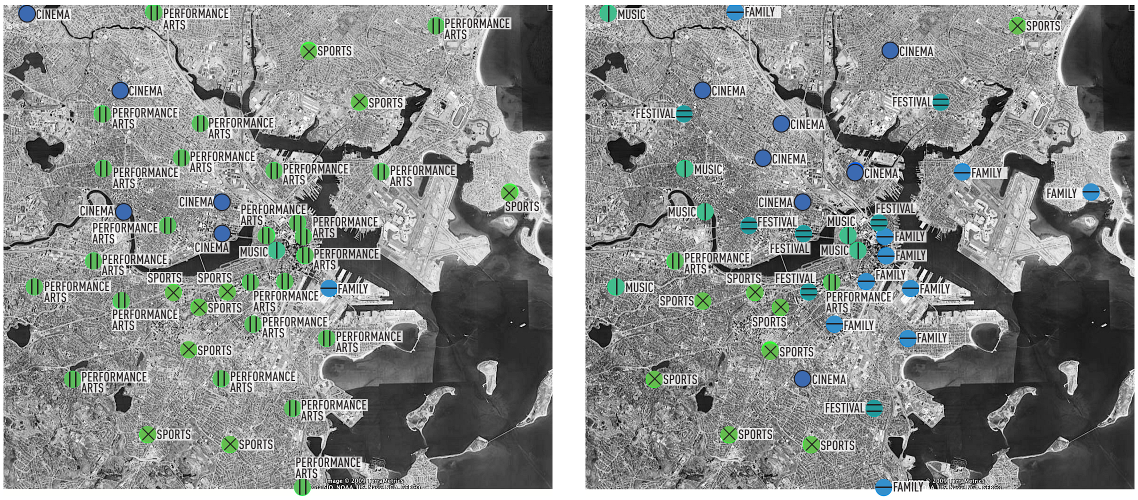
We now need to rank social events in each area of residence. To do so, we produce a score for each event j in area i ($score_{i,j}$) in three ways:

(1) Popular events in area. The simplest way 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 a score itself but how the score compares to another. The output of this approach is shown on the map of Figure 5(a), which depicts the most popular event category in every location. By visual inspection, we gather that this approach is able to identify only popular categories and ignores the remaining categories.

(2) *Term Frequency-Inverse Document Frequency.* To fix the previous problem, one could identify events that are not necessarily popular in general but are popular in the area of residence. *TF-IDF* (Term Frequency-Inverse Document Frequency) is a widely-used approach in the literature of Information Retrieval [9] that would do just that. 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 may not be necessarily 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. The output of *TF-IDF* is shown in the chart of Figure 4. The chart depicts, in each row, the popularity of the six event categories (e.g., cinema, family, music) in one specific location (zipcode) with colored bars. Each bar is proportional to the number of residents



(a) (b)

Fig. 5. In each location (zipcode), we show the event category that: (a) is the most popular one in each area; (b) have highest *TF-IDF* (highest popularity in the area of residence).

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, while residents of the central area of Boston Park Street (“`MA 02108`”’) tend to attend music events.

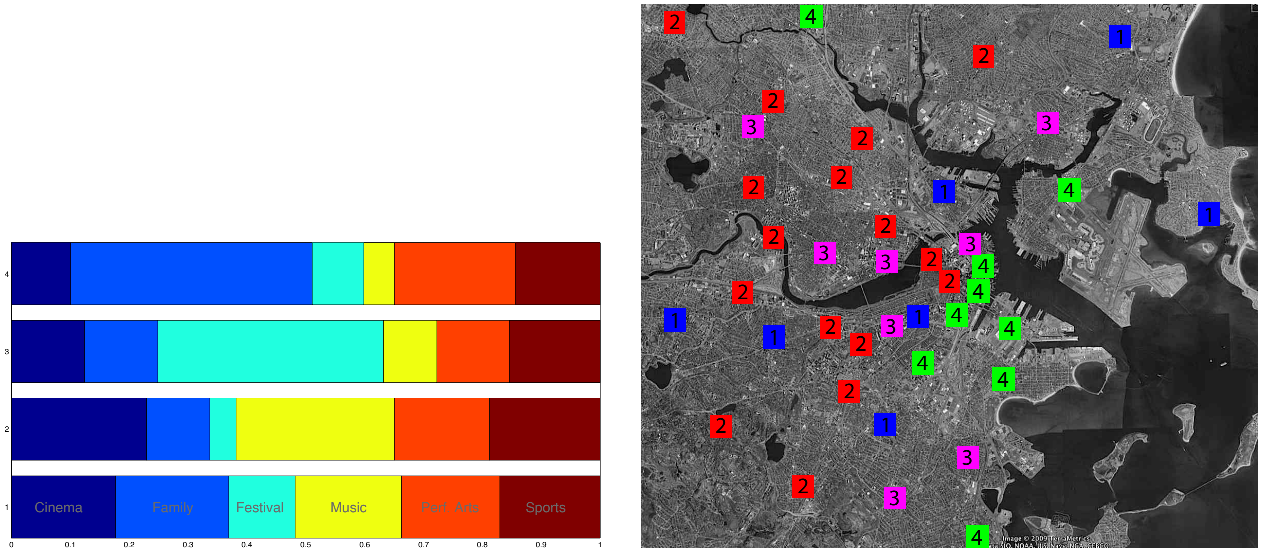
TF-IDF is the product of two quantities TF and IDF. The *term frequency* (*TF*) is a count of how often event j has been attended by residents in location i (it is what we called $m_{i,j}$). This count is then normalized to prevent a bias towards locations whose residents have attended a disproportionate number of social events: $tf_{i,j} = \frac{m_{i,j}}{\sum_k m_{i,k}}$. However, to find less-attended events, we have to be more discriminating. This is the motivation behind inverse document frequency. *IDF* aims to boost events that are less frequent. If r is the number of locations, and r_j is the number of locations from which event j is attended, then *IDF* is computed as follows: $idf_j = log(\frac{r}{r_j})$. If attendees at event j come from every location, then $r = r_j$ and its *IDF* is $log(1)$ which is 0. The problem is that, in our case, *IDF* is likely to 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 Ahren *et al.*’s did [10]. 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 ($\widehat{score}_{i,j}$):

$$\widehat{score}_{i,j} = tf_{i,j} \times idf'_j \tag{2}$$

The idea is that a location has high $\widehat{score}_{i,j}$ for the events attended more often by its residents (high $tf_{i,j}$), discounting for those events that are attended virtually by everybody since they are useless as discriminators (events that have low idf'_j). The output of the *TF-IDF* approach is shown on the map of Figure 5(b). By comparing Figures 5(a) and 5(b), one sees that *TF-IDF* does not identify only popular events, which suggests that *TF-IDF* is able to find events that better represent the specificity of certain locations.

(3) *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



(a) The eigendecomposition identifies four clusters.

(b) Predominant clusters in each area of residence.

Fig. 6. Clusters of preferences for social events generated by the eigendemposition.

approach similar to Calabrese *et al.*'s [11]. We arrange *TF-IDF* scores in one $(\texttt{area} \times \texttt{event})$ 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 6(a)): 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 6(b)). Upon those clusters, one is able to identify areas with similar interests where similar pricing schemes should be applied.

IV. Existing Techniques for Audience Measurements

The cost per thousand impressions is the standard pricing model throughout all advertising media [12]. According to this model, the price of an ad depends on: 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).

To count the number of potential viewers of a piece of advertising, one needs audience measurements. Those measurements are 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 persons 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, passengers per vehicle. Many media planners, however, dismiss the measure of DEC because it is often believed that reported traffic count does not reflect reality and, more worryingly, the measure has never been validated [1].

For this reason, the Traffic Audit Bureau has launched a project named "Eyes On" whose goal is 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 (i.e., 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 proved - the rating is destined to remain the same for a long time since updating it every year or so would be expensive.

In the academic and industrial labs, researchers have focused on how to personalize advertising mostly on situated displays. Karam and colleagues built and evaluated an architecture called BlueScreen in which individuals store data about their preferences and interests on their mobile phones and transmit the data to nearby situated displays using Bluetooth [13]. The displays are then able to tailor ads depending on the preferences of who is passing by. Narayanaswami and colleagues built a similar system and drew preliminary conclusions about the opportunities and challenges of pervasive advertising [14]. There has been little work on measuring the effectiveness of advertising on bigger displays such as electronic billboards.

V. Discussion

Privacy Concerns. While people claim to be concerned about privacy, their actions usually belie these claims. They are pleased to share pictures on Flickr.com, status updates on Twitter.com, and whereabouts on FourSquare.com with the public at large. However, apparently harmless decisions of sharing personal information may have unexpected consequences in the long term. For example, Pleaserobme.com has been publishing the location of Foursquare's users who are somewhere other than their home. The aim of this website is to make users of location-based services reflect upon whether they are 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 they underestimate what might happen to that information as it is reused by strangers. One way of addressing this concern is to avoid using personal information of single individuals and, instead, use aggregate information. In this vein, users of our systems share only their home location, and our algorithms process whereabouts that are aggregated at the level of area of residence, making it difficult to drill down to anyone's individual data. An additional form of data aggregation is the generalization of movement data proposed by Andrienko *et al.* [15], which guarantees k-anonymity when making publicly available spatio-temporal datasets. We have recently proposed an obfuscation algorithm that can run directly on a mobile phone and allows privacy-conscious users of location-based services to report, in addition to their actual locations, also some erroneous ("fake") locations [16]. The erroneous locations are selected by a randomized response algorithm in a way that makes it possible to accurately collect and process aggregated location data. Still, there is a tension 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 allow people to 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, Geambasu *et al.* [17] recently proposed a system that, by integrating cryptographic techniques with distributed hash tables (DHTs), is able to make all copies of certain data become unreadable after a user-specified time, even if one obtains a cached copy of the data.

Measure Effectiveness. The evaluation of our approaches has been qualitative, largely because of lack of ground truth. To fix this problem, we are currently working on methodologies that perform quantitative evaluations. For example, we have recently tested different algorithms for recommending social events from our mobile phone data and produced two interesting findings [3]: 1) The most effective algorithm recommends events that are popular among residents of an area; 2) The least effective, instead, 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, one 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 - one could easily determine whether a specific advertising resulted in an increase in sales.

VI. Conclusion

From location estimations of a large number of mobile phone users in Greater Boston, we are able to: (1) estimate the number of mobile phone users within a small geographic area; and (2) determine the preferences for social events in each area of residence. Those two results suggest the ability of mobile phone technologies to produce audience measurements that are more credible than static measurements currently used by the industry. One would 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:

1. The findings are general in that they come from a representative population sample and are complementary to previous work on social-networking data. Critics may rightly point out that, in the future, location data will be voluntarily shared by mobile-social networking users. The problem is those users are going to represent a specific part of the population for a long time, and this 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 may 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 [18].

2. This study suggests how mining mobile phone data generates new business models. Mobile telecommunication operators could have a 2-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. The operators AT&T, Sprint, and Orange have recently started to experiment with this model and, as a result, they are sharing aggregate mobile data with a variety of research communities [19].

3. 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 the behavior of crowds for predicting the use of space, planning accessibility, and planning emergency evacuations. Also, sociologists, urban planners, and computer scientists have long been making assertions about how individuals “cluster” in geography based on personal interests, but the quantitative proof to strengthen these arguments has not always been available [18]. Being based on quantitative and large-scale human interactions, this work may add a new dimension to their observations and scholarship, with hopes that their work will continue to gain prominence as a rigorous science.

Acknowledgements: We thank: the company *Airsage Inc.* for making this study possible; the projects EPSRC Horizon 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*, 2010.

[2] F. Calabrese, F. Pereira, G. D. Lorenzo, L. Liu, and C. Ratti, “Recommending Social Events from Mobile Phone Location Data,” in *Pervasive*, 2010.

[3] D. Quercia, N. Lathia, F. Calabrese, G. D. Lorenzo, and J. Crowcroft, “Recommending Social Events from Mobile Phone Location Data,” in *IEEE ICDM*, 2010.

[4] J. Krumm, “Real Time Destination Prediction Based on Efficient Routes,” in *Society of Automotive Engineers (SAE) 2006 World Congress*, 2006.
[5] J. Krumm and E. Horvitz, “Predestination: Inferring Destinations from Partial Trajectories,” in *UbiComp*, 2006, pp. 243–260.
[6] F. Calabrese and C. Ratti, “Real Time Rome,” *Networks and Communications Studies*, vol. 20, no. 3–4, pp. 247–258, 2006.
[7] F. Calabrese, C. Ratti, M. Colonna, P. Lovisolo, and D. Parata, “A System for Real-Time Monitoring of Urban Mobility: a Case Study in Rome,” *IEEE Transactions on Intelligent Transportation Systems*, 2010.
[8] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, “Understanding individual human mobility patterns,” *Nature*, 2008.
[9] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Addison Wesley, 1999.
[10] S. Ahern, M. Naaman, R. Nair, and J. H. Yang, “World explorer: visualizing aggregate data from unstructured text in geo-referenced collections,” in *Proceedings of ACM Conference on Digital Libraries (JCDL)*, New York, NY, 2007.
[11] F. Calabrese, J. Reades, and C. Ratti, “Eigenplaces: segmenting space through digital signatures,” *IEEE Pervasive Computing*, 2010.
[12] D. M. Chickering and D. Heckerman, “Targeted advertising on the web with inventory management,” *Interfaces*, vol. 33, no. 5, 2003.
[13] M. Karam, T. Payne, and E. David, “Evaluating bluscreen: Usability for intelligent pervasive displays,” in *ICPCA*, 2007.
[14] Narayanaswami and *al.*, “Pervasive symbiotic advertising,” in *ACM HotMobile*, 2008.
[15] G. Andrienko, N. Andrienko, F. Giannotti, A. Monreale, and D. Pedreschi, “Movement data anonymity through generalization,” in *SIGSPATIAL ACM GIS Workshop on Security and Privacy*, 2009.
[16] D. Quercia, I. Leontiadis, L. McNamara, C. Mascolo, and J. Crowcroft, “SpotME If You Can: Randomized Responses for Location Obfuscation on Mobile Phones,” in *Technical Report of The University of Cambridge*, 2010.
[17] R. Geambasu, T. Kohno, A. Levy, and H. M. Levy, “Vanish: Increasing data privacy with self-destructing data,” in *USENIX Security Symposium*, 2009.
[18] N. Eagle, “Behavioral inference across cultures: Using telephones as a cultural lens,” *IEEE Intelligent Systems*, 2008.
[19] T. Simonite, “Mobile data: A gold mine for telcos,” *MIT Technology Review*, 2010.