The Workshops of the Tenth International AAAI Conference on Web and Social Media
CityLab: Technical Report WS-16-16

This Is Not Just a Café: Toward Capturing the Dynamics of Urban Places

Myeong Lee[1], Rosta Farzan[2], Brian Butler[1]
[1]University of Maryland, College Park, MD, 20742
[2]University of Pittsburgh, Pittsburgh, PA, 15260
{myeong, bsbutler}@umd.edu, rfarzan@pitt.edu

Abstract

Social media has provided a huge amount of user-generated data in capturing urban dynamics. Among them, place-level human behavior has been largely detected through people's check-in records at certain places. Conventionally, places are characterized by a set of pre-defined features, often specified by the owner of the places. In this paper, we argue that capturing socially-meaningful features and dynamics of an urban place may also be done by analyzing human activity traces. We explore the activity-centered representation of urban places as a foundation for understanding local neighborhoods at scale. We analyze activities from several venues using data from MeetUp, a popular local event organizing service. Then, Yelp's business information was used to triangulate the analysis. The results suggest that strategies based on local event data have great potential for modeling socially-meaningful urban places at scale.

Introduction

A byproduct of advances in mobile and social technologies has been the possibility of identifying the dynamics of cities, such as human mobility, socio-economic status, and land use, through geo-tagged data generated by individual citizens. Data such as call detail records (CDR), geo-tagged Tweets, and Flickr photos have been valuable sources for detecting urban dynamics. These geo-tagged data, however, provide little knowledge about the nature and features of urban places where people meet, discuss their daily lives, and keep diverse activities. Fortunately, the emergence of venue-based social media has the potential to support understanding of urban places at scale. Yet, foundational social science research suggests that current computational approaches would benefit from a more nuanced consideration of the concept of place.

Gieryn, a sociologist, pointed out that for a place to be identified as a "place", it should be identifiable or represented by *people* (Gieryn 2000). Harrison and Tatar developed this idea and re-conceptualized "place" in the context of computing systems as being created at the confluence of loci, people, and events (Harrison and Tatar 2007). Following these concepts, a place can be understood as a socially constructed form with people and their activities in addition to spatial features. Recent studies using location-based social media such as Foursquare and Yelp suggested that data about human activities in specific venues provides great potential for modeling places and associated human behaviors, supporting understandings of urban dynamics. If we can take the concept of place into account, it would be possible to construct an urban place model with rich contents about venues.

The contents of people's activities including the time and location are at the core of the concept of the place. In this paper, we explore whether local event data can be utilized to model such concept of place. Local events such as concerts, farmer's market, and board games represent a wide range of activities happening in urban places. At the same time, local event data from event-based social media such as MeetUp, Eventbrite, and Facebook Event provide information about the venue, time, and the description of activity associated with each event. In this work, we utilize MeetUp, as one of the most popular local event organization platform, to explore the value of event data in modeling places. We further triangulated our analysis by employing Yelp to gather conventional characteristics of places defined by the businesses. Yelp data allowed us to contrast descriptions of places derived from activity traces with conventional descriptions of the same places.

Through a descriptive approach, we aim to understand whether (1) places can be quantified using description of activities happening in each place and (2) local event data are appropriate in modeling places and activities. This work can benefit urban computing researchers by providing place models that are grounded on the theoretical concept of place. Also, urban planners and policy makers can

(a) (b)

Figure 1. An example of a franchise business in different locations.
Crazy Mocha (a) Squirrel Hill branch, (b) Friendship branch (photos from Google Maps and Yelp.com)

consider local events as a means to understand how places are constructed through diverse human meet-ups.

In the following sections, we first review some related work. Then, several target places are chosen to be explored. The activities in these places are quantitatively presented and two representative cases are analyzed qualitatively. Lastly, we discuss limitations and implications of the study.

Related Work

In a recent work, Cranshaw et al, modeled an urban place as a vector of user check-ins for each venue in Foursquare. The model was used to calculate the social affinity between places based on the quantified vectors (Cranshaw et al. 2012). The place model was systematically used in capturing dynamically changing neighborhood in the city of Pittsburgh. A subsequent study extended the social affinity measure by extending the vector to temporal space, and used the construct in clustering analysis to re-draw activity-based neighborhoods (Rösler and Liebig 2013). Le Falher et al. employed additional semantics from Foursquare data in modeling vectors for venues and cities (Le Falher, Gionis and Mathioudakils 2015). They compared six cities with different neighborhoods, which were dynamically identified using the computational models. Similarly, social distance was quantified by constructing an activity model that consists of coordinates, Foursquare's business categories, and Facebook users' interests (Del Bimbo et al. 2014).

These studies were very well designed and quantified with a large amount of geo-tagged data, and made use of the contents of places and activities to some degree. However, due to the nature of Foursquare data, the amount of information for people's activities were limited. If local event data prove to be useful for quantification, they would provide richer contents about human activities that shape urban places.

Datasets and Methods

MeetUp event data were collected for six months from August 2015 to January 2016 for the city of Pittsburgh. Totally, there were 4361 events organized in 1168 venues. The venues on MeetUp were mostly recorded by users, and there were often many events organized in the same physical location with different venue IDs. That is, the number of unique venues that events were held was much less than 1168 due to the multiple venue IDs for a same venue.

Given the exploratory nature of this paper to examine whether places can be characterized through human activities, we did not examine every venue and event, but targeted venues with high similarity across different locations. Franchise businesses particularly match this criterion since they often provide consistent interior designs, services, and culture across their different locations; therefore, we focused on franchise venues in this study. Figure 1 shows an example of a franchise business in different locations. In order to select target franchise businesses, we first counted the number of events for each venue and sorted the list in descending order. From the venue with the largest number of events, each venue was examined whether it was a franchise with different locations. As a result, five franchises were chosen for the study: Crazy Mocha, Elks Lodge, King's Family Restaurant, Panera Bread, and Primanti Brothers. For the venues of these franchise businesses, duplicate venue records such as one restaurant with two different IDs were aggregated so to represent each physical place as one record. Based on the aggregated dataset, we descriptively present the data for the venues and activities.

Venues

Descriptive statistics about event information of each venue is presented in Table 1. Panera Bread, a café serving sandwiches, salads, and coffee, had the largest number of events for the six-month period in Pittsburgh. Elks Lodge was a non-trivial venue among the dataset since it was not a local business, but was a non-profit organization aiming to engage local community members to the neighborhoods via recreational and socializing events. The other venues were local coffee shops and restaurants.

The average number of events per location implies how popular each place is for people's meet-ups. On average, Crazy Mocha had the highest rate of activities and King's Family Restaurant had the least. Of course, the distribution of local events is not even across the franchise branches, which can be observed in the standard deviations. The most popular place was a Crazy Mocha branch located in the Shadyside neighborhood (74 events over six months).

Table 1: Basic Information about Target Places

	Crazy Mocha	Elks Lodge	King's Family Rest.	Panera Bread	Primanti Brothers
Classification	Coffee	Non-profit	American Rest.	Sandwich Cafe	Sandwich
# Events	94	24	19	178	25
# Locations	5	2	3	21	3
Avg. Events per Location	18.80	12.00	6.33	8.48	8.33
S.D.	31.02	12.73	6.66	9.80	8.08

Activities

In order to quantify activities that happened in each place, we adopted a topic model. Latent Dirichlet Allocation (LDA) model was used to predict topics for each event description (Blei et al. 2003). Ten unique words were extracted from each event's topic, and then the occurrence of each word was counted for each venue. For example, if five events were organized in a venue, 50 words in total would be extracted for the venue. There may or may not be duplicate words among them. As a result, a venue was presented as a list of multiple words (e.g., {workshop: 5, game: 3, eat: 4}). Table 2 shows the number of words for each franchise. The average number of topic words per location indicates the variety of activities in each franchise. For example, local events organized at Elks Lodge venues are more diverse compared to the others. This intuitively makes sense because Elks Lodge branches not only have dining areas, but also various kinds of facilities such as golf course and swimming pool, which may facilitate diverse activities.

Table 2: Number of Topic Words by Franchise

	Crazy Mocha	Elks Lodge	King's Family Rest.	Panera Bread	Primanti Brothers
# Words	70	57	39	342	39
Avg. Words per Location	14.00	28.50	13.00	16.29	13.00
S.D.	16.10	0.00	5.77	27.81	5.2

We analyzed the branches for Crazy Mocha and Panera Bread in more detail since these two franchises had the most activities than the others. Several branches were compared by looking at the descriptions of the local event data. Then, these descriptive analyses were contrasted againt the conventional identity defined by the business on Yelp. Different from MeetUp, Yelp's place information is moderated and officially overseen by the Yelp administrator. Business owners or users can claim a place, and the category and information about the place are determined by the staff after their own verification. This allows us to compare an activity-oriented place to its official identity.

Results

Since ten unique words were identified from the topic of each event (or activity) description, it was possible to characterize each venue with the frequency of words by aggregating them. The frequency can be interpreted as the weight of each word in modeling a place. If there are fifteen words associated with a venue, for instance, the venue can be represented by a 15-dimensional vector where the indices are words and the values are the number of occurrences for each word. The activity structures of Crazy Mocha and Panera branches are presented in details using both quantitative and qualitative approaches.

Crazy Mocha Locations

Crazy Mocha is officially categorized as "Coffee & Tea" place by Yelp. Table 3 shows some different patterns with top-10 words that describe Crazy Mocha venues. The Squirrel Hill branch had the highest rate of activities with 74 events during the six months of our data collection. It is possible to speculate the characteristics of this place through its high-ranked topics such as "games" and "play," which suggested that the place might have many activities about board games. When we examined the actual events in detail, it was found that 'Pittsburgh Eurogames,' a MeetUp group, has been organizing regular board game events at the Crazy Mocha Squirrel Hill branch every week. They were open to the public and provided lessons for beginners. Also, there were some socializing meetings at the venue.

Table 3. Word Frequency of Crazy Mocha Venues

	Squirrel Hill	South Side	South Shore	Friendship	Shadyside
1	Sign (50)	Please (7)	Please (1)	Free (3)	Time (9)
2	Please (36)	Bring (7)	Bring (1)	Meeting (3)	Conversation (9)
3	Bring (35)	Something (7)	Free (1)	Time (3)	Make (9)
4	Day (35)	Conversation (7)	Board (1)	Event (3)	Keep (9)
5	Games (35)	Help (7)	Card (1)	Every (3)	Learn (9)
6	Like (35)	Ill (7)	Cards (1)	Fitness (3)	Next (9)
7	Play (35)	Practice (7)	Others (1)	Heaven (3)	Open (9)
8	Players (35)	See (7)	Round (1)	Language (3)	Opportunities (9)
9	Remember (35)	Start (7)	Taking (1)	Month (3)	Speakers (9)
10	Welcome (35)	Work (7)	Winner (1)	Pa (3)	Starting (9)
# of Unique Words	46	10	10	10	10
# of Events	74	7	1	3	9

Crazy Mocha's Squirrel Hill branch was also examined qualitatively through Yelp to validate the result. There were 25 comments for the venue, and three visitors actually commented about board games in the place. Some users showed sentimental expressions for the gamers. User A described the branch as a place of study groups and board gamers, saying:

> The shop is spacious and often full. I frequently see study groups or board gamers when I'm here.

Different from the Squirrel Hill branch, the activities from the Shadyside branch were mostly about conversation, learning, and speaking. Only one MeetUp group had organized events in the Shadyside branch, and it was a series of gatherings among German-speaking people. Due to the small number of activities at the venue, however, there was no comment about German-speaking meet-ups.

Table 4. Word Frequency of Panera Bread Venues

	Waxford	Larimer	Mt. Lebanon	Robinson	Oakland
1	Please (16)	Players (18)	Life (10)	Group (5)	Games (25)
2	Beginning (15)	Around (18)	Love (10)	Work (5)	Play (24)
3	Coffee (15)	Begin (18)	Meditation (10)	Open (5)	Every (22)
4	Conversation (15)	Beginner (18)	Help (10)	Want (5)	May (22)
5	Else (15)	Cards (18)	Energy (10)	discussion (5)	Club (22)
6	Group (15)	Long (18)	Healing (10)	Anxiety (5)	Meets (22)
7	Intermediate (15)	Played (18)	Living (10)	Back (5)	Minutes (22)
8	Month (15)	Table (18)	Others (10)	Holding (5)	Per (22)
9	Pour (15)	Tables (18)	Receive (10)	Hope (5)	Scrabble (22)
10	Help (15)	Time (18)	Spirit (10)	Learn (5)	Session (22)
# of Unique Words	49	56	84	87	94
# of Events	21	29	19	13	35

Panera Bread Locations

Panera Bread is tagged as "Sandwiches, Salad, Soup" in Yelp. This franchise was the most popular one for organizing events by MeetUp groups in the city of Pittsburgh. Among the 21 Panera Bread branches on MeetUp, we analyzed the top-5 venues in the number of events. Table 4 shows the top-10 topic word frequency for the venues. Different from Crazy Mocha where one branch dominated in the number of events, local events in Panera Bread were relatively evenly distributed across the branches. The Oakland branch had the largest number of events with 35 during the six months. Twelve out of 21 branches had less than six events during the six months, which means less than one event per month.

Every event organized at the Waxford branch was about conversations. A French conversation group organized most of the events, and characterized the place with relevant words "conversation," "group," and other words indicating the level of language skills. Interestingly, "pour" was detected as one of the top-ranked words because the group described several events in French. Since the LDA algorithm was set to model English words only, it appeared to be in a topic model.

The Larimer and the Oakland branches were mainly characterized as board game places from the topic models, and several MeetUp groups actually organized board game events frequently at the places. Specifically, the Larimer branch's board games were mostly Mah Jong game, something that was implied by the words, but not explicitly captured. On the other hand, the Oakland branch's game was Scrabble, which was directly captured by the topic model. There were 22 meet-ups for the Scrabble game during the six months. The Oakland branch also had diverse activities compared to other branches such as political movement, foreign language conversation, and book reading.

The Robinson branch also had unique activities in the place. There were some educational sessions about wellness and health. However, the evidence of wellness meetups were captured in the low-frequency words. In the Mt. Lebanon branch, most events were from a spiritual group, a women's book club, and a local business network. Large numbers of the spiritual group and the women's inspirational book club resulted in many affectionate words such as "love," "life,", "energy," and "healing."

In the Yelp reviews for Panera Bread branches, however, there were few comments about activities and social aspects of each place. It was possible to see how activities and events shaped urban places, but due to the small amount of information both from Yelp and MeetUp, there are some points to discuss for the future work.

Limitations and Discussions

Through the exploratory study to characterize urban places based on MeetUp's local event data, we showed that the official identity of a place can noticeably diverge from what a place means for those appropriating the place and the socially constructed form of our living spaces. We present the limitations of this study, and discuss potentials and future work.

Dataset Size

The amount of local event data to describe an urban place is not enough if we use only one source of information. MeetUp was not exception from the limitation. The highest frequency that we observed was the board game events that were organized in a coffees shop every week. This kind of regular activity data provides an insight into the dynamics of a place to some degree. Even though, an activity that happens once a week is not enough to quantify and characterize an urban place.

Lopez, Butler, and Brusilovsky showed that local information sources are highly fragmented, and only about 20% of the overall information were available from one information source (Lopez; Butler and Brusilovsky 2014). In order to make this approach useful, there should be a strategy to automatically combine local event and other human activity data from multiple sources in the semantic level. Also, targeting a bigger city would help in collecting more data.

Other Potential Datasets

People's comments or reviews about activities have potential to be useful in realizing the concept of place. User reviews on Yelp provided hidden information about places. Of course, the information about local events that were identified from MeetUp data was mostly unavailable on Yelp. Even though, some comments shed lights on using the reviews as complementary data. For example, there was a user comment about an event that the user liked at the Larimer branch of Panera Bread. She described the reason why this place was suitable for doing specific activities:

> I have always loved Panera Bread for many reasons: the great variety and changing seasonal menu that takes them beyond a breakfast in-and-out dash, the atmosphere that is simply so conducive to group project work, finishing class assignments or just hanging out and gabbing for hour with my best friend.

Even though she does not talk about actual activities, this kind of sentences provides an insight into how the space can or already be used by people implying prospective activates, and can be potentially used in the characterization process.

Topic Models

Since most event descriptions on MeetUp are one to two paragraph texts, it is not easy to capture topics well like Wikipedia documents. Some more accuracy measures and tests are needed for the dataset. In order to do that, LDA parameters such as the number of topic models should be empirically tuned for the dataset. Furthermore, sequence-based topic models might be useful since events and activities in a place occur sequentially (Farrahi and Gatica-Perez 2012). The sequence of topics may provide more insights into characterizing places.

Another point to consider is about foreign languages from the dataset. There were many events about foreign language practice and conversations, and some of their event descriptions are written in foreign languages. Failing to deal with the languages may make it harder not only to identify the characteristics of urban places, but it also misleads the topic models due to some same-spelling words with different meanings (e.g., pour in French).

Comparison between Places

Once a vector of a place is constructed based on the topical words, it is possible to measure the distance between two places using the cosine similarity score or other distance measures. This approach allows to cluster places based on the similarity of "placeness" between two places.

Urban places characterized by human activities may also indicate the land use or the cultural aspects of a neighborhood in the future. For example, many woman-focused business and spiritual activities from the Panera Bread branch in Mt. Lebanon imply that the Mt. Lebanon region might have a cooperative culture within the neighborhood.

These kind of hidden knowledge can be beneficial to urban planners, policy makers, and business owners by providing rich knowledge about dynamically changing urban places.

Conclusion

This study explored whether local event data could be used to model urban places in the way that captures social meanings and human activities associated with the venues. The descriptive analyses suggested that event datasets have great potential for constructing computational models for urban places. Subsequently, we discussed some ways to enhance the quality of the data and potential datasets that might complement the event data. By combining user reviews, comments, and other information about activity traces, urban places could be further refined in the characterization process.

Modeling urban places based on human activity is not just for capturing people's behavior in a city, but to better understand and realize the concept of place. Since a large amount of local event data are available nowadays due to the prevalence of event-based social media and organizations' uses of the internet to advertise events, the potential advantages of using this approach would become larger.

References

Blei, D. M., Ng, A.Y., and Jordan, M. I. 2003. Latent dirichlet allocation. *the Journal of Machine Learning Research* 3: 993-1022.

Cranshaw, J., Schwartz, R., Hong, J. I., and Sadeh, N. 2012. The livehoods project: Utilizing social media to understand the dynamics of a city. In *ICWSM*.

Del Bimbo, A., Ferracani A., Pezzatini D., D'Amato F., and Sereni M. 2014. LiveCities: Revealing the Pulse of Cities by Location-Based Social Networks Venues and Users Analysis. In *WWW*.

Farrahi, K., and Gatica-Perez, D. 2012. Extracting mobile behavioral patterns with the distant n-gram topic model. In W*earable Computers (ISWC), 16th International Symposium on*. 1-8. IEEE.

Gieryn, T. 2000. A space for place in sociology. *Annual Review of Sociology,* 26: 463-496.

Harrison, S., and Tatar, D. 2007. Places: People, Events, Loci – the Relation of Semantic Frames in the Construction of Place. *Computer Supported Cooperative Work (CSCW)* 17 (2-3): 97–133.

Le Falher, G., Gionis, A., and Mathioudakis, M. 2015. Where Is the Soho of Rome? Measures and Algorithms for Finding Similar Neighborhoods in Cities, In *ICSWM*.

López, C., Butler, B., and Brusilovsky, P., 2014. Does Anything Ever Happen Around Here? Assessing the Online Information Landscape for Local Events. *Journal of Urban Technology* 21: 95–123.

Rösler, R., and Liebig, T. 2013. Using Data from Location Based Social Networks for Urban Activity Clustering. In *Geographic Information Science at the Heart of Europ*e, edited by Vandenbroucke, D, Bucher, B, and Crompvoets, J. 55–72. *Lecture Notes in Geoinformation and Cartography*. Springer International Publishing.