Inferring Travel Purpose from Crowd-Augmented Human Mobility Data

Zack Zhu
Wearable Computing Lab
ETH Zurich
zack.zhu@ife.ee.ethz.ch

Ulf Blanke
Wearable Computing Lab
ETH Zurich
ulf.blanke@ife.ee.ethz.ch

Gerhard Tröster
Wearable Computing Lab
ETH Zurich
troester@ife.ee.ethz.ch

ABSTRACT

Affordances from the urban space shape the way we interact with our environment, whether manifested as driving into the city centre for work or playing sports in designated arenas. Given today's abundance of crowd-generated digital traces on location-based social network (LBSN) platforms, an opportunity arises to grasp deeper semantic characterization of urban affordances beyond static representations found in traditional GIS systems. Complementing these perceptions of the city, travel surveys capture mobility dynamics of people with absolute trajectory recordings and explicit travel purposes. By marrying rich LBSN data with travel surveys, we ask if crowdsourced urban characteristics can be used to explain user behaviour when interacting with the city. Concretely, our objective is to model and infer the purpose of travel, or the activity at the destination of a trip, in daily life scenarios. To this end, we generate features to correspond to time, location, and demographics in order to construct a fused understanding of people's travel purposes. Using LBSN data to augment a travel survey of 87,600 trips by 10,372 people, we show that fusion of extracted features can achieve an interpersonal prediction accuracy of >75% for 9 broad classes of travel purposes covering typical aspects of life. This represents an increase of nearly 20% compared to without LBSN augmentation.

Keywords

Crowd-generated Urban Data, Human Mobility Mining, Data Fusion

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications – data mining, spatial databases and GIS

1. INTRODUCTION

Today's cities are complex systems of interconnected elements, such as of schools, shopping malls, or homes. Our lifestyle defines their temporal use as well as how we move, reside or interact with this urban system. By adapting the urban layout over time to meet our demands, spatial-temporal constraints emerge for how we interact with our urban elements.

To capture this, we introduce the notion of *urban affordance* based on Gibson's general theory of affordances [6]. Analogous to ecological niches that afford shelter or food, the urban space affords education, income through work, or socializing. Often embedded as part of a larger urban environment, places of affordances are without explicit boundaries, but influenced by nearby elements [6]. To further complicate matters, places in the urban space can provide multiple affordances with respect to time or demographic background. For example, schools provide the affordance of education for pupils, but work for teachers. Similarly, a downtown district affords employment during working hours while affording meals and socializing at noon or in the evening.

Clearly, it is difficult for static layers in traditional GIS systems to capture such interactions between the city and its residents. However, the increasing popularity of location-based social networks (LBSN), such as Foursquare[1], provide a unique platform for gathering crowd-generated perceptions of the urban environment. Such data, often generated in-situ by smartphones or other components within the Internet of Things (IoT), amass to a rich and dynamic repository of urban perception. In addition to IoT-generated data for representing the city, travel surveys, conducted by municipal governments, provide valuable samples of travel behaviour of residents. Coupling crowdsourced urban reflections with mobility recordings in terms of absolute geographical coordinates, a rich set of variables can be used to understand the affordances of a city and how people leverage them.

In this paper, we investigate the possibility of automatically inferring why people travel, based on urban affordances reflected in crowd-generated LBSN data. Along with other signals characterizing users and the geography of trips, we fuse together an augmented travel survey dataset. Then, this dataset describes *who*, via demographics of the traveller, *when*, via travel time, and *where*, via location semantics from LBSN data. It also captures the travel purpose explicitly, or *why*, serving as a valuable label of ground truth to map against. With these explicit labels, we are able to apply supervised machine learning algorithms and objectively evaluate our model's capability in inferring travel purposes. The main contribution of this paper is, therefore, an approach for enriching travel survey data with deeper semantics. We demonstrate a strong net gain (~20%) in prediction accuracy when applying our enrichment approach. We further provide a thorough importance analysis to interpret our features quantitatively for understanding the predictive powers of who, when, where towards why.

Urb-IoT'14, October 27 - 28 2014, Rome, Italy

ACM 978-1-4503-2966-8/14/10$15.00
http://dx.doi.org/10.1145/2666681.2666685

[1] `http://www.foursquare.com`

2. RELATED WORK

Early work in understanding spatial-temporal characteristics of groups of people include the work of Eagle and Pentland [4], where 100 members of MIT recorded trajectory traces via mobile phones. Later, a larger study of 100,000 users was conducted by Gonzalez et. al. [7] to identify strong spatial and temporal regularities in human trajectories. More recently, researchers are exploiting the increasing prevalence of GPS-tracked devices to interpret trajectory patterns and behaviour in terms of activities conducted in various spatio-temporal settings. These include the use of call detail records [10], digital traces from social media platforms [13], city-wide bike-sharing traces [3], pedestrian movements at festivals [1], and GPS tracking of taxicab fleets [14]. Although such datasets can provide detailed trajectories of numerous individuals, there lacks an explicit understanding of the individual's intention for making a trip. In other words, while such unlabelled data is plentiful, the label of travel purpose is missing, resulting in the inability to objectively assess the generalizability of proposed models.

As travel surveys contain valuable labels to understand the purpose of trajectories, researchers have started to leverage these datasets as we do in this paper. One notable work, by Jiang et. al. [8], utilizes an activity-based travel survey to cluster for demographic profiles of individuals. Their findings show that finer-grain profiles can be extracted beyond the original worker, student, and non-worker profiles. In comparison, our work focuses on identifying the purpose of a trip as opposed to further understanding who would conduct a trip fitting a given clustering of spatiotemporal characteristics. Another related work utilizing travel survey data is that of Krumm and Rouhana [9]. In the same direction as our work, their objective is also to infer trip purpose. However, we differentiate our work by incorporating crowd-generated LBSN data to characterize the spatial aspect of our model. We show that fine-grained venue categories and unstructured textual data provide significant gains in correctly inferring trip purposes as well as add interpretability to urban landscapes.

3. AUGMENTING TRAJECTORIES WITH CROWDSOURCED SEMANTICS

By mapping the physical coordinates of travel survey data into the virtual world of location-based social media, we are able to collect rich, location-specific, and crowd-generated data to semantically augment people's trajectories. In the following section, we describe the original Puget Sound travel survey dataset of 2006 and our method for fusing it with data from Foursquare venues. As we collect the data from Foursquare in 2014, there is a temporal discrepancy between the two datasets. We intend to address this discrepancy in future work by utilizing a new 2014 travel survey from Puget Sound, which will be available in late 2014.

3.1 Puget Sound Travel Survey

The Puget Sound Research Council travel survey (PSRC) contains consecutive 48-hour travel surveys of 10,372 individuals residing in the Puget Sound region of Washington State, USA. From the survey, a total of 87,600 trips are recorded, including those of children and seniors [12]. Compared with other large-scale human mobility recordings, the PSRC dataset also surveys for the travel purposes of trips. This is valuable as it explicitly encodes *why* a trip was made. The travel purpose categories, with occurrence frequency in brackets, are as follows: Home (37,122), Work (11,776), Personal Business (7,877), Escort (7,819), Shopping (7,524), Recreation (4,766), Eating Out (4,041), Attending School (3,468), Social (3,202), and Unlabelled (5). As a preprocessing step, we remove

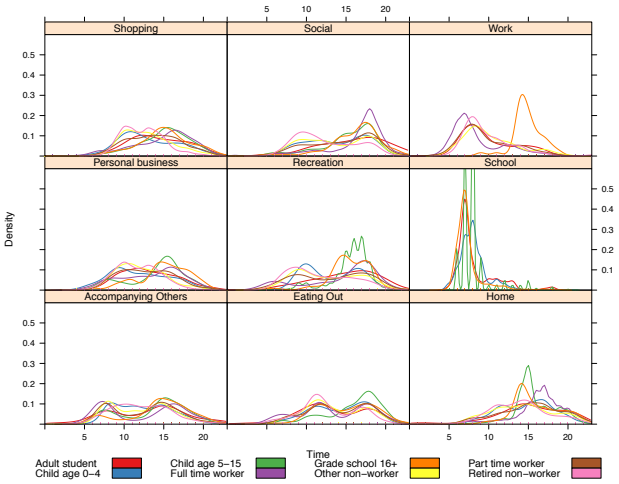

Figure 1: Temporal distribution of travel purposes for different demographical groups.

the five unlabelled instances from our experimentation. Detailed interpretations of these labels are described in [12].

To provide qualitative intuition for our extracted features, we visualize the interaction of temporal-demographic and temporal-spatial characteristics of the dataset. In Figure 1, we plot the temporal density of departure times for different travel purposes. Within the plot of each purpose, we further separate the density by demographic data given in the travel survey. Overall, we notice a high level of density overlap with respect to time and demographics. However, some demographic-temporal combinations distinguish themselves. For example, we notice retired non-workers tend to socialize early during the day while working or studying adults tend to do so late in the evening. Looking at students and adult workers, we see a sharp increase in school and work-related travel during the morning commute hours. Near the end of their day, they then depart for home after school (around 3pm) or work (around 5pm). Interestingly, we also see 16+ students heading to work (presumably their part-time jobs) in the late afternoon and evenings. This visualization shows that, although temporal-demographic data overlaps in many cases, there are some well-defined combinations with discriminating characteristics.

Just as important as time is the spatial distribution of people's travels. In Figure 2, we plot the destination points, coloured by the associated travel purposes. To capture change in travel purposes over time, the data is divided into quartiles with respect to a 24-hour period and shown separately. To avoid over-plotting, we select 5 prominent categories of travel purposes and plot a uniformly sampled distribution (10% of the data) for a zoomed-in region. From the spatiotemporal plot, we can see the change in urban affordances over time in the same spatial areas. During the first quartile, people depart for work and school. Work is heavily concentrated in the city centre (in blue) while school is evenly distributed around the region (in orange). Around mid-day, many still head for work (or perhaps returning from lunch) while other travel purposes start to rise in popularity, such as shopping (red) and socializing (purple). By the afternoon, the commute home (green) begins and the city centre transforms to afford more non-work-related purposes, such as shopping and social. In the last quartile, very few depart for work while most travel to go home. At the same time, socializing and shopping take more dominant roles in the city centre. With

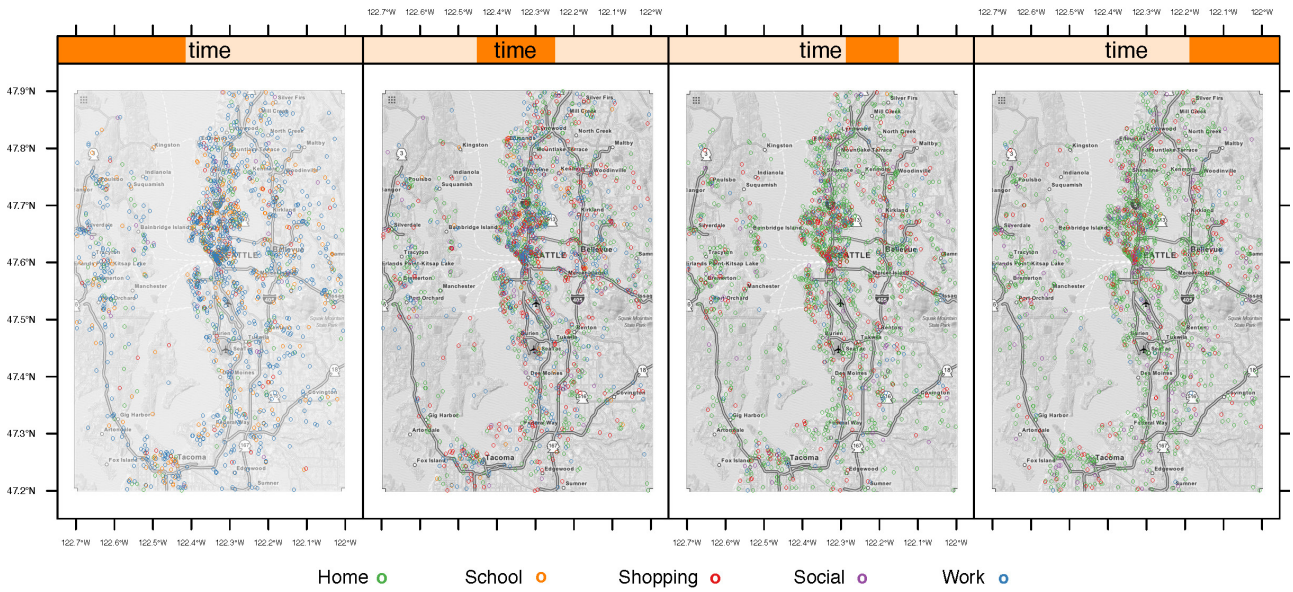

Figure 2: A spatial-temporal plot illustrating the spatial density of various travel purposes in the Puget Sound region and their progression throughout the 24-hour period.

these transitions, we show that there indeed exists spatiotemporal patterns capable of discriminating people's travel purposes.

3.2 Semantic Augmentation via Foursquare

By filling out a confidentiality agreement, we obtain the geographical coordinates of the surveyed trips from the Puget Sound Research Council. Therefore, we are able to link online data from Foursquare with the offline trajectories recorded in the PSRC dataset. To do so, we query the Foursquare Venues API[2] for nearby venue data so that we obtain a distribution of venue types around the coordinate and a concatenation of "venue tips" written for the respective venues.

The Foursquare categorization hierarchy is extensive with 583 low-level venue categories under 10 high-level venue category groupings[3]. Venue tips are unstructured text snippets that Foursquare users post as interesting and relevant information about the venue. As it is free-form text, it can range from dish recommendations for a particular restaurant to class selection strategies at a university. Therefore, the semantic augmentation we receive from Foursquare can be thought of in three levels of detail: high-level semantic venue groupings (e.g. "Arts & Entertainment"), specific venue category labels (e.g. "Indie Movie Theater"), and detailed text snippets detailing a particular aspect of this venue (e.g. *"Great old theater that doesn't have a bad seat in the house. Beer and wine avail too"*). In the latter parts of this paper, we will demonstrate the effectiveness of these three levels of semantics for inferring travel purposes and the benefits of leveraging fine-grained, semantic spatial characterization.

4. MODELLING TRAVEL PURPOSES

In the following section, we first describe concrete features we extract to represent three aspects of the augmented travel survey: demographic, temporal, and spatial. Following, we examine the classification performances gained by leveraging different feature sets and interpret their importances.

4.1 Feature Extraction

As mentioned previously, we leverage three aspects: *who, when,* and *where*. The description and derivation of these feature sets are as follows:

Demographic Features: In our list of travel purposes, some can naturally be differentiated by understanding who makes the trip. For example, we can differentiate Work and School by age group even if two trips share the same spatiotemporal features. We create binarized age group membership for the following ranges as provided directly by the PSRC dataset: *Under 5, 5 to 15, 16 to 17, 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, 65 to 74, 75 to 84, 85 and over, Don't Know, and Refused*. Since the PSRC dataset also provides an occupation description, we similarly create binary features to indicate the following types: *Full time worker, Part time worker, Retired non-worker, Other non-worker, Adult student, Grade school 16+, Child age 5-15, and Child age 0-4.* We also binarize the reported gender to add three additional features (female, male, refused). Therefore, we extract a total of 24 binary demographic features.

Temporal Features: Intuitively, timing of a trip is a strong signal for routine travel purposes. We create 48 bins to represent every 30-minute window in the 24-hour period. Concretely, we construct a binary indicator vector $T = [t_0, t_1, ..., t_{47}]$ such that an element of T is equal to 1 if it is within the beginning and ending timestamps of a trip. Additionally, the PSRC travel survey provides the duration stayed at a destination in minutes. We create a feature $D = \frac{ms}{1440}$, where ms represents the minutes stayed. Therefore, D represents the fraction of a day between 0 and 1. This results in a total of 49 temporal features.

Spatial Features: From the geographical coordinates of a trip's origin and destination, we extract nearby venue data from Foursquare. We define the notion of "nearby" via a tuned parameter r, which can be used to express an integer indicating the nearest r neighbour venues or a real number indicating a circular radius to include

[2] `https://developer.foursquare.com/overview/venues`
[3] `https://developer.foursquare.com/categorytree`

venue data from. After selecting the relevant Foursquare venues, we extract features as a function of counts:

- **Venue Category - High Level:** Given the venues within a radius, we map the venue categories to one of 10 venue groupings, as defined by the Foursquare categorization hierarchy. Therefore, we derive the feature vector $M = [m_0, m_2, ..., m_9]$ as integer counts of the following groupings: *Arts & Entertainment, College & University, Events, Food, Professional & Other Places, Nightlife Spots, Residences, Great Outdoors, Shops & Services, and Travel & Transport*. For each trip, we extract M_o and M_d for nearby features of the origin and destination, respectively, to extract 20 features.

- **Venue Category - Low Level:** Directly using the 550 types of venue categories occurring in our experiment, we similarly count the distribution within a specified radius. Due to the large number of venue types, we apply a *tf-idf* transformation, used typically in preprocessing n-gram features, to derive our feature vector $V = [v_0, v_1, ..., v_{549}]$. The purpose is to give higher weights to venues occurring less frequently throughout the whole dataset (e.g. airport) while down-weighing ubiquitous ones (e.g. cafe). We calculate each element v_t as

$$v_{t,d} = \begin{cases} \frac{1+\log tf_{t,d}}{(|D|/\ df_t)+1} & tf_t > 0 \\ 0 & else \end{cases}$$

  where t is a venue type, d represents the vicinity considered, and $tf_{t,d}$ is the raw count of venue t within the vicinity d. The term df_t denotes the number of occurrences of venue t throughout all the vicinities and $|D|$ is the total number of vicinities considered. For each trip, we derive 2×550 features for origin and destination characterization.

- **Venue Tips** For many Foursquare venues, unstructured textual tips are written by Foursquare users. Following standard text mining techniques, we concatenate all tip strings found within the radius of a geographical coordinate into one document. Then, we tokenize this string into words and extract uni- and bi-gram features. As this would generate a high number of features, we mitigate noise by stemming and removing stop-words of the English language. Finally, we apply the same *tf-idf* transformation as in the venue category processing to adjust the values of the uni- and bi-gram features. In our experiment, this feature set results in the addition of 2×418,019 extra features for fine-grained characterization of spatial coordinates.

4.2 Experimental Results

We construct a multi-class classification problem using the "one-vs-all" strategy to train individual classifiers for identifying each travel purpose. We use the L1-regularized Linear SVM [5] algorithm as implemented in Scikit-Learn [11]. The choice of this classifier is mainly due to its favourable performance with high-dimensional feature spaces, such as ones found involving textual features. The L1-regularization forces the coefficients of many features to zero to greatly reduce the resultant feature space for fast inference. Furthermore, the use of a linear kernel allows us to perform feature importance interpretation by looking at the magnitude of the feature coefficients.

For the results presented below, we report the testing accuracy using 10-fold cross-validation. It is important to note that we are assessing the interpersonal generalization ability of our model, which means we report accuracy metrics by testing on unseen users' trajectories. As such, we take special care that folds are not shared by a single user. This is important because the trips of one individual may be repeated multiple times (e.g. work to home on multiple days), which would artificially bloat accuracy if the same trips were split into both training and testing folds.

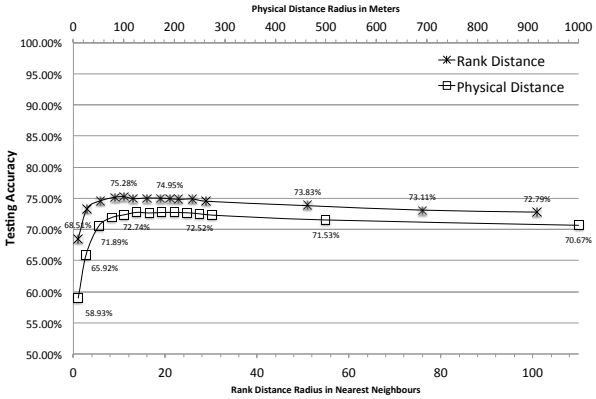

Figure 3: Tuning the parameter r as physical and rank-based radius for classification accuracy.

Geo-Coordinate Vicinity Selection: One crucial component of augmenting the travel survey dataset is to appropriately select the vicinity of a given geo-coordinate. We present two simple methods here and tune the radius parameter for each: radius selection by r-nearest neighbour and selection by a physical distance radius. Tuning r for both approaches, we plot the testing accuracy in Figure 3. It can be seen that the rank-based radius performs favourably compared to physical distance radius for all the attempted values. We believe this is due to the ability of the rank-based radius to maintain the same data density. In other words, the richness of the augmented venue feature vector is not diminished even for suburban areas with sparse points of interests. Furthermore, the circular nature of venue selection by the physical distance radius may explain poorer performance. For example, relevant neighbouring venues stretched out along the same street could be neglected while irrelevant venues on parallel streets are included. The highest accuracy is achieved with $r = 11$ (with 75.28%) using rank-based distance; this is the radius we use for the remaining results.

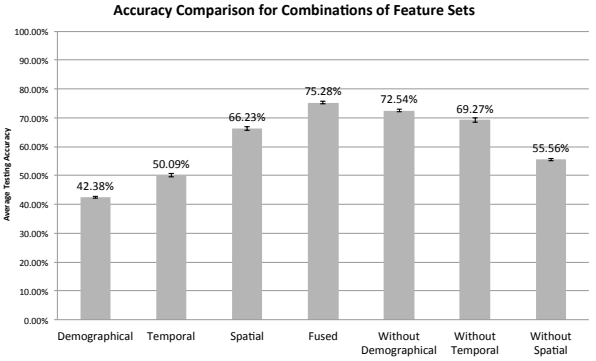

Figure 4: A comparison of accuracy for all combinations of features from demographical, spatial, and temporal aspects.

Utility of Feature Sets: In Figure 4, we plot the average testing accuracy achieved with the three feature sets and their combinations. By guessing the most frequently occurring class (Home at 42%), a trivial classifier can be constructed as a baseline for comparison. Of performances achieved with the individual feature sets, demographic features deliver similar accuracy as the trivial classifier, while temporal and spatial features achieve approximately 50% and 66% accuracy, respectively. We achieve the highest testing performance when combining all feature sets (by simple feature-level concatenation), at 75.28% accuracy. To determine the accuracy gain as a result of each feature subset, we conduct subtractive analysis to investigate the decrease in classification accuracy when a feature set is removed. We notice the largest drop of 19.72% when the spatial augmentation is removed. On the other hand, removing demographical features resulted in the least performance decrease of only 2.74%. Hence, our approach would be applicable without significant loss in performance even applied to datasets without demographic information, such as anonymized call detail records or location-based social media data.

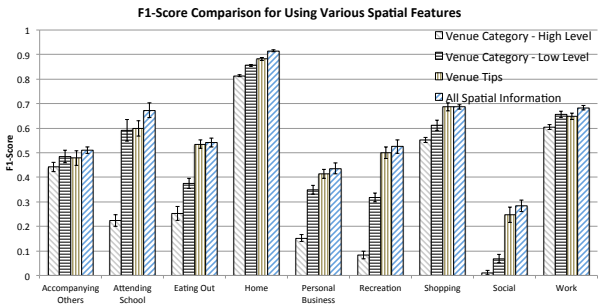

Figure 5: A comparison of F1-Scores for individual travel purpose classes when inferred using various subsets of the spatial feature set.

Spatial Feature Analysis: What is essential to performance are the location semantics derived from the crowd. We take a closer look at the effect of leveraging different levels of detail in augmentation, as described in Section 3.2. Here, we utilize only enriched trajectory data without demographics, as this is not available in many scenarios. In Figure 5, we plot the class-specific F1-score so that we account for class imbalance effects. We see a clear trend of increasing performance with increased semantic granularity of augmentation. However, the best performance is achieved when all three levels are used simultaneously. It is also clear that the two dominant classes (Home and Work) with more data tend to achieve higher F1-scores. On the other hand, the least populated class of Social has the lowest F1-score.

By using only the high-level venue categories, similar to [9], we achieve the lowest F1-scores. All classes benefit from the use of low-level venue categories, with the largest performance increase in Attending School (+0.36), Personal Business (+0.19), and Recreation (+0.24). With even more fine-grained features from venue tips, we see additional performance increases over low-level venue categories, especially for Social (+0.18), Eating Out (+0.16), and Recreation (+0.18). With crowd-generated data, multiple aspects of a vicinity can be captured beyond a static set of predefinitions. As people interact with a vicinity, their reports (venue category labelling or textual descriptions) provide a dynamic reflection. Using such rich characterizations, detailed classification boundaries can be constructed as the classification algorithm selects from a large set of meaningful features.

4.3 Feature Importance and Interpretation

Travel Purpose	*Nearby Venue Types*
Accompanying Others	Street Art, Convention, Taiwanese Restaurant, Hockey Field, Cycle Studio, Mini Golf, Australian Restaurant, Big Box Store, Elementary School, Hindu Temple, Ski Area, Private School, Disc Golf, Airport Gate, Molecular Gastronomy Restaurant, Climbing Gym, Roller Rink, Middle Eastern Restaurant, High School, Motorcycle Shop
Attending School	Warehouse, Bathing Area, Lingerie Store, Hindu Temple, Roof Deck, Shipping Store, Trailer Park, High School, Distillery, Elementary School, Camera Store, Piano Bar, Travel Lounge, German Restaurant, Cuban Restaurant, Alternative Healer, Airport, College Stadium, Middle School, Auditorium
Eating Out	Stoop Sale, Cambodian Restaurant, Mongolian Restaurant, Jazz Club, Radio Station, Law School, Pawn Shop, Dim Sum Restaurant, Candy Store, Fast Food Restaurant, Laundromat, Moroccan Restaurant, Stationery Store, Afghan Restaurant, University, Bike Rental / Bike Share, Internet Cafe, Badminton Court, Boutique, Baby Store
Home	Mattress Store, Modern European Restaurant, Capitol Building, Music Festival, College Football Field, College Communications Building, Fabric Shop, Dry Cleaner, Subway, Astrologer, Theme Park Ride / Attraction, Jewish Restaurant, Other Event, Taxi, Snack Place, Swim School, Beer Garden, Portuguese Restaurant, Disc Golf, College Engineering Building
Personal Business	Gymnastics Gym, College Technology Building, Bistro, Piercing Parlor, Planetarium, Organic Grocery, Luggage Store, Roof Deck, Cycle Studio, Library, College Math Building, Hakka Restaurant, Post Office, Outdoor Supply Store, National Park, College Library, Leather Goods Store, Vineyard, Residence, Rental Car Location
Recreation	Boxing Gym, Paella Restaurant, Big Box Store, Cultural Center, Soccer Stadium, College Soccer Field, Leather Goods Store, Gym Pool, Fair, Squash Court, Golf Course, Nature Preserve, Music Festival, South American Restaurant, Caribbean Restaurant, Science Museum, Organic Grocery, Newsstand, Mosque, Convention
Shopping	Airport Lounge, Grocery Store, Warehouse Store, Hockey Arena, Volcano, Indonesian Restaurant, Spanish Restaurant, Law School, Smoothie Shop, Gun Shop, Palace, Szechuan Restaurant, Aquarium, Light Rail, Supermarket, College Gym, Department Store, Afghan Restaurant, Shoe Repair, Donut Shop
Social	Comfort Food Restaurant, College Baseball Diamond, Country Dance Club, Cycle Studio, RV Park, Big Box Store, Argentinian Restaurant, Ski Chalet, Church, Street, Religious School, Distribution Center, Hot Dog Joint, Record Shop, Dry Cleaner, German Restaurant, Bistro, Synagogue, Voting Booth, Internet Cafe
Work	English Restaurant, Watch Repair Shop, Cricket Ground, Outlet Store, College Soccer Field, Print Shop, Country Dance Club, Castle, Pedestrian Plaza, Festival, College Technology Building, Afghan Restaurant, Nudist Beach, Embassy / Consulate, Gun Range, Health Food Store, College Science Building, Warehouse, Mac & Cheese Joint, Prison

Table 1: We list the top 20 most important venue category features that represent the vicinity of destinations of each travel purpose class. Although noisy, we notice intuitive interpretations of nearby venues that fit the neighbourhoods for the different types of purposes.

As we use a linear kernel SVM, we are able to rank the importance of features via their coefficients [2]. We fit the complete dataset to a model using all features described so far. Since L1-regularization is used, feature selection is conducted implicitly as most coefficients shrink to zero. For the remaining features with non-zero coefficients, we group by feature sets as described previously and plot the distribution of importance ranking for each class in Figure 6 (smaller the rank, the more important). For each box plot, the median is depicted by a black dot while features within the interquartile range (between the 25th and 75th percentile) are enclosed in the rectangular box. The whiskers depict data reaching as far as 1.5 times the interquartile range and further outlying points are plotted as circles.

For all classes, the median rank of tip features consistently rank higher than other feature groups, speaking to the usefulness of leveraging crowd-generated tips data. On the other hand, we notice features from high-level venue categories ranking the lowest. For the temporal features, a handful of outliers rank highly for some classes, such as Shopping, Recreation and Eating Out. This suggests that some activities are inherently more routine with respect to time than others. By fusing all features together, we ensure the classification boundaries selectively leverage temporal, spatial and demographical characteristics simultaneously.

Using the same feature rankings, we can also interpret the model to qualitatively grasp vicinities characterizing different travel purposes. We plot the top 20 most important venue types for each class in Table 1. Here, we see some purposes are afforded in vicinities less distinguishable than others, namely, Home and Work. However, this is compatible with our intuition as these two classes can

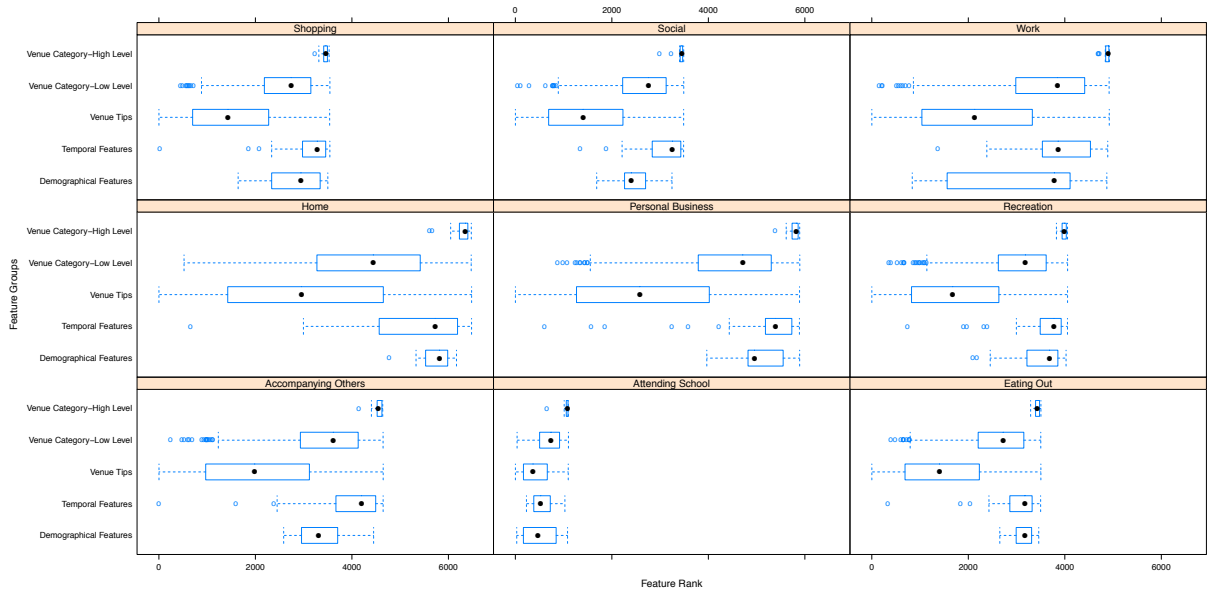

Figure 6: We rank the importance of features by the absolute value of their coefficients. We plot the rank distribution of various feature sets against 9 classes of travel purposes. Due to the use of L1-regularization for in the Linear SVM classifier, the majority of features are dropped due as weight coefficient shrink to 0.

be embedded in a variety of environments. Homes are found in suburban residential areas as well as city centre condominiums. Similarly, work does not necessarily take place only within the central business district of a city. On the other hand, vicinity composition does seem more descriptive for purposes such as Attending School, Eating Out, or Recreation. Intuitively, these are more constrained to certain elements in the urban space. For example, Attending School takes place near a variety of educational establishments while Recreation is fulfilled by venues such as Cultural Centres or various sports arenas. From these vicinity interpretations, we find fine-grained venue categories are quite useful for a data-driven understanding of urban affordances and their mixing elements.

5. CONCLUSION AND FUTURE WORK

In this work, we illustrated an approach to enrich semantic understanding of human trajectory from crowd-generated urban characterizations. We evaluate on a large-scale travel survey and observe a maximum of 75% accuracy when utilizing features corresponding to who, when, and where. By analyzing the importance of various feature sets, we illustrate the importance of urban space characterization and temporal features for inferring travel purposes, while demographic data brings the least improvement. As such, we believe trajectory data alone (containing geo-coordinates and time), which may be obtained anonymously using IoT devices, could be used to estimate people's travel purposes with a sufficient degree of accuracy. In future work, we intend to extend our methodology to work in tandem with current travel survey methods. For example, by deploying mobile apps to query travel study participants with model-computed estimates of travel purpose.

6. ACKNOWLEDGEMENTS

We are grateful to Neil Kilgren of the Puget Sound Regional Council for providing the travel survey data. We also thank the anonymous reviewers. This work was partially supported by the Hasler Foundation through the SmartDAYS project.

7. REFERENCES

[1] Blanke, U., Franke, T., Tröster, G., and Lukowicz, P. Capturing crowd dynamics at large scale events using participatory gps-localization. ISSNIP '14 (2014).

[2] Chang, Y.-W., and Lin, C.-J. Feature ranking using linear svm. In *WCCI Causation and Prediction Challenge* (2008), 53–64.

[3] Coffey, C., and Pozdnoukhov, A. Temporal decomposition and semantic enrichment of mobility flows. LBSN '13 (2013), 34–43.

[4] Eagle, N., and (Sandy) Pentland, A. Reality mining: Sensing complex social systems. *Personal Ubiquitous Comput. 10*, 4 (Mar. 2006), 255–268.

[5] Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., and Lin, C.-J. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research 9* (2008), 1871–1874.

[6] Gibson, J. J. The theory of affordances. *Hilldale, USA* (1977).

[7] Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. Understanding individual human mobility patterns. *Nature 453*, 7196 (2008), 779–782.

[8] Jiang, S., Ferreira, J., and Gonzalez, M. Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery 25*, 3 (2012), 478–510.

[9] Krumm, J., and Rouhana, D. Placer: Semantic place labels from diary data. UbiComp '13 (2013), 163–172.

[10] Noulas, A., Mascolo, C., and Frias-Martinez, E. Exploiting foursquare and cellular data to infer user activity in urban environments. MDM '13 (2013), 167–176.

[11] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research 12* (2011), 2825–2830.

[12] Puget Sound Regional Council. 2006 Household Activity Survey: General User's Guide. `http://www.psrc.org/assets/4358/2006HHSurveyUserGuide.pdf`, 2006.

[13] Wu, L., Zhi, Y., Sui, Z., and Liu, Y. Intra-urban human mobility and activity transition: Evidence from social media check-in data. *PloS one 9*, 5 (2014), e97010.

[14] Zheng, Y., Liu, Y., Yuan, J., and Xie, X. Urban computing with taxicabs. UbiComp '11 (2011), 89–98.