Graph Convolutional Neural Networks for Human Activity Purpose Imputation

Graph Convolutional Neural Networks for Human Activity Purpose Imputation from GPS-based Trajectory Data

Henry Martin[*1] martinhe@ethz.ch
Dominik Bucher[*1] dobucher@ethz.ch
Esra Suel[2] esra.suel@sdsc.ethz.ch
Pengxiang Zhao[1] pezhao@ethz.ch
Fernando Perez-Cruz[2] fernando.perezcruz@sdsc.ethz.ch
Martin Raubal[1] mraubal@ethz.ch

Abstract

Automatic location tracking of people has recently become a viable source for mobility and movement data. Such data are used in a wide range of applications, from city and transport planning to individual recommendations and schedule optimization. For many of these uses, it is of high interest to know *why* a person visited at a given location at a certain point in time. We use multiple personalized graphs to model human mobility behavior and to embed a large variety of spatio-temporal information and structure in the graphs' weights and connections. Taking these graphs as input for graph convolutional neural networks (GCNs) allows us to build models that can exploit the structural information inherent in human mobility. We use GPS travel survey data to build person specific mobility graphs and use GCNs to predict the purpose of a user's visit at a certain location. Our results show that GCNs are suitable to exploit the structure embedded in the mobility graphs.

1 Introduction and Related Work

In recent years, human mobility analysis has become increasingly important since it is closely associated with opportunities (e.g., on-demand travel and shared mobility services, location-aware recommender systems) and challenges (e.g., traffic congestion, air and noise pollution, decisions on transportation infrastructure investments) in modern society. Advanced information and communication technology presents a unique opportunity to better understand highly complex mobility behavior [19]. Specifically, the prevailing availability of Global Navigation Satellite System (GNSS) based movement trajectory data allows low cost collection of trajectory data from very large numbers of participants over long periods of time. To take advantage of the available data, researchers both from domain sciences and machine learning are developing methods for cleaning raw data, identification of trips (i.e., movements) and activities (i.e., staypoints), and prediction of transport modes and activity types [22; 4].

Previous studies have proposed a number of methods to identify activity types from various trajectory datasets [4; 14; 13]. For instance, [16; 15] trained a random forest model on GPS trace data to improve trip purpose identification and [21; 3] used Bayesian theory-based models to infer trip purposes using taxi trajectory data. However, these studies often neglect high regularity of human activity patterns while predicting activity types and mostly rely on extracted features of activities (e.g., average stay duration, time of day, points of interest in vicinity, etc.). In this paper, we propose

[*] equal contribution
[1] Department of Civil, Environmental and Geomatic Engineering
[2] Swiss Data Science Institute (ETHZ/EPFL), Zürich, Switzerland

Workshop on Modeling and Decision-Making in the Spatiotemporal Domain, 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

a new method based on graph convolutional neural networks (GCNs) [6] for the prediction of activity types (i.e., trip purposes) from GPS trajectory data generated by personal smartphones. This allows incorporating the high regularity of human mobility in terms of locations visited at specific times of day and the frequency of observed direct trips between two locations [18]. For example, a person might prefer to go grocery shopping at their favorite supermarket after work every Friday; and such regularity in behaviour is potentially useful for inferring activity types from location data. To the best of our our knowledge, this is the first application of GCNs for the task of predicting activity purposes from trajectory data. Furthermore, we use a unique dataset that includes raw trajectories from 139 users over one year in Switzerland, in which the participants labeled each staypoint with the purpose of their visit.

2 Data and Methods

2.1 Data

We use semantically enriched tracking data from the *SBB Green Class* pilot study[1]. Here, 139 Swiss users were tracked over the course of one year with an app on their smartphone. The tracking app[2] segmented the movement data into staypoints (a user does not move out of a certain area during a given period) and movement trajectories and presented it to the users for validation and labeling. The trajectories are labeled with the used mode of transport and the staypoints with an activity purpose $l_p \in \{\textit{home}, \textit{work}, \textit{errand}, \textit{leisure}, \textit{wait}\}$. The study participants generate between 2 and 20 staypoints every day which results in between 2000 and 4000 staypoints per user over the course of the tracking period.

2.2 Methods

Activity areas. From prior work on human mobility analysis, we know that humans tend to visit a fixed set of spatial locations that evolves over time [1; 18]. To exploit the spatio-temporal patterns and structure that lies within human mobility and the sequential visits of these places, we explicitly model the human mobility behavior between these locations. Based on the point-cloud like staypoint data, we create frequently visited places which we call activity areas using the DBSCAN clustering algorithm [7; 8].

Graph network construction. To exploit the spatial information encoded in the relation between activity areas and the spatio-temporal structure that lies within the sequential visit patterns, we construct a set of person specific, undirected and connected graphs $G_{u,j} = (V_u, E_{u,j})$. We define $G_{u,j}$ as the j^{th} mobility graph of user u with a user specific, finite ($|V_u| = n_u$) set of nodes $V_u = \{V_{u,i}\},\ i \in [1, 2, ..., n_u]$, where n_u is the number of nodes per user and $V_{u,i}$ denotes a specific node of that user. The nodes are connected with m sets of weighted edges $E_{u,j},\ j \in [1, 2, ..., m]$. For every graph $G_{u,j}$ we define $W_{u,j} \in \mathbb{R}^{n_u \times n_u}$ as its weighted adjacency matrix.

To construct the mobility graphs for one user, we define every activity area as a node $V_{u,i}$ in his graphs $G_{u,j}$ (cf. Figure 1). The creation of a simplified graph by clustering (spatially) similar staypoints to activity areas can be interpreted as a coarsening step comparable to the graph coarsening step from [6]. To showcase the GCNs ability to learn on different graphs (with the same set of nodes), we use two different definitions of spatio-temporal connectivity to create the edges between the nodes. Specifically, we use the transition frequency between nodes and the spatial (Euclidean) distance between all nodes as two different sets of weighted edges. The transition frequencies are calculated by counting how often a user directly traveled between two activity areas. Other suitable definitions for the edges between nodes could be based on the travel times (e.g., one graph per mode of transport), angles, ticket prices or soft similarity measures such as cultural similarity.

Feature extraction. For each node (activity area) $V_{u,i},\ i \in (1, 2, ..., n_u)$, we extract 30 features[3] based on the aggregated information of the associated staypoints. Similarly, we assign labels to

[1] `https://www.sbb.ch/de/abos-billette/abonnemente/greenclass/ueber-sbb-green-class/pilotprojekte.html`

[2] `https://motion-tag.com/en/mobility/`

[3] *Mean stay duration*, *maximal duration*, *minimal duration*, *total number of staypoints* within $V_{u,i}$, mean *longitude* and *latitude*, average *distance to public transport stops (train, tram, bus)*, distribution of *arrival and*

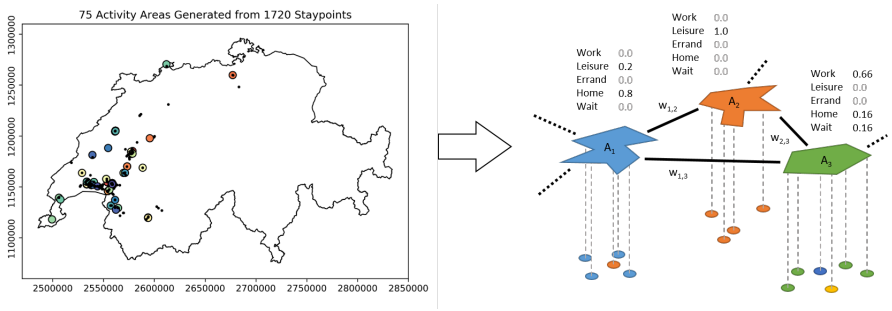

Figure 1: To create activity areas (A_i) from individual staypoints, we employ a DBSCAN clustering algorithm. Activity areas can be connected ($w_{i,j}$) in various ways, e.g., with their weights corresponding to the number of times a user directly traveled between two activity areas. The labels of the activity areas correspond to the distribution of the underlying staypoints.

the nodes based on the aggregated staypoints. This process is visualized in Figure 1 on the right. Staypoints associated with the same node can have different labels; we therefore assign a distribution over all classes to each node. The probability of label l_p at node i (short for $V_{u,i}$) is defined as $p_{i,l_p} = n_{i,l_p}/n_i$, where n_i is the number of all staypoints and n_{i,l_p} is the number of staypoints with label l_p at the node i. As the node labels are unbalanced, we introduce sample weights based on the inverse label weights $w_{l_p} = n_L/(|L| \cdot n_{l_p})$ (where n_L is the cumulative weight of all labels and n_{l_p} is the cumulative weight for a single label $l_p \in L$). These are later used to compute a weighted cross entropy loss as the cost function. Out of the 136 users we select a subset of 82 users with similar labeling behavior (number of staypoints per class is over 100). During the experiment, we randomly select 41 out of the 82 users for training, 5 users for validation and 36 completely unseen users for testing. Additionally, we standardize all training features by removing their mean and scaling them to unit variance and apply the standardization estimated on the training data to the test data.

Activity purpose imputation. The breakthrough success of convolutional neural networks (CNN) in many areas like image, speech, text or video processing [12; 10; 9] is based on their ability to exploit the structure that lies within the data by using self-learned instead of handcrafted features. In recent publications, the concept of graph convolutional neural networks (GCN) emerged to generalize this ability to arbitrary graph structures and manifolds [2; 6]. Applications of GCNs on traffic forecasting [5] and human action classification [20] showed the potential of GCNs to use information embedded in the graph structure and to exploit spatio-temporal structure for their task.

We follow the approach of [6] to approximate graph convolutions formulated in the Fourier domain using a truncated expansion in terms of Chebyshev polynomials. As it was done in [11], we only use the first order approximation. For readability we now consider only the mobility graphs of a single user and omit the index u, however all graph related parameters stay user specific. This results in the following propagation rule for graph convolution layers:

$$H_{G_j}^{(l+1)} = \sigma(D_j^{-\frac{1}{2}} W_j D_j^{-\frac{1}{2}} H_{G_j}^{(l)} M^{(l)}) \tag{1}$$

where l is the layer index, $H_{G_j}^{l}$ is the input from the previous layer with $H_{G_j}^{0} = X_j$ (corresponding to the input feature matrix of graph G_j), D_j is the degree matrix corresponding to the weighted adjacency matrix W_j. D_j is used to normalize the weighted adjacency matrix. $M^{(l)}$ denotes the weight matrix for the l^{th} neural network layer and $\sigma(\cdot)$ is the ReLu activation function. Note that the network parameters $M^{(l)}$ are independent of a specific graph G_j (or user). Figure 2 shows the network used in this work. We employ two graph convolutional layers on each of the two graphs (G_1 and G_2, one graph for every set of edges), whose outputs are then combined using a weighted sum:

$$H^{(S)} = H_{G1}^{(2)} M_{G1}^{(S)} + H_{G2}^{(2)} M_{G2}^{(S)} \tag{2}$$

Finally, a fully connected layer creates the label predictions $Y = H^{(S)} M^{(Y)} + b^{(Y)}$ (where $b^{(Y)}$ denotes a bias term).

departure time at staypoints (classified into *night*, *early morning*, *late morning*, *early afternoon*, *late afternoon*, *early evening*, *late evening*), and the distribution over *weekdays and weekends* (Mon-Sun).

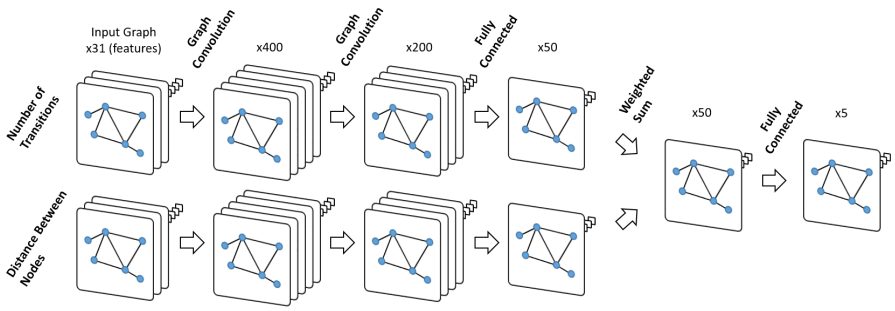

Figure 2: The computation graph used within this work: The transition and distance graphs are each convoluted twice (layer sizes of 400 and 200). Their outputs are then combined using a weighted sum (layer size 50). Finally, a fully connected layer produces the output label distributions.

3 Results

The here presented model predicts a distribution of labels for each activity area $V_{u,i}$. To measure the performance of our model we employ the earth mover's distance (EMD; also referred to as the Wasserstein metric) [17]. In essence, this metric reports the minimal shift of probability masses required to reach a given target distribution from a predicted one. Table 1 shows the results for the here introduced GCN, for a multi-layer perceptron (MLP), a random forest (RDF) classifier, a multi-output random forest regressor and as a baseline, the average label distribution for each node[4]. All classifiers are able to outperform the trivial baseline and both neural network implementations outperform the random forests which might be because the random forests can not take advantage of the additional information provided by the soft class labels.

Table 1: EMD comparison of different models.

Learning Method	EMD
GCN	**0.076**
MLP	0.097
RDF	0.146
RDF (regression)	0.146
Average Label Distr.	0.200

The GCN performs better than the MLP implementation, even though both are trained and tuned similarly. This indicates that the GCN is able to take advantage of the additional information embedded in the graphs (connections and weights) and use them to exploit local structure.

4 Conclusion and Future Work

In this work, we presented a GCN-based approach for imputing human activity purposes from GPS trajectory data. Our results show that we can use multiple personalized graphs to model human mobility behavior and to embed a large variety of spatio-temporal information and structure in their weights and connections. We could also show that we can exploit this structure using GCNs.

For our future work on the exploitation of spatio-temporal structure using graph based modeling in combination with GCNs, we plan to experiment with different graph building methods and compare them to the clustering approach applied in this work. For example, the here presented study merely considers the features of staypoints for predicting activity types. Taking into account other contextual information (e.g., points of interest or the road network) could significantly influence the predictive powers of a GCN model. Finally, we would like to analyze the impact of using more complex GCN models (e.g., use a k^{th}-order approximation instead of a first order approximation for the graph convolution).

[4]MLP: implementation in tensorflow, cross-entropy-loss, parameters: 3 layers (400,200,50), dropout: p=(0.8,0.5) between layers; RDF classif.: implementation only supports one-hotted label training, parameters: trees=500, balanced+scikit learn default; RDF reg.: output normalized that classes add up to 1, parameters: trees=500+scikit learn default. All classifiers except for the GCN only use the node features.

Acknowledgments

This research was supported by the Swiss Data Science Center (SDSC) and by the Swiss Innovation Agency Innosuisse within the Swiss Competence Center for Energy Research (SCCER) Mobility.

References

[1] Laura Alessandretti, Piotr Sapiezynski, Vedran Sekara, Sune Lehmann, and Andrea Baronchelli. Evidence for a conserved quantity in human mobility. *Nature Human Behaviour*, page 1, 2018.

[2] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[3] Chao Chen, Shuhai Jiao, Shu Zhang, Weichen Liu, Liang Feng, and Yasha Wang. Tripimputor: real-time imputing taxi trip purpose leveraging multi-sourced urban data. *IEEE Trans Intell Transp Syst*, 99:1–13, 2018.

[4] Cynthia Chen, Jingtao Ma, Yusak Susilo, Yu Liu, and Menglin Wang. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation research part C: emerging technologies*, 68:285–299, 2016.

[5] Zhiyong Cui, Kristian Henrickson, Ruimin Ke, and Yinhai Wang. High-order graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *arXiv preprint arXiv:1802.07007*, 2018.

[6] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems*, pages 3844–3852, 2016.

[7] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[8] David Jonietz and Dominik Bucher. Continuous trajectory pattern mining for mobility behaviour change detection. In *LBS 2018: 14th International Conference on Location Based Services*, pages 211–230. Springer, 2018.

[9] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1725–1732, 2014.

[10] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[11] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[13] Lin Liao, Dieter Fox, and Henry Kautz. Location-based activity recognition. In *Advances in Neural Information Processing Systems*, pages 787–794, 2006.

[14] Feng Liu, Davy Janssens, Geert Wets, and Mario Cools. Annotating mobile phone location data with activity purposes using machine learning algorithms. *Expert Systems with Applications*, 40(8):3299–3311, 2013.

[15] Lara Montini, Sebastian Prost, Johann Schrammel, Nadine Rieser-Schüssler, and Kay W Axhausen. Comparison of travel diaries generated from smartphone data and dedicated gps devices. *Transportation Research Procedia*, 11:227–241, 2015.

[16] Lara Montini, Nadine Rieser-Schüssler, Andreas Horni, and Kay W Axhausen. Trip purpose identification from gps tracks. *Transportation Research Record*, 2405(1):16–23, 2014.

[17] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. A metric for distributions with applications to image databases. In *Computer Vision, 1998. Sixth International Conference on*, pages 59–66. IEEE, 1998.

[18] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.

[19] Jorim Urner, Dominik Bucher, Jing Yang, and David Jonietz. Assessing the influence of spatio-temporal context for next place prediction using different machine learning approaches. *ISPRS International Journal of Geo-Information*, 7(5):166, 2018.

[20] Sijie Yan, Yuanjun Xiong, and Dahua Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. *arXiv preprint arXiv:1801.07455*, 2018.

[21] Pengxiang Zhao, Mei-Po Kwan, and Kun Qin. Uncovering the spatiotemporal patterns of co2 emissions by taxis based on individuals' daily travel. *Journal of Transport Geography*, 62:122–135, 2017.

[22] Yu Zheng. Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(3):29, 2015.