

Learning Urban Nightlife Routines from Mobile Data

Ada Pozo
Idiap Research Institute and EPFL
Switzerland

Thanh-Trung Phan
Idiap Research Institute and EPFL
Switzerland

Daniel Gatica-Perez
Idiap Research Institute and EPFL
Switzerland

ABSTRACT

The use of smartphone sensing for public health studies is appealing to understand routines. We present an approach to learn nightlife routines in a smartphone sensing dataset volunteered by 184 young people (1586 weekend nights with location data captured between 8PM and 4AM.) Human activity is represented at two levels, namely as the types of places visited and as the areas of the city where those places are. Routines extracted with two topic models (Latent Dirichlet Allocation and Hierarchical Dirichlet Process) are semantically meaningful and represent different moments of the weekend night, depicting activities such as pub crawling. The inference capacity of the routine representation is demonstrated with two classification tasks of value for alcohol research (alcohol consumption throughout the night, and heavy alcohol consumption.) The results suggest that nightlife routine mining could be used as a complementary tool to traditional survey-based methods in public health studies, and also inform other institutional actors interested in understanding and supporting youth well-being.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile devices**.

ACM Reference Format:

Ada Pozo, Thanh-Trung Phan, and Daniel Gatica-Perez. 2020. Learning Urban Nightlife Routines from Mobile Data. In *19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020)*, November 22–25, 2020, Essen, Germany. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3428361.3428396>

1 INTRODUCTION

There is growing interest in the use of smartphone sensing for public health studies, as these devices enable the collection of everyday life sensor data that allow to reason about people’s activities and their connection to health and well-being variables. Mobile data has been used to estimate mood and stress levels [1, 11], to assess possible alcohol consumption [3, 32, 55, 59], or smoking [45].

Routines are used by people to arrange their everyday life. As defined in health research, routines are regularities in activities and the spatial, temporal and social contexts in which these activities unfold [29]. The extraction of human routines from mobile data has gathered substantial attention [19–21, 28, 39, 49, 60]. Routines

describe people’s regular activities in a summarized manner, and highlight deviations from these trends. In addition, routines can be used for inference tasks, such as understanding the typical activities that a person may perform [4, 6].

While previous research in routine discovery has made significant progress [19–21, 28, 39, 49, 60], open questions remain regarding the use of methods for specific time periods and populations of interest. We investigate the use of location data to model young people nightlife habits. We specifically address two open issues. First, unlike most previous works, which have largely focused on daytime habits, our objective is to study routines of youth in the nightlife setting. Nightlife activities on weekends tend to have a higher degree of flexibility than on weekdays. Yet these leisure activities do not have the same regularity and have fewer frequent patterns than daytime routines do [12]. The data we analyze comes from a mobile crowdsensing experience capturing nightlife behaviour of 200+ young people [54], in which sensor data was collected between 8PM and 4AM on Friday and Saturday nights. As a second open issue, most previous works rely on manual labelling to obtain semantic information for routines [5, 8, 20, 22, 28, 60]. In contrast to this, we follow an approach to automatically represent places in two ways, namely by adding semantic information to raw location traces via place matching, and by associating location traces to regions of the city that function as mid-level descriptors of human activity.

First, our pipeline uses existing methods to extract stay points from location traces [41], and creates two place representations: one based on matching stay points to semantic venues adopting a Hidden Markov Model [7]; and another one that assigns stay points to urban areas called stay regions. In the second stage of the pipeline, these representation are fed as input to two topic models, namely Latent Dirichlet Allocation (LDA) [9] and Hierarchical Dirichlet Process (HDP) [61], employed in the past to model human behaviours [20, 28, 60], to extract nightlife routines. We show the inference capacity of the routines discovered in two classification tasks of interest for public health: consuming alcohol in a given night; and engaging in heavy drinking in a given night.

The contributions of this work are two-fold:

1. Our framework extracts routines that characterize youth nightlife behavior using real-life mobility data of over 200 study participants. Two place representations (semantic individual places and stay regions) and two topic models (LDA and HDP) are studied and compared. The discovered routines represent different periods of the night, and characterize typical behaviours, like pub crawling or going out for dinner.

2. We assess the value of routines on two classification tasks relevant to alcohol research that infer whether users drink on a given night and if so, if users engage in heavy drinking. Our method infers both tasks with 72% accuracy. For the two classification tasks,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MUM 2020, November 22–25, 2020, Essen, Germany

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8870-2/20/11...\$15.00

<https://doi.org/10.1145/3428361.3428396>

we analyse what routines and features are more discriminant, and demonstrate that the routine representation is valuable.

The paper is organized as follows. Section 2 introduces related work. Section 3 provides a description of the datasets we employed. Section 4 describes the methods for place representation. Section 5 introduces the topic model approach for routine mining. Section 6 describes the inference tasks. Section 7 presents and discusses the results. Finally, conclusions are drawn in Section 8.

2 BACKGROUND AND RELATED WORK

2.1 Basic Definitions

A **routine** is a spatio-temporal regularity that is semantically meaningful, characterized by frequent stays during concrete time periods in places with the same semantic meaning. This could involve staying at the same exact place with the same location or going to different places with the same connotations.

A **location point** is a single estimate of user location. It is characterized by its coordinates, timestamp, and accuracy.

A **stay point** is a group of consecutive location points spatially close where the user spent a period of time [41]. It is defined by the center coordinates, the timestamps when the user arrived and left, the worst accuracy of all location points forming the stay point, and the distance between its two farthest location points.

Finally, a **stay region** is a cluster of spatially close stay points that characterizes an area of a city or town.

2.2 Semantic enrichment of trajectories

One frequent step in mobility-related tasks is to enrich raw location traces with semantic information, often manually [4, 5, 8, 20, 22, 26, 28, 60], but also in an unsupervised manner [2, 23, 40, 42, 51, 65, 66, 68]. With the latter approach, traces are first divided into move or stay points. This segmentation is performed using grid-cells [2, 66], density-based clustering [38, 64], or spatio-temporal clustering [23, 40, 42, 51, 65]. In our work, we use the algorithm proposed in [42] with some changes. This algorithm uses a time-based clustering technique that detects a stay point if the user stays in a small geographic region for a period of time.

The next step is to match each point to a place. In general, an external place database is used as reference. The main approaches can be divided into geometry-based [2, 66] or probability-based [7, 8, 23, 25, 40, 65]. Geometry-based approaches associate a location point with the places it intersects with. These approaches suffer from their inability to perform place disambiguation if one point intersects with more than one place. Probabilistic-based matching addressed this problem known as place ambiguity [7], by choosing the most probable venue among the candidates of a location point, e.g. using Hidden Markov Models (HMM) [7, 8, 40, 51, 65]. In this case, the problem of matching is finding the most likely sequence of places, represented as HMM states. In our work, we adopt the place matching algorithm proposed in [7] to enrich our sequence of stay points with semantic information.

2.3 Routine discovery

There has been extensive work regarding the use of sensor data for both low-level activity discovery [43, 44, 56, 57, 69] and to extract routines [8, 19, 20, 27, 28, 39, 49, 60, 67]. Early approaches

used Principal Component Analysis [19, 24, 30]. More recent work [13] used a neural network based on the Continuous Bag-of-Words model to handle the semantics of a point of interest at different intervals of the day, clustering these representations to derive different weekday and weekend habits. Lv *et al.* [39] define location matrices per day (each entry indicating how long each user stays in a concrete place during that day), and cluster them to generate routine activities that show the probability of a user being in each place over time. Likewise, Xu *et al.* [64] define matrices indicating where each user was at different timeslots, finding similarity stay patterns.

Topic models, primarily Latent Dirichlet Allocation (LDA) [8, 14, 20, 22, 28, 67], but also Hierarchical Dirichlet Process (HDP) [44, 60, 69] and others [21, 27, 49, 50, 63] have been used to discover user activities and routines. Farrahi *et al.* [20] represented a user’s day as a sequence of cell tower labels indicating whether the user is at home, work, or elsewhere. Words are generated by concatenating locations and attaching a time label, and fed into LDA. Huynh *et al.* [28] feed accelerometer data into LDA to find low-level activities such as “sitting”. Sun *et al.* [60] employed HDP on raw sensor data to discover routines. While our work takes inspiration of the bag-of-words model of [20], there are four differences. First, we study nightlife weekend routines, compared to most previous work. Second, we study both HDP and LDA to overcome the need to pre-specify the number of topics. Third, we add the semantic information in a fully unsupervised way using twelve place categories and also the regions of the city. Finally, we used the mined routine representations for two nightlife-related classification tasks.

2.4 Ubicomp and alcohol consumption

Several works have explored the relation between alcohol consumption and young people behaviours and location using self-reported data [16, 17, 34–36]. These papers found that heavy drinking takes place both in private places and in nightlife venues (bars, pubs, nightclubs). Other work [16, 17] found that heavy drinking nights started mainly at private places (62%) and nightlife venues (22%), finding a relation between heavier drinking and the duration of those nights. Mobile phones have also been employed to collect data about drinking behaviours via SMS [33]. Santani *et al.* [54] proposed to investigate nightlife in Switzerland in the Youth@Night mobile crowdsourcing study, which recorded different types of smartphone sensor data in the background and asked users to report drinking events using short surveys to described consumed drinks, social companions, visited places, and their ambiance.

As novelty, our work uses nightlife routines as a representation of a night period to infer alcohol consumption and heavy drinking. Compared to previous work [55], we also use the Youth@Night dataset (see Section 3), but only the location sensor data is used here. No additional logs, like Bluetooth, applications, or accelerometer, are employed. Our features use solely mobility data, explored in other works only in terms of statistics about stay points (number, accuracy, duration), traveled distance, and raw speed. In our case, more complex mobility information is used in the form of routines. These routines are then used as input to infer relevant information for public health studies on alcohol consumption.

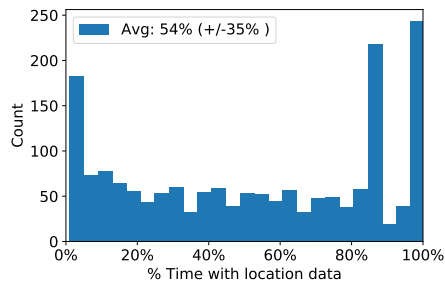


Figure 1: Distribution of the percentage of time with location information available across user nights.

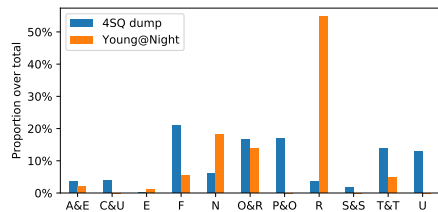


Figure 2: Place category distribution in Youth@Night and Foursquare data.

3 DATASETS

3.1 Youth@Night dataset

The Youth@Night study [54] investigated nightlife behaviour of young people in Switzerland, focusing on two cities: Zurich and Lausanne. Between September and December 2014, this study collected smartphone data from volunteers about their activities on Friday and Saturday nights from 8PM to 4AM. The study was approved by the ethical boards of the cantons of Zurich and Vaud. For the rest of this paper, we will refer to these weekend nights as user nights. All participants were young people between 16 and 25 years old. Using an Android application, participants reported drink surveys in-situ, indicating the drinks they consumed (name, size, alcoholic or not), where they did so (city and place category), and who they were with. In total, this campaign collected 2531 drink surveys from 241 participants during 2867 user nights.

In parallel, smartphone sensor data was collected in the background. Complete details on the list of sensors can be found in [54]. We focus on the location data, estimated using both GPS and GSM, and captured at high frequency for one minute, every two minutes. Every location point includes coordinates (latitude and longitude), radio of accuracy in meters (higher values indicate more uncertainty about users’ location) and timestamp. Participants could use their devices freely, and thus turn off their GPS or wireless sensor. This implies that location is sparse and with frequent gaps. Fig. 1 shows the distribution of the percentage of time when location was available per user night. It is clear that the dataset comprises both user nights with a high percentage of location data, and user nights with little data. In total, there is location for 1586 user-nights coming from 184 unique participants.

3.2 Foursquare dataset

The database used to semantically enrich location traces consists of Foursquare (4SQ) venues in Switzerland [47], as 169K 4sq venues retrieved in 2016. Each venue is assigned to a category according to a hierarchy tree¹. The list includes Arts & Entertainment (A&E); College & University (C&U); Event (E); Food (F); Nightlife Spot (N); Outdoors & Recreation (O&R); Professional & Other Places (P&O); Residence (R); Shop & Service (S&S); Travel & Transport (T&T); and Unknown (U). Note that Foursquare considers college residences to be under the College & University category, and hotels to be under Travel & Transport. We use such categories to label our venues, and if we do not have a category for a venue, we categorize it as Unknown. We filter this dataset to remove venues non-related to nightlife, with 128K venues remaining after filtering.

Fig. 2 show the distribution of venues across categories for the Foursquare data. Compared to the distribution of places reported by users in the surveys it is clear that there is a significant difference. This is especially important for the Residence venues, which represent over 50% of the places in Youth@Night surveys and are less than 5% of the Foursquare data. Information about the users’ home was directly retrieved from the mobile survey, rather than using temporal statistics [7, 8, 65].

We also use data from Google Places to obtain semantic labels for particular cases, according to the following procedure:

1. Propagation from surveys: We intersect the places where users reported to be in surveys and the extracted stay points. For each drinking event occurring during a stay point, we add a venue to our place dataset. The location of the venue is the stay point area, whereas the category is the one reported by the participant. On the surveys, users could match their location to a Foursquare venue, or simply indicate the type of place. Whenever the Foursquare venue was available, we query the Foursquare API² to obtain the venue information. If this was not available, we use the manual tagged type. With this approach, we added 1117 places to the database.

2. Adding Google Places: Since 4sq tends to have more information about cities than for rural areas, it was found that several stay points tended to remain unmatched. To reduce this, we queried Google Places API³ to find additional nightlife related venues. Because the location dataset is from 2014, and this matching was performed in 2019, we analysed how much nightlife related places actually changed since the time when the Foursquare venue database was collected. We concluded that the differences are not significant enough, and thus we assume that the venues are mostly similar. It is important to notice that we are not interested in specific venues and rather focus on venue categories, i.e., a food-related place that changed names but continued to be a food place remains the same in terms of category. We query Google Places API with a radius of 150m around stay points that do not have any possible candidates. We only consider nightlife related places, ignoring shops or other daytime venues. With this approach we find 1065 places. We convert Google types to 4SQ categories using a correspondence table (not shown for space reasons). With this, the final place dataset has 130k venues.

¹<https://developer.foursquare.com/docs/resources/categories>

²<https://developer.foursquare.com/docs/api>

³<https://developers.google.com/places/web-service/intro>

4 SEMANTIC PLACE REPRESENTATION

4.1 Stay Point Extraction

This is the first step to process the location data. We use the method in [42], with slight changes. This algorithm uses spatio-temporal clustering to obtain stay points given a list of consecutive location points. The algorithm finds a stay point when three constraints are met:

1. The maximum distance between any two location points is not larger than a distance threshold D_{max} . This value limits how big the areas covered by stay points can be. A large threshold will tend to merge different places, whereas a smaller one risks dividing a place visit into several small visits.

2. The time difference between the first and the last location points is larger than a minimum time threshold T_{min} . In other words, the user stays at least T_{min} in the stay point.

3. The time difference between consecutive location points is smaller than T_{max} . This ensures that the location points under consideration are close in time. This is particularly important for the dataset at hand, since it is frequent to find gaps in the location information. Thus, we need to ensure that no stay points are found using location points with great separation in time. For example, users may turn the GPS on when they leave a bar and turn it off when they reach their destination (e.g. a club). A stay point should not be found if users turn the GPS back on again several hours later when they want to move somewhere else, since we cannot know if they stayed there the whole time.

A stay point is defined by its center coordinates, the timestamps when the user arrived and left, the worst accuracy of all points forming the stay point, and its radius (distance between its two farthest location points.)

4.2 Place Representation 1: Place Matching

Extracted stay points are then matched to concrete venues. This is a difficult task due to the inherent ambiguity of stay points to indicate that users could be associated to several candidate places. To mitigate this, we adopt the algorithm in [7] for place matching, based on two steps: candidate place generation and place matching.

Candidate place generation. For each stay point, we obtain a list of candidate places where the user may have been. We consider all places within a radius r of the center of the stay point. If r is larger, we have more candidates and more stay points are matched. However, the larger the radius, the less likely the user was really there. Therefore, a small r results in more unmatched venues but more confident matching.

Place matching using HMMs. We then match a stay point to the best candidate. To deal with place ambiguity, the work in [7] proposed an HMM, describing a series of hidden states (candidate places) that emit observations (in our case, stay points). An HMM is defined by the initial, transition, and emission probabilities. The initial probabilities for a user to start a night in one of the possible places is computed as the fraction of candidate places with each 4SQ category. Transition probabilities of moving between each 4SQ category are computed between consecutive candidates. Finally, for the emission probabilities of 4SQ registering some coordinates if the user was in that place, each candidate place has its own emission probability, favoring venues overlapping with the accuracy radius

of the location point, with the maximum emission given when the stay point or the place area are contained in one another.

Once the the HMM is learned, the Viterbi algorithm is used to find the most likely sequence of places [62]. Given the probabilistic nature of HMMs, if one of the stay points has no candidates, none of the following stay points of the night can be matched either. We address this issue by resetting the Viterbi algorithm when a place has no candidates and starting again in the next stay point. This does not add much noise to the matching process, as location information is already missing throughout the night.

4.3 Place Representation 2: Stay Regions

Given the inherent ambiguity of place matching, mistakes occur in the process, e.g., considering that a user was in a bar instead of on the street next to the bar. This matters for our task as we focus on both nightlife and the particular demographics of young people, for whom the places they go to might not always match the venues found in Foursquare or Google Places. To mitigate the effects produced by inaccurate place labeling, we also investigate the use of stay regions to describe a city. Our alternative is to define user-visited regions that have semantic meaning, e.g. lakeside, residential area, etc. This representation reduces the ambiguity of concrete matched venues where the user is, while providing an idea of the types of places users visit. We implement this task in two steps:

Cluster stay points into stay regions. We use hierarchical clustering to derive stay regions from stay points. We use stay points instead of 4SQ venues since they represent better the areas of interest of the studied population. We employ average linkage and the spatial distance between the centre of stay points as distance measure. A stay region is created among stay points with a distance smaller than 500 m. Note that although the centre of the stay points is considered for the distance, stay regions are defined using the area covered by the stay points with their radius. Therefore, they may overlap with one another.

Obtain stay region type. This assigns a semantic label that represents the type of a region. For this purpose, we consider the 4SQ categories associated with the stay points inside the cluster. We count the number of stay points per category and then apply TF-IDF to consider infrequent categories as more relevant. To do this, we make an analogy and consider each cluster as a document and the categories as the vocabulary. Finally, we choose the categories that constitute 70% of the stay points in our cluster as the type of that cluster. This results in clusters typically represented by 1 to 3 4SQ categories. They are denoted as $4SQ_Cat_1 + \dots + 4SQ_Cat_n$, where each $4SQ_Cat$ is an abbreviation of a 4SQ category.

We implement the stay region extraction using Python's library "scipy" hierarchical clustering [31] and "scikit-learn" TF-IDF [46].

5 TOPIC MODELS FOR ROUTINES

With user nights represented as either sequences of places (Representation 1) or regions within the city (Representation 2), we aim to discover routines that describe these nights using topic modelling. Topic models have been successfully used to characterize text collections through unsupervised learning. Each document is described as a mixture of topics and in turn, each topic is also described as a mixture of words from a fixed vocabulary. Topic models assume

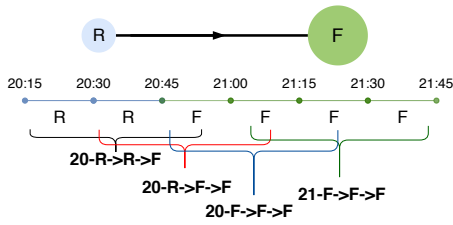


Figure 3: Example of a user night and BOW representation, using place representation 1 (place matching).

that a document is generated by first selecting a topic from the topic distribution and then choosing a word from the word distribution of that topic. Using this procedure, topic models have been used to discover activity and routine patterns [8, 20, 28, 60]. In our case, an analogy is made where a document is a night of user data, words are sequences of locations, and topics represent routines.

We first recall the two methods we used, namely Latent Dirichlet Allocation (LDA) and Hierarchical Dirichlet Process (HDP), and then describe how documents and words are created.

5.1 Latent Dirichlet Allocation (LDA)

LDA samples documents and words from Dirichlet distributions [9]. Each document has its own mixture of topics. The number of topics is known and has to be fixed beforehand. In our work, we use the topic coherence C_v to select the number of topics [53]. Topic coherence is a measure of the interpretability of the topics. We use Gibbs sampling for inference, with Mallet’s wrapper implementation in Python [52]. The two hyperparameters α_{LDA} and β are set to $\alpha_{LDA} = K/50$ and $\beta = 0.01$, where K is the number of topics. A lower value of α_{LDA} implies that fewer topics are used to model documents. Likewise, the lower the value of β , the fewer words per topic.

5.2 Hierarchical Dirichlet Process (HDP)

One key issue with LDA is that the number of topics has to be defined beforehand. HDPs solve this problem by allowing the number of topics to be infinite a priori through sampling from Dirichlet Processes [61]. During inference, the appropriate number of topics for the current documents is found. Thus, it can be considered a nonparametric generalization of LDA, where the number of topics is inferred by the model. We use Wang *et al.*’s C++ implementation with Gibbs sampling [10]. Concentration parameters γ and α_{HDP} are both set to 1, the default values. γ takes the role of α_{LDA} in LDA, whereas α_{HDP} has a similar role as β .

5.3 Bag-of-word representation (BOW)

We represent each user night as a document. Following the BOW model of [20], we represent words as a sequence of locations with a time label. To construct words, we first create a vector per night with the location of the user every 15 minutes, when that information is available. If a user was in more than one type of place in that period, the place where he spent the longest time is used. The two place representations described in the previous section can be used for this purpose. Then, we use a sliding window over each night

Table 1: Ground truth proportions for classification tasks. The number of user nights in each class and the corresponding percentage (in parenthesis) are shown.

	Drinking alcohol	Heavy drinking
Yes	629 (53%)	412 (39%)
No	538 (47%)	649 (61%)

that concatenates each three consecutive location symbols. A time label is attached to preserve the sequence of locations. We divide the night in 8 timeslots of 1 hour each and attach the corresponding time label to the sequence of places. Despite topic models ignoring word order, this allows to keep track of when the events happen while at the same time avoiding vocabulary explosion, due to a too-fine-grained representation. The final word representation has one time label and at most three location labels with the format Time-Loc1-> Loc2->Loc3. An example of the word creation process using place Representation 1 (place matching) is depicted in Fig. 3: consider a user who at 20:15 is home and then goes out to a Food venue from 20:45 until 21:45. This interval of the night is represented as RRRFFF and the corresponding words would be 20-R->R->F, 20-R->F->F, 20-F->F->F, 21-F->F->F.

6 INFERENCE TASKS AND FEATURES

This section defines the tasks used to assess the inference power of the extracted routines, and the manually engineered features related to mobility used in conjunction with such routines.

6.1 Definition of Tasks

We study whether the routine representation is informative in inference tasks relevant for public health. We investigate two binary classification tasks, namely whether the user consumes alcohol on a given night, and whether that consumption is heavy. Table 1 shows the ground truth proportions for each task. Due to missing information, we have a different number of samples in each task.

Alcohol Drinking. This binary task infers whether the user drank alcohol on a given night, as in [55]. As mentioned in Section 3.1, users reported the drinks consumed using “in-situ” drink surveys and “retrospective” forgotten drink surveys. We consider the surveys related to alcoholic beverages and label a user night as “drink alcohol” if people consumed at least one alcoholic drink.

Heavy drinking. This binary task infers whether the night involved heavy drinking, as recently proposed in [48]. From established alcohol research, users are said to have drunk heavily in a night if a man drank more than 5 standard drinks, or a woman consumed more than 4 standard drinks [15]. A standard drink is an alcoholic beverage containing more than 10 grams of pure alcohol. The number of grams in the alcohol consumed is computed as $number-of-grams = size [ml] * content [\%] * 0.793 [g/ml]$, where $size$ is the size of the drink, $content$ is the alcoholic percentage of the drink, and 0.793 denotes the density of alcohol. Both the size and the type of drink were obtained using the Youth@Night surveys. Both tasks have potential relevance in public health studies [58].

6.2 Features

The primary features we use for classification are the topic distributions across user nights. We consider the 2 place representations: the 4SQ categories associated with stay points (representation 1, referred to as SP), and the type of stay region each stay point is in (representation 2, referred to as SR). We use the topics obtained with LDA (30 topics) and HDP (48 topics for SP, 58 for SR).

In addition, we extract handcrafted features at the user-night level to improve the performance and show other aspects of mobility not reflected by routines:

- 1–16 5-bin histograms of the time between stay points, duration and accuracy of the user night.
- 17–18 Times of the first and the last stay points of the night.
- 19–20 Number of unique venues and stay regions visited during the night.
- 20–21 Number of unique 4SQ categories and stay region types visited during the night.
- 22 Percentage of time when users are stationary, i.e. their speed is smaller than 1 km/h.
- 23 Percentage of time when location data is available.
- 24 Distance travelled during the night.

These features add complementary information to the topic representation. For example, the number of stay regions and venues indicate whether the user was static or moved frequently. When we use these features (referred to as handcrafted) we standardize all of them to the $[-1, 1]$ range.

7 RESULTS AND DISCUSSION

We first present the results of semantic place representation on the nightlife data. Then, experiments are conducted on the routines extracted with LDA and HDP. The routine classification performance is presented for the two classification tasks. Finally, we discuss the results and limitations of our work.

7.1 Semantic place representation

Stay point extraction. We set the minimum duration of a stay point to $T_{min} = 5$ minutes, the maximum time between stay points to $T_{max} = 4$ hours and the maximum distance from the center of the stay point to any other point to $D_{max} = 100$ m. [18]. As preprocessing steps, we remove location points outside Switzerland and all location points with poor accuracy. This keeps the majority (83%) of the location data. Table 2 presents statistics of the extracted stay points. People tend to visit 1-5 places per night, staying between 1 and 2 hours. This is in line with previous work that found that people typically visit 1 to 4 places when they go out and on normal days [17, 18, 36].

Place representation 1: place matching. As mentioned in Section 4, setting the parameter r for candidate place search is a trade-off between places that cannot be matched and confidence in the matching results. We use a two-step approach to minimize the number of unmatched places. Initially, we set $r = 100$ m., which is the maximum area of our stay points. Then, for the places with no candidates we do a second candidate search with $r = 150$ m.

To verify place matching performance, we use the venues propagated from the drink surveys as ground truth and evaluate the performance using 4-Fold Cross-Validation (CV). For each fold, we

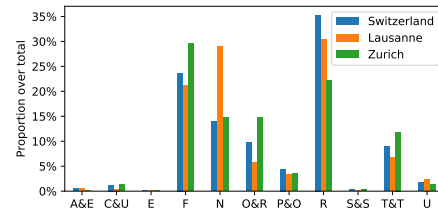


Figure 4: Distribution of the proportion of stay points matched to 4SQ categories across venues. Switzerland refers to all matched stay points, Lausanne and Zurich refer to stay points inside the limits of each city.

include 75% of the propagated places in the places dataset, and the other 25% for verification. The obtained weighted average of F1-Score across all 4sq categories is $F1 = 0.42$. The best performing category was Residence, with $F1 = 0.59$. This is probably due to having more samples of this category on average for each fold. The overall place matching performance is explained by the difficulty of the task, which is an unresolved issue in the literature [7, 40]. Most previous works rely on large datasets to alleviate it [7, 23, 40, 51]. In our case, we use real-life nightlife data that has not been studied in the past in as much depth as daytime, and is thus less represented in place datasets as we show here for the Swiss context.

Finally, we run the algorithm with all the propagated venues. Only 121 (2%) stay points remain unmatched to any venue. In Fig. 4, the label "Switzerland" depicts the distribution of 4SQ categories across the matched stay points. As expected with the places reported in surveys (Fig. 2), Residence is the most frequent label, followed by Food and Nightlife Spot. A random data sample shows a large variation of missing data across nights.

Place representation 2: Stay regions. SRs allow a description of the city that gives a more general idea of where people are. This approach loses some of the semantics associated to knowing the venue type where the participant goes to, but in return it offers other information that may indicate, for example, whether the user is in the city centre or the suburbs, in a going out area or in a quieter region, etc.

The clusters found in the areas of Lausanne and Zurich, where our study on nightlife is based, can be summarized as follows. Lausanne is a small city that has one main area to go out, Lausanne-Flon, covered mostly by stay region #1 (not shown). As a consequence, young crowds mostly move around this area and the clusters tend to overlap there. Stay region #2 shows one of the characteristics of Swiss cities: on the lower floors of buildings there is commerce, whereas on the upper floors there are flats where people live. Stay region #3 covers the main train station and the surrounding area, filled with cafés. Lastly, stay region #4 is a cluster depicting a residential neighborhood of high density. On the other hand, Zurich is a bigger city. Thus, one can find more clusters where people go out, each with a different activity. As in Lausanne, we have a stay region #1 covering the main train station, which characterizes the activities in the cluster with its public transport and many restaurants. Stay region #2 covers the old town, and a pedestrian area that mostly has bars. It also covers the front part of the technical

Table 2: Statistics of extracted stay points.

#Users	#User Nights	Average	Total	#Stay Points / User night			Stay point duration		
		#Nights / User	# Stay points	Mean	Std	Median	Mean	Std	Median
184	1586	8.61	4747	2.99	2.64	2.0	1h. 22 min.	1h. 44 min.	36 min.

university, where there are open spaces to gather. Another area of interest is stay region #3, which is by the lake and where typically young people drink. Stay region #4 is normally a commercial and financial district, but that also has bars and restaurants open in the evening. Finally, stay region #5 covers what once was the red-light district of Zurich. This area is now a popular zone where young people often go out.

7.2 Routine extraction with topic models

This section presents the routines obtained using LDA and HDP, where each user night is described as a document and the locations at different times are the words. For the rest of this paper, we will refer to the models that use the stay points and their matched places as words with SP, and to models that use the sequence of stay regions with SR.

SP Representation. The routines for both LDA and HDP can be summarized as follows.

LDA Routines. We use topic coherence to select the optimal number of latent topics with the LDA model. Using the Elbow method, a value of $K = 30$ topics is selected.

Some of the topics found are:

- Topic 5 and 7 represent being on a Nightlife Spot around 20:00-23:00 and 23:00-01:00, respectively.
- Topic 28 models going out for dinner around 20:00-22:00.
- Topic 11 and 6 model being at a Residential venue at 01:00-03:00 and 21:00-23:00, respectively.
- Topic 2 models being at a Professional & Other Places venue (e.g. parking lot) at any point during the night.
- Topic 23 models being at a Travel & Transport place throughout the night, mostly in its first half until midnight.
- Topic 29 models being around outdoor places towards the second half of the night (22:00-03:00).

Latent topics found by LDA are able to capture low-level routines that represent different parts of the user nights, despite the noise present on the semantic labels. This shows that a user night is in general described as a mixture of topics that captures the overall trends. They are defined in general by around 2-3 words, indicating they are quite polarized. Furthermore, the time intervals that routines refer to are clearly marked. Venues that are more frequent (Residence, Nightlife Spot, etc.) have different routines referring to them at distinct times, while less frequent venues, such as Travel & Transport are more spread out during the night.

HDP Routines. Next, we find routines using HDP. This model finds 48 topics:

- Topics 0 and 3 model being at a residence in the first and second half of the night, respectively.
- Topic 1 models being out in a food-related place, especially around 21:00-22:00 (e.g. going out for dinner.)

- Topic 5 models being in an outdoors place in the first half of the night and until 01:00, with less strength after.
- Topic 4 models activities in the second half of the night, either on a transport related venue or in a nightlife venue.
- Topic 6 models being at any hour at night around a Professional & Other Places venue, e.g. a parking lot.
- Topic 7 represents the routine being in a College & University-related place early (20:00) and late at night (01:00-03:00).
- Topic 34 models the routine being at a Travel & Transport place in the first hours of the night (20:00-22:00).

Similar to LDA topics, the topics found by HDP are interpretable and represent low-level routines capturing different parts of the night. Nevertheless, the topics are less polarized, needing at least 7 words to describe them. Consequently, topics tend to be more spread out in time.

SR Representation. The routines for both LDA and HDP for the SR representation can be summarized as follows.

LDA Routines. As before, we use topic coherence to choose the correct number of topics with LDA. $K = 40$ topics is chosen using the Elbow method.

Topics found tend to be quite spread out over time although stay regions types are correctly grouped together:

- Topic 0 shows the routine of being in a region composed of Outdoors & Recreation and Residence venues around 21:00-00:00. It also represents, although less strongly, being in the same area in the second half of the night.
- Topic 5 represents the routine of being in a mixed region composed of Food, Professional & Other Places, and Nightlife Spot places in the second half of the night.
- Topic 7 represents the routines of being in a residential area from 23:00 to 01:00.
- Topic 30 represents the routine of being in a nightlife and food-related area in the second half of the night (23:00-02:00).

Compared to the stay point level (SP) the routines found tend to be quite less polarized in terms of time, spanning more hours. This might be because the vocabulary is sparser, but also because fewer stay regions are visited each night in comparison to individual venues. Topics also seem to be representative of fewer user nights, but when they are, it is just as strongly as before.

HDP Routines. HDP finds 58 topics with documents at the SR level. This is an increase of 10 topics with respect to the SP level, and is explained by the lower concentration of user nights due to the bigger vocabulary. Some of the found topics are:

- Topic 0 represents the routine of being in a residential area at any point through the night.
- Topic 11 represents the routine of being in a mixed region composed of Nightlife Spot, Outdoors & Recreation, and Unknown venues in the second half of the night.

- Topic 14 represents the routine of being in a region composed mostly by Residence and Nightlife Spot venues around 21:00-01:00.
- Topic 17 represents the routine of being in a region composed of Nightlife Spot and Food-related venues between 21:00 and until 01:00.

In this case, the biggest difference that can be observed with respect to stay point level is in terms of Topic 0. Whereas before we had different topics related to being early or late in residences, now there is only one routine covering the whole night in a residential area.

7.3 Results of classification tasks

In this section, we present the results of user-night-level classification of (1) alcohol consumption, and (2) heavy drinking.

All user nights with ground-truth information for each task are used, with no filtering of cases where we have almost no location information. This implies that we make inferences both on nights with sufficient amount of location information and on nights with significantly less data (see Fig. 1). Concretely, considering our BOW model, the maximum number of words per document is 16. 25% of nights have 4 words or less and 25% have 13 or more, which illustrates how varied the data is.

We use 90% of the data to train and the remaining 10% as test set. To tune the parameters for each model we use 5-fold Cross-Validation (CV) on the train set. As explained in Section 6, each user night is represented in two ways: with the topic proportions only (Routines), and with the topic proportions and the handcrafted mobility features (Routines + HC). Routines are mined either describing the user night as a sequence of stay points (SP) or stay regions (SR) with both LDA and HDP, which gives a total of four possible combinations. It is worth reminding that features from SP indicate the concrete type of venues but give no indication of the area of the city, whereas SR indicate the area of the city with no information on the concrete place and its semantics. In addition, HDP tends to obtain more general routines that span several hours, while LDA tends to give routines happening on specific hours.

We evaluate the performance of the routine features to infer the correct class using three different classifiers:

- Logistic regression (LR), with L2 regularization ranging from 10^{-3} to 10^3 .
- Support Vector Machines (SVM), with regularization ranging from 10^{-3} to 10^3 , either RBF or linear kernel, and with RBF kernel coefficient in the range from 0.1 to 1.
- Random Forests (RF), with the number of trees ranging from 10 to 500.

Accuracy is used as the performance measure to report results and its maximum value is used to choose the best parameters with 5-fold CV. Since in some cases the datasets are imbalanced (see Table 1), subsampling of the majority class is used during training to obtain a balanced dataset. In this section, only the best result is shown for each combination of features. All classifiers are implemented in Python using "scikit-learn" [46].

Routines. The first row for each task in Table 3 (Routines) presents the results using only the topic proportions of routines as features. It can be seen that for the two classification tasks, the baseline

majority-class classifier is outperformed. It is worth noting that the best model for the binary inference of alcohol consumption uses SR with HDP routines, that is, the indicators of a user drinking seem to be more related to the areas of the city and general routines. On the contrary, SP with LDA is preferred for binary classification of heavy drinking. This suggests that heavy drinking might be associated with being in concrete types of venues at more specific hours, which matches previous findings in the alcohol research literature [17, 36]. In addition, it should be mentioned that, only with the routines, the results of inferring heavy drinking are slightly better than the inference of drinking or not on a given night.

To analyse in more depth what routines are actually capturing, we plot the most important features for the classifiers. The feature importance with SVM using RBF kernel cannot be directly obtained. Hence, we focus on the best achieving models with Random Forest and Logistic Regression. Regarding Random Forest, we obtain the Gini impurity score, which is a measure of the total decrease in node impurity weighted by the probability of reaching that node and averaged over all trees of the ensemble. With Logistic regression, we use the feature weights in the model.

The first row of Fig. 5 represents the top 5 features for the best performing classifier of each task. One can see that with heavy drinking, all routines refer to being in a Residence venue around the beginning and middle of the night. This is similar to other studies in alcohol research that found that heavy drinking nights tend to start in private places and that most heavy drinking nights take place there as well, at least partially [16, 17, 34, 36]. Regarding drinking alcohol or not, results suggest that being in a region where there is only nightlife (Topic 21) is related to alcohol consumption, whereas the rest of the topics among the top 5 indicate the opposite. It is especially noticeable, considering the heavy drinking results, that this happens with routines related to being in an area that is only residential all night (Topic 0) or early (Topic 43).

Routines and handcrafted features. The last row of each task (Routines + HC) in Table 3 shows the classification performance by training the models on both routines and handcrafted features. Overall, there is an improvement in performance when adding the handcrafted (HC) features, particularly for alcohol consumption, which achieves 72% accuracy with all classifiers. The binary inference of drinking alcohol is again better when focusing on the areas of the city (SR) with more general routines extracted with HDP, whereas heavy drinking classification is again better with specific places (SP) and routines extracted with LDA.

The second row of Fig. 5 shows the top 5 features for these models. It is clear that routines lose importance after adding the hand-crafted features, with only one or two in the top 5 for all models. For the two classification tasks, the time of the first stay point of the night seems to be a good indicator. This is likely that this is related to an early start of the night, which has been associated with drinking in the specialized literature [16]. Both classification tasks also have the number of different venues and 4SQ categories of such venues as pointers of drinking. In the literature, there are similar findings suggesting that the number of drinks consumed increase with the number of venues visited [16], and that there is a connection between the number of venues visited and drinking [17, 36]. It is interesting that the classifiers seem to capture this.

Table 3: Classification accuracy on the test set with different tasks for the best achieving combination of topics. Model refers to place representation used to generate routines (sequence of stay points - SP or stay regions - SR) and the topic model (HDP or LDA). Best performance in each task across all experiments shown in *italics*. LR - Logistic Regression, SVM - Support Vector Machines, RF - Random Forest Classifier, Majority - Baseline naive classifier that always infers the majority class.

	Features	Model	LR	SVM	RF	Majority
Drink Alcohol	Routines	SR - HDP	0.658	0.667	0.590	0.53
	Routines + HC	SR - HDP	<i>0.718</i>	<i>0.718</i>	<i>0.718</i>	
Heavy Drinking	Routines	SP - LDA	0.664	0.673	0.682	0.61
	Routines + HC	SP - LDA	<i>0.720</i>	0.673	0.664	

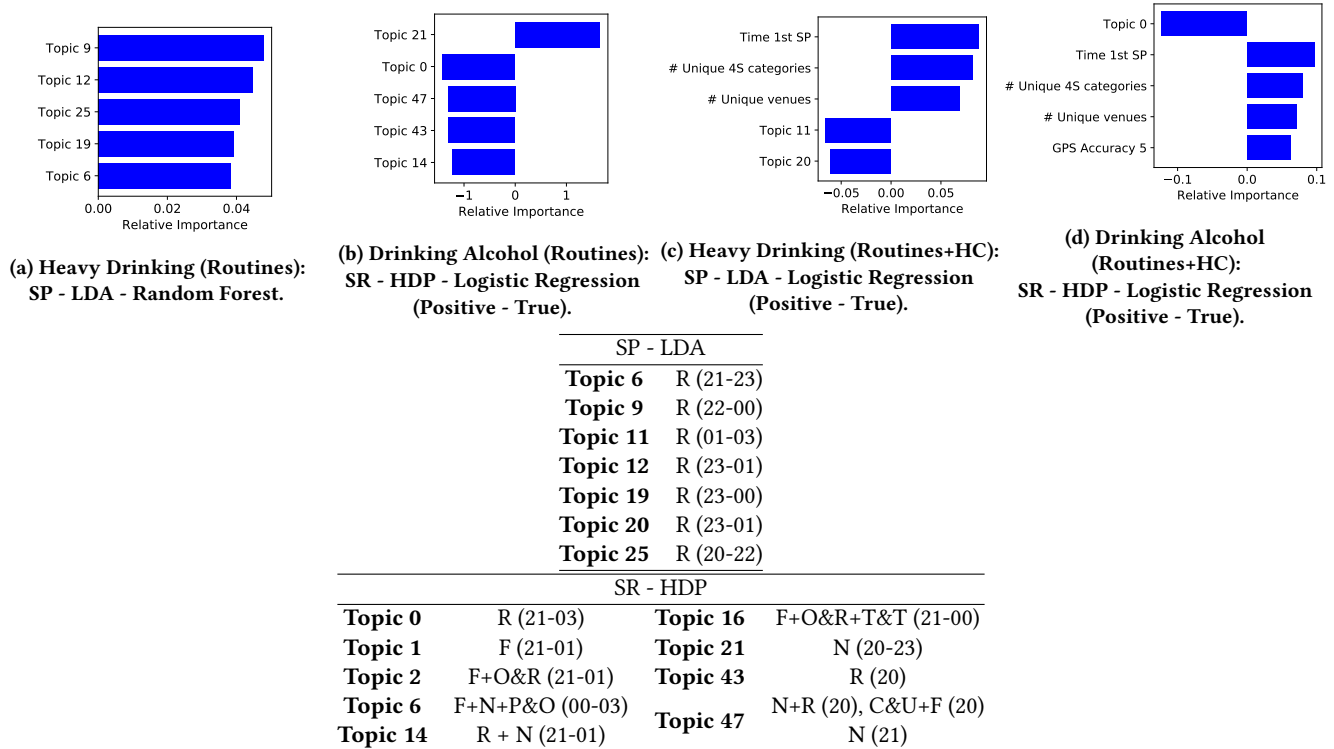


Figure 5: Feature importance for the classification tasks. The model used to obtain the importance is shown in each case. (a), and (b) depict the results using only the topic proportions as features. (c), and (d) depict the results with topic proportions and handcrafted features. The description of the topics is shown over this caption. The numbers in parentheses correspond to time intervals according to hour of the day, from 00 (midnight) to 23 (11PM). With heavy drinking, both kinds of features prefer SP and LDA topics, while for the other task SR and HDP perform better. The relative importance with Logistic Regression refers to the weight of each feature and with Random Forest to the Gini impurity score.

7.4 Discussion

We finish this section by discussing the results and their implications, as well as the limitations of our work.

Place representation. We first enriched the location traces with information about the type of venues likely visited by people and in which area of the city they were. This only considered the locations where people stayed for a certain time as relevant, disregarding periods when they were moving. This unused information could indicate important aspects such as transportation means (walking,

car, train) or short- vs. long-term travel. Integrating this information to describe a trajectory remains an open research problem [37].

After stay point segmentation, we were able to match almost all stay points to venues. We evaluated this matching both by manual inspection and with an experiment. The results indicated that exact matching is a difficult task. In addition to the difficulty of unsupervised matching in general, the performance can be explained by the nature of the data. Study participants were fully free to turn off the location sensors and their full device, which frequently resulted in gaps in the location signal. This implies shorter sequences and less accuracy, known to be difficult conditions for matching [23, 51]. In

addition, it was not possible to find a place dataset that matched the distribution of venues visited by the studied population. This is also known to complicate the place matching task, as results are related to the availability of places around stay points [23]. A way to improve matching could use the typical time of stay [40]. Nevertheless, since large-scale temporal information is needed to generalize well, a study of this possibility is left for future work. Finally, we studied an alternative to this problem with the use of stay regions to represent user nights. Despite the loss of more precise information, with the stay region approach we bypassed the noise introduced by erroneous place matching and kept semantic information about likely activities performed by the user. Stay regions showed a different trend in each city: concentrated for the small-city case of Lausanne, and more spread for the case of Zurich, where stay regions captured the larger dimensions associated to a larger metropolitan area. Importantly, both the stay point and the stay region representations were useful to produce useful inputs for routine mining and thus for classification.

Routine mining. We showed that some of the extracted routines correspond to bar hopping, late-night going out (e.g. for clubbing), or staying at home, all of which are practices reported in the social science literature of youth and nightlife [16, 17, 34, 36]. The routine representation is thus relevant in and of itself, but also useful for further inference tasks. While the method discovers meaningful nightlife routines, some limitations can be discussed. First, the extracted nightlife routines are not representative of the whole society, but only of a limited demographic: young people willing to participate in a study of this kind. Even though this limits the generality of the results, we argue that these routines still represent the trends of a population of interest for public health studies. Second, from the modeling perspective, the routines sometimes captured trends such as being in a place at a concrete time. This is partly due to the basic assumption of topic models. A possible direction would be to cluster the topic proportions across nights [60]. Finally, the routine analysis could be enriched by studying deviations from routine trends.

Classification tasks. The results in the two binary classification tasks proved that the discovered routines, despite using only location data, were discriminant to some extent. The majority class classifiers were outperformed in the two cases. Regarding the routine mining model, LDA and HDP showed to be the best performing model depending on the inference task. Regarding night-level alcohol consumption binary classification, the results matched several findings in the literature about the relation between the duration of nights and the number and types of visited places, on one hand, and the consumption of alcohol on the other hand [16, 17, 34, 36]. Regarding night-level heavy drinking binary classification, we found that specific types of places are an indicator of heavy drinking, which agrees with the findings of alcohol research studies in other world regions [17, 36, 36]. Moreover, we found that the number of stay regions and places visited is related to higher consumption of alcohol, in the same line of other alcohol research [16]. In contrast to such previous results, which were all based on retrospective surveys, in this study phone sensing data was captured automatically and in situ (which can be seen as an alternative to address some of the recall problems associated with self-reports) and modeled as routines. Furthermore, we achieved an accuracy comparable to [55]

on alcohol/no-alcohol drinking classification, and to [48] on heavy drinking classification, while using only automatically extracted mobility features.

Applications. We identify two possible applications for our work. First and foremost, our goal is the development of computational methods that can support research in public health. We showed that the discovered routines have connections with previous findings based on survey-based studies [16, 17, 34, 36]. From this perspective, our work combining smartphone sensing and topic models can be seen as bringing methodological novelty to established methods of data collection and analysis in alcohol research. In the second place, the connection of routines with the urban space also makes our work potentially useful to city offices interested in supporting both youth and a vibrant use of the public space. City stakeholders could contextualize and use the results given their understanding of the local circumstances (e.g., in Switzerland, alcohol consumption in the public space is allowed, and drinking of wine and beer is legal from the age of 16). In these applications, citizens need guarantees to trust institutions with privacy and ethical frameworks given the sensitive nature of location data.

8 CONCLUSIONS

We investigated a framework for the discovery of nightlife routines. Location traces were represented with two approaches, which resulted in the description of a night as a sequence of either venue types or of city areas. Then, the corresponding trajectories extracted from 1586 user-nights were used to mine routines representing nightlife activity patterns. Two topic models (LDA and HDP) used to extract routines were analysed and compared. The obtained routines represent different parts of a night, and model the particularities of youth activities in each time period. Each topic model extracted slightly different routine aspects, and we found that neither can be said to outperform the other. With public health applications in mind, we showed how routines can be used for two specific tasks related to alcohol intake at the night level: a binary consumption task (alcohol drinking or not), and a heavy drinking task. Comparable accuracy to more complex methods that use a combination of sensor data was reached using routine information in combination with other handcrafted features. Future work includes the use of information about moving episodes between stay points to compute means of transport like walking or taking public transport. Additionally, we would like to study higher level routines by building on top of the topic representations developed in this work.

ACKNOWLEDGMENTS

This work was funded by the Swiss National Science Foundation (SNSF) through the Dusk2Dawn Sinergia project. We thank Florian Labhart (Idiap and Addiction Switzerland) for discussions.

REFERENCES

- [1] Saeed Abdullah, Mark Matthews, Ellen Frank, Gavin Doherty, Geri Gay, and Tanzeem Choudhury. 2016. Automatic detection of social rhythms in bipolar disorder. *Journal of the American Medical Informatics Association* 23, 3 (May 2016), 538–543. <https://doi.org/10.1093/jamia/ocv200>
- [2] Luis Otavio Alvares, Vania Bogorny, Bart Kuijpers, Jose Antonio Fernandes de Macedo, Bart Moelans, and Alejandro Vaisman. 2007. A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th*

- annual ACM international symposium on Advances in geographic information systems - GIS '07. ACM Press, Seattle, Washington, 1. <https://doi.org/10.1145/1341012.1341041>
- [3] Zachary Arnold, Danielle LaRose, and Emmanuel Agu. 2015. Smartphone inference of alcohol consumption levels from gait. In *2015 International Conference on Healthcare Informatics*. IEEE, 417–426.
- [4] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K Dey. 2016. Modeling and understanding human routine behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 248–260.
- [5] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K. Dey. 2016. Modeling and Understanding Human Routine Behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). ACM, New York, NY, USA, 248–260. <https://doi.org/10.1145/2858036.2858557>
- [6] Nikola Banovic, Anqi Wang, Yanfeng Jin, Christie Chang, Julian Ramos, Anind Dey, and Jennifer Mankoff. 2017. Leveraging human routine models to detect and generate human behaviors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6683–6694.
- [7] Luke Birmingham and Ickjai Lee. 2019. Mining place-matching patterns from spatio-temporal trajectories using complex real-world places. *Expert Systems with Applications* 122 (May 2019), 334–350. <https://doi.org/10.1016/j.eswa.2019.01.027>
- [8] Nicola Bicoocchi and Marco Mamei. 2014. Investigating ride sharing opportunities through mobility data analysis. *Pervasive and Mobile Computing* 14 (Oct. 2014), 83–94. <https://doi.org/10.1016/j.pmcj.2014.05.010>
- [9] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
- [10] blei lab. 2010. Hierarchical Dirichlet Process (with Split-Merge Operations). <https://github.com/blei-lab/hdp>.
- [11] Luca Canzian and Mirco Musolesi. 2015. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, Osaka, Japan, 1293–1304. <https://doi.org/10.1145/2750858.2805845>
- [12] H. Cao, Z. Chen, F. Xu, Y. Li, and V. Kostakos. 2018. Revisitation in Urban Space vs. Online: A Comparison across POIs, Websites, and Smartphone Apps. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* 2, 4 (2018), 156:1–156:25. <https://doi.org/10.1145/3287034> This work was supported in part by the National Key Research and Development Program of China under grant 2017YFE0112300, the National Nature Science Foundation of China under 61861136003, 61621091 and 61673237, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology.
- [13] Hancheng Cao, Fengli Xu, Jagan Sankaranarayanan, Yong Li, and Hanan Samet. 2019. Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population. *IEEE Transactions on Mobile Computing* (2019).
- [14] Federico Castanedo, Diego López de Ipiña, Hamid K. Aghajan, and Richard Kleihorst. 2014. Learning routines over long-term sensor data using topic models. *Expert Systems* 31, 4 (Sept. 2014), 365–377. <https://doi.org/10.1111/exsy.12033>
- [15] Kelly E. Courtney and John Polich. 2009. Binge drinking in young adults: Data, definitions, and determinants. *Psychol Bull.* 135, 1 (2009), 142–156.
- [16] Paul Dietze, Paul A Agius, Michael Livingston, Sarah Callinan, Rebecca Jenkinson, Megan SC Lim, Cassandra JC Wright, and Robin Room. 2017. Correlates of alcohol consumption on heavy drinking occasions of young risky drinkers: event versus personal characteristics. *Addiction* 112, 8 (2017), 1369–1377.
- [17] Paul M Dietze, Michael Livingston, Sarah Callinan, and Robin Room. 2014. The big night out: What happens on the most recent heavy drinking occasion among young Victorian risky drinkers? *Drug and alcohol review* 33, 4 (2014), 346–353.
- [18] Trinh Minh Tri Do and Daniel Gatica-Perez. 2013. The places of our lives: Visiting patterns and automatic labeling from longitudinal smartphone data. *IEEE Transactions on Mobile Computing* 13, 3 (2013), 638–648.
- [19] Nathan Eagle and Alex Sandy Pentland. 2009. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology* 63, 7 (2009), 1057–1066.
- [20] Katayoun Farrahi and Daniel Gatica-Perez. 2011. Discovering routines from large-scale human locations using probabilistic topic models. *ACM Transactions on Intelligent Systems and Technology* 2, 1 (Jan. 2011), 1–27. <https://doi.org/10.1145/1889681.1889684>
- [21] Katayoun Farrahi and Daniel Gatica-Perez. 2014. A probabilistic approach to mining mobile phone data sequences. *Personal and Ubiquitous Computing* 18, 1 (Jan. 2014), 223–238. <https://doi.org/10.1007/s00779-013-0640-8>
- [22] Laura Ferrari and Marco Mamei. 2011. Discovering daily routines from Google Latitude with topic models. In *2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE, Seattle, WA, USA, 432–437. <https://doi.org/10.1109/PERCOMW.2011.5766928>
- [23] Barbara Furlletti, Paolo Cintia, Chiara Renso, and Laura Spinsanti. 2013. Inferring human activities from GPS tracks. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing - UrbComp '13*. ACM Press, Chicago, Illinois, 1. <https://doi.org/10.1145/2505821.2505830>
- [24] Gabriel Goulet-Langlois, Haris N. Koutsopoulos, and Jinhua Zhao. 2016. Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C: Emerging Technologies* 64 (March 2016), 1–16. <https://doi.org/10.1016/j.trc.2015.12.012>
- [25] Qihang Gu, Dimitris Sacharidis, Michael Mathioudakis, and Gang Wang. 2017. Inferring Venue Visits from GPS Trajectories. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - SIGSPATIAL '17*. ACM Press, Redondo Beach, CA, USA, 1–4. <https://doi.org/10.1145/3139958.3140034>
- [26] Baris Guç, Michael May, Yucel Saygin, and Christine Körner. [n.d.]. Semantic annotation of GPS trajectories.
- [27] Samiul Hasan and Satish V. Ukkusuri. 2014. Urban activity pattern classification using topic models from online geo-location data. *Transportation Research Part C: Emerging Technologies* 44 (July 2014), 363–381. <https://doi.org/10.1016/j.trc.2014.04.003>
- [28] Tân Huynh, Mario Fritz, and Bernt Schiele. 2008. Discovery of activity patterns using topic models. In *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*. ACM Press, Seoul, Korea, 10. <https://doi.org/10.1145/1409635.1409638>
- [29] Margaret M. Jastran, Carole A. Bisogni, Jeffery Sobal, Christine Blake, and Carol M. Devine. 2009. Eating routines. Embedded, value based, modifiable, and reflective. *Appetite* 52, 1 (2009), 127 – 136. <https://doi.org/10.1016/j.appet.2008.09.003>
- [30] Shan Jiang, Joseph Ferreira, and Marta C. González. 2012. Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery* 25, 3 (Nov. 2012), 478–510. <https://doi.org/10.1007/s10618-012-0264-z>
- [31] Eric Jones, Travis Oliphant, Pearu Peterson, et al. 2001–. SciPy: Open source scientific tools for Python. <http://www.scipy.org/> [Online; accessed May 2019].
- [32] Hsin-Liu Cindy Kao, Bo-Jhang Ho, Allen C Lin, and Hao-Hua Chu. 2012. Phone-based gait analysis to detect alcohol usage.. In *UbiComp*. 661–662.
- [33] Emmanuel Kuntsche and Florian Labhart. 2012. Investigating the drinking patterns of young people over the course of the evening at weekends. *Drug and alcohol dependence* 124, 3 (2012), 319–324.
- [34] Kypros Kypri, Mallie J Paschall, John D Langley, Joanne Baxter, and Beth Bourdeau. 2010. The role of drinking locations in university student drinking: Findings from a national web-based survey. *Drug and Alcohol Dependence* 111, 1-2 (2010), 38–43.
- [35] Kypros Kypri, Mallie J Paschall, Brett Maclelland, and John D Langley. 2007. Intoxication by drinking location: A web-based diary study in a New Zealand university community. *Addictive Behaviors* 32, 11 (2007), 2586–2596.
- [36] Tina Lam, Simon Lenton, Rowan Ogeil, Lucinda Burns, Alexandra Aiken, Tanya Chikritzhs, William Gilmore, Belinda Lloyd, James Wilson, Dan Lubman, et al. 2017. Most recent risky drinking session with Australian teenagers. *Australian and New Zealand journal of public health* 41, 1 (2017), 105–110.
- [37] Miao Lin and Wen-Jing Hsu. 2014. Mining GPS data for mobility patterns: A survey. *Pervasive and mobile computing* 12 (2014), 1–16.
- [38] Yaqiong Liu and Hock Soon Seah. 2015. Points of interest recommendation from GPS trajectories. *International Journal of Geographical Information Science* 29, 6 (June 2015), 953–979. <https://doi.org/10.1080/13658816.2015.1005094>
- [39] Mingqi Lv, Ling Chen, and Gencai Chen. 2013. Mining user similarity based on routine activities. *Information Sciences* 236 (July 2013), 17–32. <https://doi.org/10.1016/j.ins.2013.02.050>
- [40] Mingqi Lv, Ling Chen, Zhenxing Xu, Yinglong Li, and Gencai Chen. 2016. The discovery of personally semantic places based on trajectory data mining. *Neurocomputing* 173 (Jan. 2016), 1142–1153. <https://doi.org/10.1016/j.neucom.2015.08.071>
- [41] Raul Montoliu. 2012. Discovering Mobility Patterns on Bicycle-Based Public Transportation System by Using Probabilistic Topic Models. In *Ambient Intelligence - Software and Applications*, Paulo Novais, Kasper Hallenborg, Dante I. Tapia, and Juan M. Corchado Rodriguez (Eds.). Springer Berlin Heidelberg, 145–153.
- [42] Raul Montoliu, Jan Blom, and Daniel Gatica-Perez. 2013. Discovering places of interest in everyday life from smartphone data. *Multimedia Tools and Applications* 62, 1 (01 Jan 2013), 179–207. <https://doi.org/10.1007/s11042-011-0982-z>
- [43] U. Naeem, A.-R. Tawil, and I.I. Kennedy. 2015. A dynamic segmentation based activity discovery through topic modelling. In *IET International Conference on Technologies for Active and Assisted Living (TechAAL)*. Institution of Engineering and Technology, London, UK, 6 –6 . <https://doi.org/10.1049/ic.2015.0136>
- [44] Thuong Nguyen, Sunil Gupta, Svetha Venkatesh, and Dinh Phung. 2014. A Bayesian Nonparametric Framework for Activity Recognition Using Accelerometer Data. In *2014 22nd International Conference on Pattern Recognition*. IEEE, Stockholm, Sweden, 2017–2022. <https://doi.org/10.1109/ICPR.2014.352>
- [45] Abhinav Parate, Meng-Chieh Chiu, Chaniel Chadowitz, Deepak Ganesan, and Evangelos Kalogerakis. 2014. Risk: Recognizing smoking gestures with inertial sensors on a wristband. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. ACM, 149–161.
- [46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

- [47] Thanh-Trung Phan and Daniel Gatica-Perez. 2017. Healthy #fondue #dinner: analysis and inference of food and drink consumption patterns on instagram. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia - MUM '17*. ACM Press, Stuttgart, Germany, 327–338. <https://doi.org/10.1145/3152832.3152857>
- [48] Thanh-Trung Phan, Florian Labhart, Skanda Muralidhar, and Daniel Gatica-Perez. 2020. Understanding Heavy Drinking at Night through Smartphone Sensing and Active Human Engagement. In *Proceedings of Int. Conf. on Pervasive Computing Technologies for Healthcare (Pervasive Health)*.
- [49] Tian Qin, Wufan Shangguan, Guojie Song, and Jie Tang. 2018. Spatio-Temporal Routine Mining on Mobile Phone Data. *ACM Transactions on Knowledge Discovery from Data* 12, 5 (June 2018), 1–24. <https://doi.org/10.1145/3201577>
- [50] Tian Qin, Guojie Song, and Sizhen Du. 2016. Routine Mining Based Anomaly Detection in Mobile Phone Data. In *Proceedings of the 2016 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 180–188.
- [51] Weijun Qiu and Ayomi Bandara. 2015. GPS Trace Mining for Discovering Behaviour Patterns. In *2015 International Conference on Intelligent Environments*. IEEE, Prague, Czech Republic, 65–72. <https://doi.org/10.1109/IE.2015.17>
- [52] Radim Rehrák and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. ELRA, Valletta, Malta, 45–50. <http://is.muni.cz/publication/884893/en>.
- [53] Michael Röder, Andreas Both, and Alexander Hinneburg. 2015. Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining*. ACM, 399–408.
- [54] Darshan Santani, Joan-Isaac Biel, Florian Labhart, Jasmine Truong, Sara Landolt, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2016. The night is young: urban crowdsourcing of nightlife patterns. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*. ACM Press, Heidelberg, Germany, 427–438. <https://doi.org/10.1145/2971648.2971713>
- [55] Darshan Santani, Florian Labhart, Sara Landolt, Emmanuel Kuntsche, Daniel Gatica-Perez, et al. 2018. DrinkSense: Characterizing youth drinking behavior using smartphones. *IEEE Transactions on Mobile Computing* 17, 10 (2018), 2279–2292.
- [56] Julia Seiter, Oliver Amft, Mirco Rossi, and Gerhard Troster. 2014. Discovery of activity composites using topic models: An analysis of unsupervised methods. *Pervasive and Mobile Computing* 15 (Dec. 2014), 215–227. <https://doi.org/10.1016/j.pmcj.2014.05.007>
- [57] Julia Seiter, Wei-Chen Chiu, Mario Fritz, Oliver Amft, and Gerhard Troster. 2015. Joint segmentation and activity discovery using semantic and temporal priors. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, St. Louis, MO, USA, 71–78. <https://doi.org/10.1109/PERCOM.2015.7146511>
- [58] Oliver Stanesby, Florian Labhart, Paul Dietze, Cassandra Wright, and Emmanuel Kuntsche. 2019. The contexts of heavy drinking: A systematic review of the combinations of context-related factors associated with heavy drinking occasions. *PLOS ONE* 14, 7 (2019), 29. <https://doi.org/10.1371/journal.pone.0218465>
- [59] Brian Suffoletto, Pedram Gharani, Tammy Chung, and Hassan Karimi. 2018. Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment. *Gait & posture* 60 (2018), 116–121.
- [60] Feng-Tso Sun, Yi-Ting Yeh, Heng-Tze Cheng, Cynthia Kuo, and Martin Griss. 2014. Nonparametric discovery of human routines from sensor data. In *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, Budapest, Hungary, 11–19. <https://doi.org/10.1109/PerCom.2014.6813938>
- [61] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. 2006. Hierarchical Dirichlet Processes. *J. Amer. Statist. Assoc.* 101, 476 (Dec. 2006), 1566–1581.
- [62] Andrew Viterbi. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE transactions on Information Theory* 13, 2 (1967), 260–269.
- [63] Xingxing Xing, Man Li, Weisong Hu, Wenhao Huang, Guojie Song, and Kunqing Xie. 2014. A Spatial-temporal Topic Segmentation Model for Human Mobile Behavior. In *WAIM 2014*.
- [64] Fengli Xu, Tong Xia, Hancheng Cao, Yong Li, Funing Sun, and Fanchao Meng. 2018. Detecting Popular Temporal Modes in Population-scale Unlabelled Trajectory Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 46 (March 2018), 25 pages. <https://doi.org/10.1145/3191778>
- [65] Zhixian Yan, Dipanjan Chakraborty, Christine Parent, Stefano Spaccapietra, and Karl Aberer. 2013. Semantic trajectories: Mobility data computation and annotation. *ACM Transactions on Intelligent Systems and Technology* 4, 3 (June 2013), 1.
- [66] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wang-Chien Lee, Tz-Chiao Weng, and Vincent S. Tseng. 2010. Mining user similarity from semantic trajectories. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks - LBSN '10*. ACM Press, San Jose, California, 19. <https://doi.org/10.1145/1867699.1867703>
- [67] Yanxia Zhang, Jeffrey Olenick, Chu-Hsiang Chang, Steve W. J. Kozlowski, and Hayley Hung. 2018. The I in Team: Mining Personal Social Interaction Routine with Topic Models from Long-Term Team Data. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval - IUI '18*. ACM Press, Tokyo, Japan, 421–426. <https://doi.org/10.1145/3172944.3172997>
- [68] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In *Proceedings of the 18th International Conference on World Wide Web (Madrid, Spain) (WWW '09)*. ACM, New York, NY, USA, 791–800. <https://doi.org/10.1145/1526709.1526816>
- [69] Yin Zhu, Yuki Arase, Xing Xie, and Qiang Yang. 2011. Bayesian nonparametric modeling of user activities. In *Proceedings of the 2011 international workshop on Trajectory data mining and analysis - TDMA '11*. ACM Press, Beijing, China, 1.