2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Extracting user habits from Google maps history logs

Christos Sardianos
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: sardianos@hua.gr

Iraklis Varlamis
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: varlamis@hua.gr

Grigoris Bouras
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: it21239@hua.gr

***Abstract*—The exponential growth in the usage of smart devices, such as smartphones, interconnected wearables etc., creates a huge amount of information to manage and many research and business opportunities. Such smart devices become a useful tool for user movement recognition, since they are equipped with different types of sensors and processors that can process sensor data and extract useful knowledge. Taking advantage of the GPS sensor, they can collect the timestamped geographical coordinates of the user, which can then be used to extract the geographical location and movement of the user. Our work, takes this analysis one step ahead and attempts to identify the user's behavior and habits, based on the analysis of user's location data. This type of information can be valuable for many other domains such as Recommender Systems, targeted/personalized advertising etc. In this paper, we present a methodology for analyzing user location information in order to identify user habits. To achieve this, we analyze user's GPS logs provided through his Google location history, we find locations that user usually spends more time, and after identifying the user's frequently preferred transportation types and trajectories, we find what type of places the user visits in a regular base (such as cinemas, restaurants, gyms, bars etc) and extract the habits that the user is most likely to have.**

***Keywords*-habit recognition, trajectory patterns, significant places.**

I. Introduction

Nowadays, mobile devices and smartphones have earned an essential role in our society, penetrating many facets of everyday life. This is obvious if we consider that 80% of the Internet users own a smartphone [1], while mobile application usage is growing by 6% every year. Modern smart devices are equipped with multiple ambient sensors that provide great amounts of data, which can be analyzed to discover useful information such as daily user patterns, trajectory patterns etc. According to an MIT Technology Review[1], the collection and analysis of information from simple cellphones can provide surprising insights into how people move about and behave. Moreover, if we consider that users are sharing location information with other users, providing real-time updates to other users and benefiting from other users' location sharing, we can assume that this type of user network can be consider a social network with location-based information. This assumption led to the rise of Location-Based Social networks [2].

The information about the places that a user has visited can be exploited in many ways, for example, for promoting POIs that match user's profile [3], for recommending alternatives to his/her fellow users and friends in a location-based social network [4], for extracting useful statistics about the popularity of POIs, etc. A recommender system can take advantage of user location history only, can add content information from social networks, from explicit and implicit user preferences, from third-part services that provide information about POIs (e.g. Yelp, Foursquare, Open Street Maps etc.), or even can take advantage of the temporal information behind each check-in [5].

In this work, we consider two different types of information that can be extracted from user GPS data: i) location information, for places that the user has visited, along with the timestamp and duration of the visit, ii) trajectory information, concerning the movement of a user from one location to another.

We process user GPS logs and extract this information at the first step and then enrich and abstract this information in order to extract interesting user behavioral patterns, comprising:

- frequently visited POIs and POI types,
- frequent trajectories and preferred transition type,
- temporal patterns that associate user preference with the day of the week or the timezone of each activity.

In section II that follows, we summarize the most important works that relate to the extraction of user habits from user location data and take advantage of this information for generating personalized recommendations. In section III we provide an overview of the proposed method and in section IV we provide the implementation details of each processing step. Finally, section V provides a demonstration of the proposed method on data obtained from Google Maps history of a user and section VI summarizes our progress so far and the next steps of this work that are expected to lead to a recommender system that delivers the right recommendation at the right moment.

[1] http://www.technologyreview.com/featuredstory/513721/big-data-from-cheap-phones

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain 978-1-5386-6051-5/18/$31.00
978-1-5386-6051-5/18/$31.00 © 2018 IEEE

II. Related Work

The concept of mining useful knowledge from GPS logs has been discussed several times in the related literature. The survey work of Zheng [6] in trajectory data mining, summarizes all paradigms of trajectory mining and the issues that must be considered. More specifically, when it comes to user trajectories, early works [7] analyze GPS logs in order to mine interesting locations and travel sequences, employ user location history to measure user similarity [8], or identify and assign significance to semantic locations based on GPS records [9], whereas more recent works [10] take advantage of user location history and a richer content (i.e. user sentiment, user interest and location properties) in order to better match users to locations and recommend POIs of interest to users. Collaborative applications have boosted the interest of GPS log mining and introduced location and activity recommender systems and location-based social networks. The survey work of Bao et al [11] provides a taxonomy of recommender systems that build around social and location based information for supporting Location-Based social networks. Such applications, are based on user location and trajectory data [12]–[14], the overall user behavior and social circle in order to recommend POIs [15] or trajectories [16] that connect places of interest.

Despite the long interest on user location and trajectory history, there have only been a few works that take advantage of user's latent behavior patterns in order to provide personalized recommendations. It is worth to mention the work of He et al [15], who attempt to jointly model next POI recommendation under the influence of user's latent behavior pattern. Authors adopt a third-rank tensor to model the successive check-in behaviors of a user in POIs and fuse it to the personalized Markov chain of observed successive user check-ins, in order to improve POI recommendations. They generalize user check-in history at day of week level in the time dimension and at POI category in the POI dimension. However, in their model they limit user profiling only on the POIs and not on the trajectories between POIs.

In a slightly different context, authors in [5] introduce the concept of temporal matching between user profile and POI popularity. They profile the temporal pattern of area activity around POIs, using information from taxi pick-ups and drop-offs and propose that every user has a latent daily-repeated personalized temporal regularity, which decides when he/she is likely to explore POIs every day.

The method that we propose in this paper, is based on the analysis of user GPS data in order to extract useful information concerning the user's behavior and habits. It combines methods and techniques from the related literature and proposes an implementation for extracting spatio-temporal user patterns, which can then be used as a basis for temporal and interest-aware activity recommender systems. However, it differs from existing systems in that it extracts user habits in the form of temporal patterns that repeatedly occur in user logs and refer to the same locations or location types and the same movement type across repeated trajectories.

III. Problem definition - Method Overview

The proposed method for extracting user habits from user GPS logs, follows a step-wise approach, which begins with solving simple problems, such as the detection of user stay points and user trajectories between consecutive stay points, continues with the semantic annotation and enrichment of the extracted stay points and trajectories and ends with the abstraction of these movements to user movement patters that are repeated periodically with increased frequency. In the following subsections, we explain the steps of the proposed method.

A. Extraction of stay points and user trajectories

A *stay point* SP stands for a geographic region, where a user stays over a certain time interval [7]. The extraction of a stay point depends on two parameters, a time interval threshold (T_{threh}) and a distance threshold (D_{threh}). For example, for the set of GPS points depicted in Figure 1, that correspond to the consecutive positions of a moving user, the stay point SP can be regarded as the virtual location containing a subset of consecutive GPS points $SP = \{p_m, p_{m+1}, ..., p_n\}$, where $\forall i, j \in [m, n]$ it holds that $Distance(p_i, p_j) \leq D_{threh}$ and $|p_i.T - p_j.T| \geq T_{threh}$, where $p_i.T$ is the timestamp associated with point p_i. The centroid of all GPS points that grouped under the same stay point location, is used as the GPS coordinates of stay point SP and a maximum radius for the cluster is also kept with the stay point. Finally, since the stay point represents the stay of the user in a location for a time period, we keep the start and end time-stamp with the stay point. So, a stay point is characterized as: SP=$< lat, lon, radius, t_{start}, t_{end} >$.

Figure 1. Stay point example.

Consequently, a *user trajectory* Tr_{ij} is defined as the sequence of GPS points between two stay points SP_i and SP_j. Tr_{ij} defines a route comprising a series of GPS locations in chronological order. Stay points denote the end of one trajectory and the beginning of another. The timestamp difference between two consecutive points (GPS locations' timestamps) in a trajectory, is exceeded, so that they cannot be considered to belong in the same stay point. Figure 2 presents an example of a user trajectory connecting two consecutive stay points. So a user trajectory is characterized by a set of GPS coordinates a start and an end time-stamp as: Tr=$< \{(lat_i, lon_i)\}, tr_{start}, tr_{end} >$.

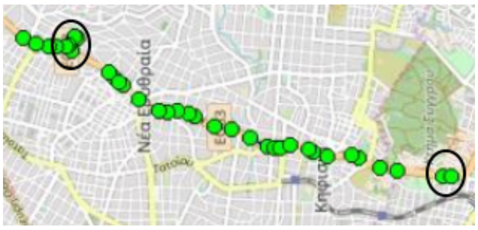

Figure 2. Trajectory example.

B. Semantic enhancement of stay points and user trajectories

When a stay point or a trajectory is extracted, the proposed method applies a semantic enrichment process, which attaches additional information that can be useful for extracting user habits at a later stage. The information for characterizing stay points can be provided from different sources. However, the main sources are POI information services, which provide semantic annotations for popular points of interest as well as for common types of POIs (e.g. leisure places, sport facilities, public buildings, transportation hubs etc). Although POI information services cover a large amount of locations, there are still stay points that can not be characterized without the user's intervention. Following the practice of popular on-the-go driving directions applications (e.g. Google Maps), users can add more personal semantic information for their own points of interest such as "Home" or "Work". These explicitly provided annotations usually refer to frequently accessed stay points that cannot be easily mapped to a known POI. As described in Section III, a stay point SP is defined as any geographic region where the user stayed for at least a short time of period (e.g. a store or an office) or even a location where the user slightly moved around, but not away from it (e.g. a park or a stadium etc.). So after extracting the possible stay points of the user, we access POI information services and search for nearby POIs, at a short range and always within the stay point limits. Using the same services, we are able to characterize the stay point by the type, such as: Park, Stadium, Leisure ground, Athletic center, Beach, Shopping center, Cafeteria etc. At the end of this step a stay point is characterized as:
$SP =< lat, lon, POI_{type}, POI_{category}, t_{start}, t_{end} >$

The semantic annotation of user trajectories, mainly refers to the application of data mining techniques to the trajectory points information (latitude, longitude and timestamp) and the detection of user type of movement across the trajectory. This allows to detect at a later stage, the preferred way of movement for specific trajectories or overall, whether the user uses public or private means of transportation etc. We treat the problem of detection of user movement type across a trajectory as classification problem [17], [18], and build on our previous work on the topic [19], [20]. The movement categories can be: motionless, walking, running, riding a bike, driving a car, being on the bus, being on a train/metro, and any other type of movement. At the end of this step a trajectory is characterized as:
$Tr =< \{(lat_i, lon_i)\}, mov_{type}, POI_{start}, POI_{end}, t_{start}, t_{end} >$

C. Abstracting user data to user habits

By the term user habits we describe a routine of behavior that the user repeats regularly and which tends to occur subconsciously [21], [22]. To adopt this to the scenario of user trajectory data, habits are repetitive user activities such as: being in the same stay point at the same time or day of a week, taking the same trajectory at the same time or week day.

The analysis of user trajectories at user group level, will allow to detect how often places are visited and understand which locations are the most popular on week days or weekends, in the morning or afternoon and which locations are visited for a few moments or for longer periods [23]. When examined at user level, the trajectories can define a set of user behaviors, highlight user habits and allow recommender algorithms or similar applications to provide user-tailored *location and time based* recommendations. Finally, in a collaborative environment (e.g. in location based social

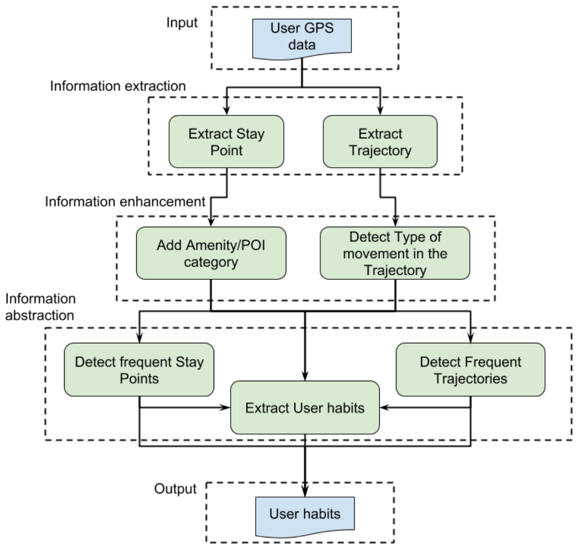

Figure 3. Basic architecture of the application.

networks) users can be compared based on their detected habits and collaborative filtering can be fused with user location and time context (and habits) information.

The extraction of user habits begins with the detection of frequent occurring patterns in user GPS logs (e.g. frequently accessed POIs and trajectories) and continues with the abstraction of information at various levels of granularity in time (e.g. time zones, days, etc), POI type, and movement type. The result of this process comprises frequently occurring staying or movement patterns for one or more users. Defining our approach for analysis, we can describe it as a three-tier analysis with multiple tasks in each tier. In the sections that follow we will describe our methodology in more detail. Figure 3 describes the basic architecture of our application.

IV. System implementation

A. Data collection

The first step of the proposed method is the collection of user GPS data. Although in previous works we employed actual GPS data collected by the user smart-phone, in this work we use data from Google Maps history, which are imported as KML (Keyhole Markup Language) files to the system. KML is a file format based on the XML standard and uses a tag-based structure with nested elements and attributes to express geographic data (such as locations, image overlays, video links and modeling information like lines, shapes, 3D images and points) in location browsers such as Google Earth and Google Maps.

The same processing pipeline can be applied to the actual GPS data collected by the smart-phone instead of using the KML file. The information extracted from the analysis of a user activity in a certain time-frame can be stored in the phone and all the actual GPS data for this frame can be erased. For example, when the user commutes to work, we can store information about the trajectory (e.g. start/end time and location, and probably a few intermediate points) and erase all the intermediate GPS data.

B. Information extraction

1) Extraction of user stay points: The initial task on the analysis of user location data is to identify the locations, where the user stays for a certain amount of time. For this purpose, following the visit point extraction method described in [24], we employ DBSCAN [25], a density-based clustering algorithm, which finds clusters of dense points using a range threshold eps and a minimum number of points $MinPts$ within this range as parameters. We implement a spatio-temporal version of DBSCAN (the distance of two points is a linear combination of geographic distance and time distance), which clusters together neighboring (in space and time) GPS traces and ignores all other points (considers them noise). Depending on the frequency of recorded GPS spots, the distance threshold of interest and a moving speed threshold, we can compute an acceptable value for $MinPts$. The parameterization of the algorithm has been described in [20].

A spatial clustering of user location data, using a density-based clustering algorithm, is expected to identify areas with very dense recorded spots and areas where the user passed through at a quick pace. This clustering will provide information for places that user spends time during the day, even when the user is standing or walking around (in a park) or running (in a stadium). The use of time distance in the distance measure of DBSCAN will change the resulting clusters, and will allow to distinguish between a stay point and a point that the user crosses several times but in different (distant) timestamps. A time-ignorant DBSCAN will detect a single cluster for all points, whereas a time-aware version will detect separate clusters. An incremental version of DBSCAN, allows to cluster the most recent GPS traces of a user and detect the stay points as they occur, and consequently assign all intermediate points to the trajectory. The steps of the stay point extraction process are depicted in Figure 4.

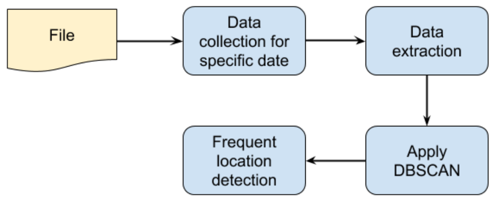

Figure 4. Workflow for identifying frequent user locations.

Figure 5 depicts the results of the stay point extraction process on a map. On the left part all the GPS points, before the detection of stay points are shown on the map in red, whereas on the right part the detected stay points only are marked with green color.

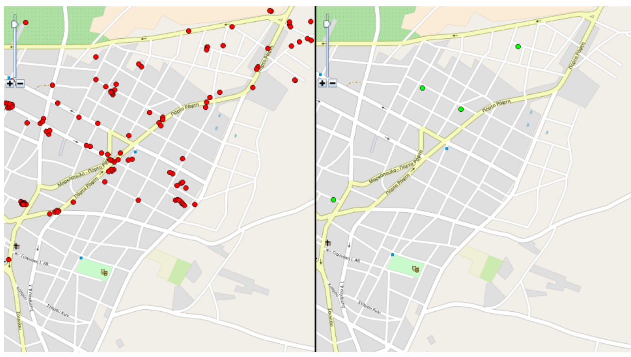

Figure 5. All GPS points of user (on the left) and the results of detected stay points (on the right).

2) Detection of user trajectories: The detection of user trajectories is binded to the detection of stay points and trajectories are directly defined as the sets of GPS tracks between two consecutive stay points. Considering the fact that a large set of stay points may exist in user's GPS logs, a respectively large set of trajectories are formed among the different stay points. So the result of this information extraction step comprises two sets: a set of user's stay points and a set of trajectories that join consecutive stay points.

C. Information enhancement

1) Semantic characterization of stay points: After user's stay points and trajectories are detected, they are annotated with additional information, which is collected by third party services or extracted by data mining algorithms, and which can be used in the user habits extraction step. The characterization of stay points step employs the OpenStreetMaps service, which offers an API for retrieving information about various POIs in a geographical area. For this purpose, a bounding box is created for each stay point, using the GPS coordinates of the stay point as the bounding box center and a range that does not exceed the radius of the respective cluster. The OpenStreetMaps API is accessed to retrieve POIs within the geographical area defined by the bounding box and it responds with an XML formated result, as depicted in Figure 6.


```xml
<?xml version="1.0" encoding="UTF-8"?>
<osm version="0.6" generator="CGImap 0.4.0 (22136 thorn-03.openstreetmap.org)" copyright="OpenStreetMap
contributors" attribution="http://www.openstreetmap.org/copyright" license="http://opendatacommons.
org/licenses/odbl/1-0/">
 <bounds minlat="38.0340500" minlon="23.7369800" maxlat="38.0346400" maxlon="23.7375200"/>
 <node id="519954080" visible="true" version="3" changeset="10846515" timestamp="2012-03-02T10:12:51Z"
 user="armitatz" uid="414661" lat="38.0357216" lon="23.7355228">
  <tag k="highway" v="traffic_signals"/>
 </node>
 <node id="2750852492" visible="true" version="1" changeset="21374126" timestamp="2014-03-28T22:34:57Z"
 user="Chris Makridis" uid="1227858" lat="38.0341596" lon="23.7374317">
  <tag k="addr:housenumber" v="4"/>
  <tag k="addr:postcode" v="14341"/>
  <tag k="addr:street" v="Βρυούλων"/>
  <tag k="amenity" v="restaurant"/>
  <tag k="name" v="Εν Αιθρία"/>
  <tag k="website" v="www.enaithria.com"/>
 </node>
```


Figure 6. Sample of the XML response with tags for stay points.

The file contains all possible Points-of-Interest marked with tags that characterize the type of POIs inside the bounding box of the stay point, such as:

- amenity
- public transport
- shop
- sport
- leisure

From the locations returned by the POI service, the closest to the stay point is used to characterize the stay point. The result of this process is that the user stayed at a specific POI (which is of certain type and category) for a specific time period.

2) Detecting type of movement in a trajectory: For the semantic annotation of a user trajectory, we process all consecutive GPS traces in order to detect user movement speed, user direction and user speed changes, we also check if the traces are near a public transportation (PT) stop or on a known PT route. Building on our previous work on the topic [19], [20] we classify each trace individually and then classify the trajectory as a whole. If there exists a set of pre-classified movement samples we train a personalized model for each user, else a pre-trained generic model is used, which uses a set of direct (latitude, longitude, timestamp) and indirect features (speed, speed changes, distance from transportation related POIs such as bus stops or metro stations) in order to characterize how the user moved across a trajectory [20]. Since our original classifier is incremental and annotates the last trajectory part with the detected movement type it is frequent that for a long trajectory, more than one movement types have occurred (e.g. the user drives but stops at the traffic lights or is stuck behind a bus for part of the trajectory). When the next stay point is detected and the trajectory is completed, a post-processing step aggregates this information for all segments of the trajectory and assigns the movement type that most likely matches to the specific trajectory.

To provide an example, let's assume the daily commute of a user to work as depicted in Figure 7. The user drives from home to the nearest train station, parks the car and takes the train from station A to the nearest station (station B) at work, then walks to commute to work. The stay point detection algorithm will detect four stay points (home, parking lot of train station A, train station B, work) and three trajectories that connect them. The first trajectory will be annotated as driving, the second as moving by metro/train and the last as walking. The sub-trajectory from the train station's parking lot to the station building will not be detected if the two places are close to each other. The walk of the user from the parking lot to the train station will also be considered as part of user's visit to the specific stay point (i.e. train station A).

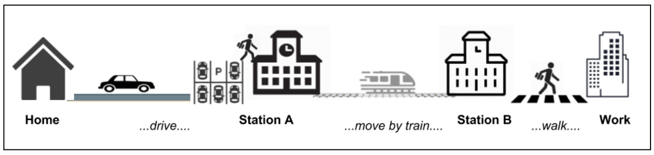

Figure 7. An example of a user commuting to work. The detection and annotation of stay points and trajectories.

D. Information abstraction

1) Detection of frequent user stay points: The next phase of the process is to extract the frequent stay points and trajectories of a user based on his/her location history. By analyzing the stay points of a user for a long time period, we can find if the user tends to visit specific stay points more frequently. To achieve this, we cluster stay points using a distance-based only DBSCAN and result with clusters containing points that have been visited many times by the user. We rank clusters in descending order of size and keep the top ranked stay points for a user. Along with the GPS coordinates of each stay point, we have time-stamp information concerning the start and end time of the user's stay. A first step of information abstraction is to find the preferred days or time zones for a user to visit a stay point. The result of this step can be similar to the following: *The user has visited train station A n times this month. The preferred days are week days, and the preferred time zones are early in the morning and the afternoon.*

2) Detection of frequent trajectories: Similar to the analysis of stay points, the analysis of user trajectories will highlight the preferred movement paths of the user and the preferred way of movement. We focus only on the frequent stay points for a user and considering the set of trajectories that the user has followed to go from one stay point to another. We apply the clustering-based sequential mining (CBM) algorithm [26] over the set of trajectories and the output of this process is the set of most frequent trajectories followed by the user. In detail, the CBM algorithm is based on the clustering of the set of points that belong to the trajectory, so in our case the input of the algorithm is the set of user trajectories that connect user frequently accessed stay points and two parameters, namely s and ξ. Parameter s defines the square area occupied by each cluster on the map and parameter ξ defines the minimum number of points that a cluster has to contain in order to be considered as active (that means that is frequently part of the user's trajectories).

The frequent trajectories are characterized by the start and end location and the set of start (departure) and end (arrival) time-stamps and type of movement for each trajectory instance, as follows:

$Freq_{Tr} =< POI_{start}, POI_{end}, \{(mov_{type}, t_{start}, t_{end})_i\} >$

The processing pipeline is summarized in Figure 8.

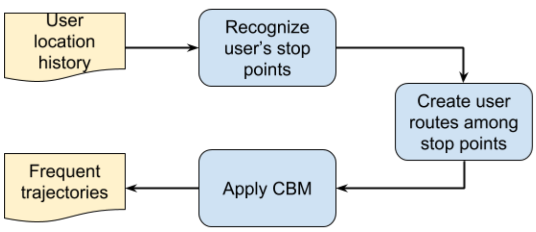

Figure 8. Frequent trajectories extraction pipeline.

E. Extracting user habits

The last processing step aims to use the information extracted from the previous steps and find frequently occurring activity patterns in user logs, which will define the user's habits. The analysis of information concerning frequently visited locations, which the user visits periodically, and frequent trajectories that the user follows to reach these destinations, will discover user's tendencies on commuting (e.g. home $\rightarrow$ work $\rightarrow$ gym $\rightarrow$ restaurant $\rightarrow$ home) at specific dates and times.

The process of user habits extraction, can be treated as an association rules extraction problem. Before extracting frequent patters and interesting rules, it is important to process the annotated stay point and trajectory information and get different levels of abstraction. For example, different places that have been visited by the user and have been annotated as restaurants, bars or cafeterias can be generalized to the category amenity and lead to rules with stronger support, the time-stamp information can be mapped to day zones (e.g. morning, afternoon, evening) or days, using different levels of granularity. The output of this information abstraction step is fed to the association rule extraction algorithm, which in our case is the Apriori algorithm [27].

V. Real case demonstration

In order to demonstrate our proposed method for extracting user habits from user trajectory data, we developed an application that takes Google Maps History files (KML files) as input and processes them following the process described in the previous sections. The application is written in Java and is available as a standalone Java program (Figure 9).

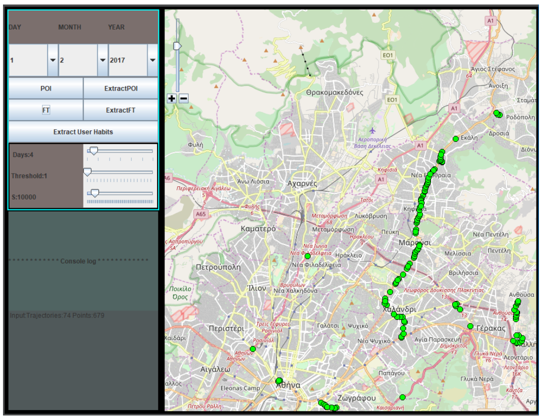

Figure 9. Snapshot of the application.

Through the application the user has a set of options that can trigger the analysis of the location data contained in the KML file, which may span several days or months. The user can either analyze the loaded data and depict them over an OpenStreetMap map embedded in the user interface, or extract the results of the information extraction and enhancement step to output files for further analysis. Using these files as input, we can extract the user's frequent stay points and trajectories, which can be displayed over an OpenStreetMap layer or exported to separate files.

The analysis of user stay points leads to a set of user's tendencies like the ones displayed in Figure 10.

These files are then fed to the Apriori algorithm to extract user habits. The input data that we use consist of the

```
Latitude      Longitude     Time
37.8883883    23.9400781    Sun Jan 01 18:11:51 EET 2017 - Sun Jan 01 22:58:04 EET 2017

Location frequence:45
Frequent Day: Sunday

Latitude      Longitude     Time
37.9033496    23.7499417    Sun Jan 08 20:00:25 EET 2017 - Sun Jan 08 22:51:04 EET 2017

Location frequence:40
Frequent Day: Sunday

Latitude      Longitude     Time
37.8608787    23.7534746    Sun Jan 01 23:35:14 EET 2017 - Mon Jan 02 00:35:33 EET 2017

Location frequence:37
Frequent Day: Monday

Place name:Barón
Amenity:cafe  ID:3942492230
```

Figure 10. An example user habit based on frequent stay points: User visits the "Baron cafe" on Sunday and Monday evenings.

user's extracted stay points combined with the time-stamp of occurrence after converting the actual start and end time-stamps into daytime zone (e.g. morning, afternoon, night) and day (e.g. weekday, weekend), the type of movement at that moment and/or the type of the amenity. The set of categorical features are fed into the Apriori association rules algorithm. So, when the Apriori algorithm is fed with information in the form of: $\{DayZone, DayType, MoveType, Dest\_POI_{Type}\}$, where $DayType$ can be weekday or weekend and the $Dest\_POI_{Category}$ can be the type of POI detected as frequent stay point (e.g. cafeteria), the extracted user habits are similar to those depicted in Figure 11.

```
1. MoveType = Metro ⇒ IsWorkingDay = true
2. DayZone = Evening  Tag = public_transport 331 ⇒ IsWorkingDay = true
3. DayZone = Evening  MoveType = Metro  Tag = public_transport ⇒ IsWorkingDay = true
4. DayZone = Evening  MoveType = Metro ⇒ IsWorkingDay = true
5. Tag = shop ⇒ IsWorkingDay = true
                    ...
                    ...
```

Figure 11. Sample of the extracted habits after the Apriori execution.

Based on the sample result of Figure 11 we can assume for example, based on rules No. 2, 3 and 4 that the user tends to commute by metro in the workday evenings, while based on rule No. 5 user also tends to visit a shop on working days.

VI. Conclusions and Next Steps

Considering the implicit interaction among users who are sharing location information with other users, we can assume that these users form a type of social network with location-based information. In this work, we presented an application that processes user GPS logs to extract useful information that is enriched and abstracted, in order to extract rules and patterns that describe user's habits. These behavioral patterns include frequently visited POIs, frequently used trajectories and associations among user preferences with day of week and/or timezone of activity. As far as the results show, we can use the GPS logs to identify interesting patterns for the user daily activity. This type of information could be exploited furthermore in many types of applications and fields of research, such as recommender systems for providing personalized recommendations.

Having that said, we consider this is a field of interest with lots of potential and it is in our intention to adapt our work so far in order to lead us to a recommender system that would deliver real-time and real-life recommendations based on user habits.

Moreover, the parameter selection of the algorithms has been made after experimentation on the specific dataset. A more thorough evaluation of different parameter settings is part of our next work on the field.

Acknowledgements

This work has been developed in the frame of the MASTER project, which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777695.

This paper was made possible by National Priorities Research Program (NPRP) grant No. 10-0130-170288 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] D. Bosomworth, "Mobile marketing statistics 2015," *Leeds: Smart Insights (Marketing Intelligence) Ltd*, 2015.

[2] P. Symeonidis, D. Ntempos, and Y. Manolopoulos, "Location-based social networks," in *Recommender systems for location-based social networks*. Springer, 2014, pp. 35–48.

[3] W.-Y. Zhu, W.-C. Peng, L.-J. Chen, K. Zheng, and X. Zhou, "Modeling user mobility for location promotion in location-based social networks," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015, pp. 1573–1582.

[4] Y. Yu and X. Chen, "A survey of point-of-interest recommendation in location-based social networks," in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, vol. 130, 2015.

[5] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2013, pp. 363–372.

[6] Y. Zheng, "Trajectory data mining: an overview," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 6, no. 3, p. 29, 2015.

[7] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from gps trajectories," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 791–800.

[8] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, "Mining user similarity based on location history," in *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*. ACM, 2008, p. 34.

[9] X. Cao, G. Cong, and C. S. Jensen, "Mining significant semantic locations from gps data," *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 1009–1020, 2010.

[10] H. Gao, J. Tang, X. Hu, and H. Liu, "Content-aware point of interest recommendation on location-based social networks." in *AAAI*, 2015, pp. 1721–1727.

[11] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," *Geoinformatica*, vol. 19, no. 3, pp. 525–565, 2015.

[12] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1029–1038.

[13] Y. Zheng, X. Xie, and W.-Y. Ma, "Geolife: A collaborative social networking service among user, location and trajectory." *IEEE Data Eng. Bull.*, vol. 33, no. 2, pp. 32–39, 2010.

[14] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 1082–1090.

[15] J. He, X. Li, L. Liao, D. Song, and W. K. Cheung, "Inferring a personalized next point-of-interest recommendation model with latent behavior patterns," in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[16] D. Quercia, R. Schifanella, and L. M. Aiello, "The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city," in *Proceedings of the 25th ACM conference on Hypertext and social media*. ACM, 2014, pp. 116–125.

[17] F. Sparacino, "The museum wearable: Real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences." 2002.

[18] N. Bu, M. Okamoto, and T. Tsuji, "A hybrid motion classification approach for emg-based human–robot interfaces using bayesian and neural networks," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 502–511, 2009.

[19] I. Varlamis, "Evolutionary data sampling for user movement classification," in *Evolutionary Computation (CEC), 2015 IEEE Congress on*. IEEE, 2015, pp. 730–737.

[20] S. Tragopoulou, I. Varlamis, and M. Eirinaki, "Classification of movement data concerning user's activity recognition via mobile phones," in *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*. ACM, 2014, p. 42.

[21] M. Y. Mind, "the mental fitness guide by gillian butler and tony hope," 1995.

[22] N. Eagle and A. S. Pentland, "Reality mining: sensing complex social systems," *Personal and ubiquitous computing*, vol. 10, no. 4, pp. 255–268, 2006.

[23] D. Gubiani and M. Pavan, "From trajectory modeling to social habits and behaviors analysis," in *Recent Trends in Social Systems: Quantitative Theories and Quantitative Models*. Springer, 2017, pp. 371–385.

[24] M. Lv, L. Chen, Z. Xu, Y. Li, and G. Chen, "The discovery of personally semantic places based on trajectory data mining," *Neurocomputing*, vol. 173, pp. 1142–1153, 2016.

[25] S. Kisilevich, F. Mansmann, and D. Keim, "P-dbscan: a density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos," in *Proceedings of the 1st international conference and exhibition on computing for geospatial research & application*. ACM, 2010, p. 38.

[26] A. A. Shaw and N. Gopalan, "Finding frequent trajectories by clustering and sequential pattern mining," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 1, no. 6, pp. 393–403, 2014.

[27] R. Agarwal, R. Srikant *et al.*, "Fast algorithms for mining association rules," in *Proc. of the 20th VLDB Conference*, 1994, pp. 487–499.