Retrieving Points of Interest from Human Systematic Movements

Riccardo Guidotti[1,2(✉)], Anna Monreale[1,2], Salvatore Rinzivillo[2], Dino Pedreschi[1], and Fosca Giannotti[2]

[1] KDDLab, University of Pisa, Largo B. Pontecorvo, 3, Pisa, Italy
{riccardo.guidotti,anna.monreale,dino.pedreschi}@di.unipi.it
[2] KDDLab, ISTI-CNR, Via G. Moruzzi, 1, Pisa, Italy
{salvatore.rinzivillo,fosca.giannotti}@isti.cnr.it

Abstract. Human mobility analysis is emerging as a more and more fundamental task to deeply understand human behavior. In the last decade these kind of studies have become feasible thanks to the massive increase in availability of mobility data. A crucial point, for many mobility applications and analysis, is to extract interesting locations for people. In this paper, we propose a novel methodology to retrieve efficiently significant places of interest from movement data. Using car drivers' systematic movements we mine everyday interesting locations, that is, places around which people life gravitates. The outcomes show the empirical evidence that these places capture nearly the whole mobility even though generated only from systematic movements abstractions.

1 Introduction

The study of human mobility can offer insight into human behavior [5,14]. Traces of human mobility can be collected with a great number of different techniques such as GPS (Global Positioning System) or GSM (Global System for Mobile Communications). The result is a huge quantity of data: about tens of thousand people moving along millions of trajectories. Mobility data can provide a complete description of the places visited and the routes followed by individual users. There are many potential opportunities, and movement data have been recognized by private and public institutions as a valuable source of information to evaluate the habits of people in terms of mobility.

Recent researches in mobility analysis have been extended in order to identify the behaviors that people constantly follow, such as groups of trajectories with common routes [13] or popular destinations [2]. Indeed, a central point in these studies is the concept of *place of interest* [7] in urban mobility environment, i.e. certain places or areas attract individual movements due to their importance. It is worth to point out that people move from one place to another, therefore "places" are not only static geographical objects, but they are also part of people life. The way people move towards these places affects the overall mobility of the environment. Thus, in order to study the relationships between people movements and the places of interest, it is mandatory to have a method that takes into account people's mobility to extract the locations they frequent routinely.

C. Canal and A. Idani (Eds.): SEFM 2014 Workshops, LNCS 8938, pp. 294–308, 2015.
DOI: 10.1007/978-3-319-15201-1_19

Online static datasets of places of interest can be easily exploited to analyze data, and there are plenty of works that enhance their potential. However, capturing real-time human mobility is challenging and often requires expensive frameworks and infrastructures. At any rate, places of interest directly extracted from movement data are more reliable and trustful than those readable from the Web or from public sources. This happens because the last ones are static, rarely updated and, overall, usually related to commercial activities such as bars, hotels, museums and so on.

The method proposed in this paper allows us to extract the places of interest around which our life gravitates. These places are extracted considering how people's everyday systematic mobility is regulated and influenced by them. Using mobility data as a proxy of human mobility and the idea of mobility profiles [13], we introduce a new notion of *Points of Interest* (POIs) explaining how they can be extracted. We test our method on a real case study considering big datasets of *GPS* trajectories. The outcomes show the empirical evidence that these POIs represent nearly the whole mobility even though they are generated only from a systematic movement abstraction. Finally, we propose a wide range of applications for which POIs extracted in such a way can be extremely useful.

The remainder of this paper is organized as follows. Section 2 presents a set of papers extracting places of interest from mobility data. In Sect. 3 are reported some basic concepts to understand the methodology presented in the following. Section 4 illustrates the procedure for the POIs extraction, while in Sect. 5 are reported the experimental results obtained using real datasets. In Sect. 6 are illustrated some possible applications for the proposed methodology. Finally, Sect. 7 concludes the paper.

2 Related Work

In the following are reported some recent works in which the extraction of places (or regions) of interest is a fundamental point. Each one of them explains its own extraction method starting from different types of data. In [2] the authors propose a visual analytic procedure for studying mobility data. Their procedure extracts relevant places from movement data because, for their aim, there is not a predefined set of places (e.g. compartments of a territory division) from which the analyst can select places of interest. In [6,10], the authors generate regions of interest with the purpose of predicting human movements using mobility pattern mining. The regions of interest are obtained by discretizing the working space in a regular grid with cells of small size. Then, the cells not visited are discarded and, by following a density based principle, the cells conceptually belonging to the same points are merged. Similarly, since it is impossible to translate a continuous surface into a graph, different authors in [4] and [12] discretize the territory in cells to apply social network analysis techniques on mobility data. In [8] is proposed an approach based on supervised learning to infer people's motion models from their GPS logs. The authors, first analyze different features to understand the kind of movement performed by the users (car, bus, bike etc.), and then use clustering algorithms to detect stopping points areas. To estimate the physical location of users from traces of mobile devices associated with access points in a wireless network, the authors in [9] characterize popular regions evaluating access points paths with GPS traces. Finally, in [15] it is proposed an approach that is capable of uncovering semantically relevant keywords for describing a location. Also in this case the locations correspond to the access points areas.

3 Preliminaries

Movements are usually performed by people in specific areas and time instants. These people are called *users* or *drivers* and each movement is composed by a sequence of spatio-temporal points (x, y, t) where x and y are the coordinates, while t is the time stamp. We call *trajectory* the movements of a user described by a sequence of spatio-temporal points:

Definition 1 (Trajectory). *A* trajectory *m is a sequence of spatio-temporal points $m = \{(x_1, y_1, t_1), \ldots, (x_n, y_n, t_n)\}$ where the spatial points (x_i, y_i) are sorted by increasing time t_i, i.e., $\forall 1 \leq i \leq k$ we have $t_i < t_{i+1}$.*

The set of all the trajectories traveled by a user u makes her *individual history*:

Definition 2 (Individual History). *Given a user u, we define the* individual history *of u as the set of traveled trajectories denoted by $H_u = \{m_1, \ldots, m_k\}$.*

Using the above definitions and following the profiling procedure proposed in [13], we can retrieve the systematic movements of a user u. Thus, we group the trajectories using a density-based clustering (i.e., Optics [3]) equipped with a *distance function* defining the concept of trajectory similarity:

Definition 3 (Trajectory Similarity). *Given two trajectories m and p, a trajectory distance function* dist *and a distance threshold ε, we say that m is* similar *to p ($m \sim p$) iff* $\text{dist}(m, p) \leq \varepsilon$*.*

The obtained result is a partitioning of the original dataset from which we filter out the *clusters* with few trajectories and the one containing noise. Finally we extract a *representative trajectory* from each remained cluster. These representative trajectories are called *routines* and the set of routines is called *mobility profile*. More formally:

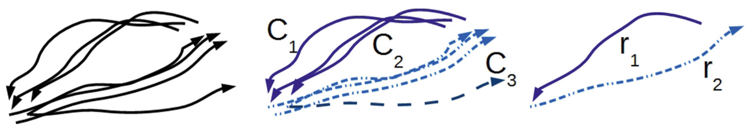


Fig. 1. The user *individual history* (black lines), the clusters identified by the grouping function (C_1, C_2, C_3) and the extracted *individual routines* (r_1, r_2) forming her *individual mobility profile*.

Definition 4 (Routine and Mobility Profile). *Let H_u the individual history of a user u, ms a minimum size threshold,* dist *a distance function and ε a distance threshold. Given a mobility profiling function* $\text{profile}(H_u, ms, \text{dist}, \varepsilon) = \mathcal{M}$*, such that $\mathcal{M} = \{M_1 \dots M_k\}$ where $M_i \subset H_u$, we define a* routine r_i *as the medoid trajectory of a group M_i. The set of routines extracted from $\mathcal{M}$ is called* mobility profile *and is denoted by $P_u = \{r_1 \dots r_k\}$.*

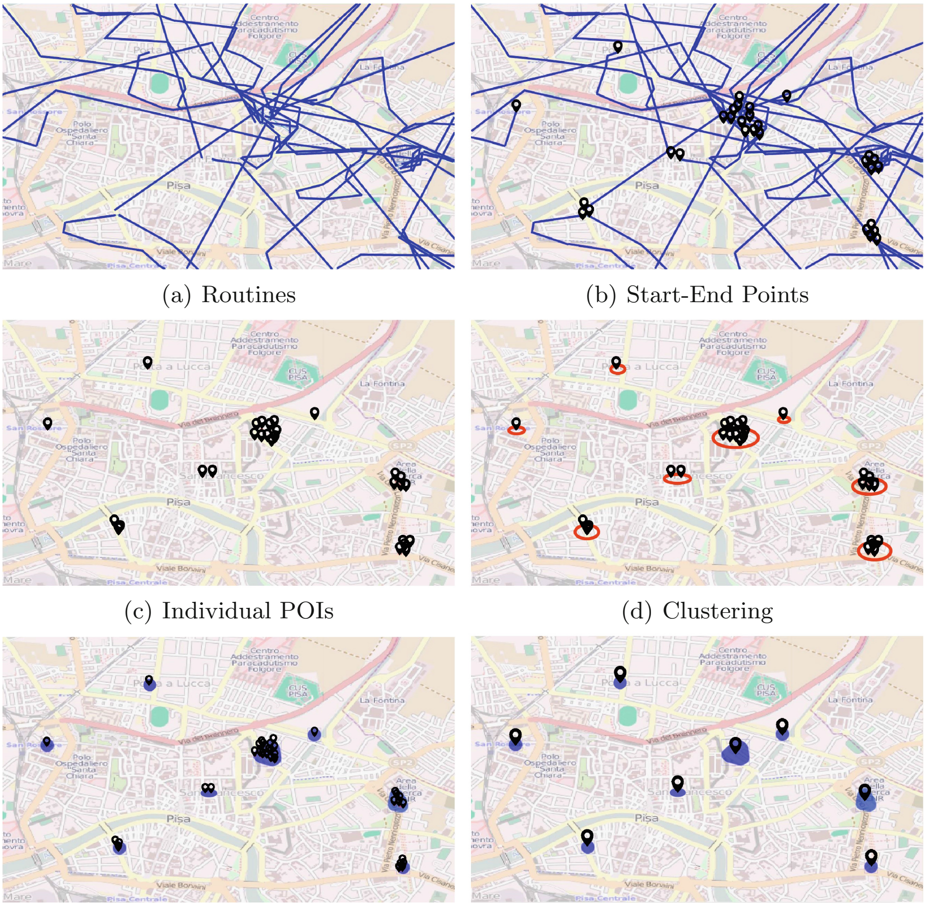
A *mobility profile* describes an abstraction in space and time of the systematic movements: the user's real movements are represented by a set of trajectories delineating the generic paths followed. Moreover, the exceptional movements are completely ignored due to the fact they will be not part of the profile. Figure 1 depicts an example of mobility profile extraction.

4 Mobility Points of Interest Extraction

In mobility data studies, places or regions can be extracted from raw data through regular territory division. However, relevant places for human mobility do not have regular shapes. Indeed, they may have arbitrary shapes and sizes and irregular spatial distribution. They might even overlap in space; hence, approaches based on dividing the territory into non-overlapping areas (as in [1] and [10]) are not appropriate.

What we are looking for in our study are places of interest that approximate as better as possible human mobility and consequently human behavior. Commonly, *a Point of Interest (POI) is a specific point location that "someone" may find useful or interesting.* Most consumers use the term POIs when referring to hotels, campsites, fuel stations or any other category used in modern navigation systems. In fact, the term is widely used in cartography, especially in electronic variants including GIS, and GPS navigation software. A GPS point of interest specifies, at minimum, the latitude and longitude of the POI. Digital maps for modern GPS devices (e.g. TomTom and Garmin) or GPS navigator applications (e.g. Google Maps and Waze), typically include a basic selection of POIs for the map area. Moreover, there are websites specialized in the collection, verification, management and distribution of POIs which end-users can load onto their devices to replace or supplement the existing POIs. While some of these websites are generic, and collect and categorize POIs for any interest, others are more specialized in a particular category (e.g. as speed cameras).

All the aforementioned type of POIs are strictly related with commercial activities (e.g. bars, restaurants, hotels and shopping centers), public facilitates (e.g. hospitals, schools and universities), leisure sites (e.g. museums and amusement parks). These places are useful to organize a holiday trip or to find a place to spend the evening. At any rate, they do not consider everyday human mobility. Indeed, people constantly follow the same periodic movement during their working day with systematic patterns. Thus, people's visited locations are influenced by the systematic movements of their everyday life. Looking individually at each user, everyone has as most visited POIs her own home, her working place, her



(a) Routines (b) Start-End Points

(c) Individual POIs (d) Clustering

(e) Buffering (f) Collective POIs

Fig. 2. Mobility POIs extraction method.

habitual shopping centers, and maybe her gym and her friends' homes. These POIs are the real interesting locations in individuals routinely life. Thus, from hereafter we will refer to a POI with this latter concept: *a POI is a place that "someone" may find relevant or interesting in her everyday systematic life.*

We propose a new method to extract these POIs in order to understand which are our significant locations and to study how certain places are affected by human systematic mobility. In the following, we illustrate the mobility data-driven procedure to extract POIs from trajectory data. Figure 2 depicts in detail all the steps to retrieve POIs. The systematic behavior of each user can be modeled with the mobility profiles presented in the previous section. Thus, the systematic daily mobility of each user is characterized by her routines (Fig. 2-a). These routines necessarily begin and end somewhere. For profiled users, having a mobility that gravitates around these locations, it results that these places are surely very important for them (Fig. 2-b). We identify these places as *individual POIs* (Fig. 2-c):

Definition 5 (Individual Point of Interest). *Given the mobility profile P_u of a certain user u, then the* individual POIs *of u is the set I_u such that*

$$I_u = \{p | p = start(r) \ \vee \ p = end(r) \ \forall r \in P_u\}$$

where $start(.)$ and $end(.)$ are two functions that given a routine return its start point and its end point, respectively.

We remark that, in this paper, a POI has the meaning of "a place frequently visited by someone" and not the meaning of a public attraction. Therefore, our extraction method allows us to infer not only typical attraction points (because surely there is at least someone working there), but also important places for individual users, such as their home, which are not available in typical public sources. We are able to capture this information thanks to the fact that the GPS signal tells us the position of the nearest parking from the location visited by the user. As we will observe in the following, typically each user frequently visits two places that are with high probability home and work.

From Definition 5 we can notice that, given two different drivers u and v, which systematically park their cars close each other, we have that $I_u \cap I_v = \emptyset$, since each individual POI is represented by GPS coordinates and it is nearly impossible that there is a perfect correspondence. However, these users are following a similar systematic mobility behavior towards the same location, as a consequence the two individual POIs should be geographically considered as a unique *collective* POI. To this aim, given a set of car drivers and considering a certain spatial tolerance, we compute a density-based clustering on the individual POIs and then, we turn each valid cluster and each noise point into a buffered convex shape area representing a collective POI.

Definition 6 (Collective Point of Interest). *Given a set I of individual POIs, then the* collective POIs *set C_I is defined as*

$$C_I = buffer(convex(clustering(I, \varepsilon)), \varepsilon')$$

In the above definition, ε and ε' are distance values and $clustering(I, \varepsilon)$ is a density-based clustering function that returns clusters composed of individual POIs (Fig. 2-d). Note that two points are considered close enough if their distance is lower than ε. The clusters returned can also be composed of noise points because each noise point represents an individual POI supported by at least a routine and thus, it is relevant for at least one user. $convex(.)$ is a function returning the convex shape of the clusters of the input points. If the cluster contains only one point, then itself is returned. Lastly, $buffer(.)$ is a function that applies a spatial buffer of ε' to the set of input shapes and points (Fig. 2-e). In the following, for the sake of simplicity, we will call a collective POI simply POI. In other words, we can think to a POI as a geographical area with a

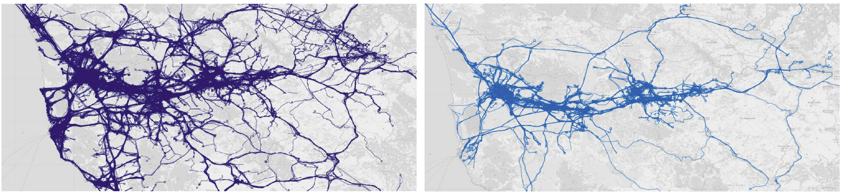


Fig. 3. (Left) A sample of the considered trajectories in Pisa province. (Right) Mobility profiles extracted in Pisa province.

certain extension that is visited frequently by at least one user (Fig. 2-f). Note that two different POIs a and b could be overlapped because of the buffering phase. Anyway, $\varepsilon' < \varepsilon$ ensures that the center of a is not included in b because otherwise the clustering algorithm would have put them in the same cluster because they would have been distant no more than ε.

5 Mobility Case Study

To extract the latent POIs in human systematic mobility we applied the method described above on large provincial trajectory datasets. First of all, we briefly report some consideration about the dataset used and the mobility profile extraction. Then, we describe the study performed to extract reliable POIs and what they represent on the analyzed area. Finally, we show why the extracted POIs represent the overall mobility even though they are built starting from the systematic movements abstractions that are mobility profiles.

5.1 Mobility Dataset

As a proxy of human mobility, we use real GPS traces collected for insurance purposes by *Octo Telematics S.p.A*[1]. This dataset contains 9.8 million car travels performed by about 160,000 vehicles active in a geographical area focused on Tuscany in a period from 1st May to 31st May 2011. Figure 3-left depicts a sample of the considered trajectories. The mobility dataset is geographically and temporally too various to be used for our purposes. Thus, it was split following different principles based on time and geography. In real world, different events may change how people move on the territory. Such events can be unpredictable or not frequent, like natural disaster, but most of them are not. The most regular and predictable event is the transition between working days and non-working days. During Saturday and Sunday, people usually leave their working mobility routines for different paths. Following this concept we filtered out weekend trajectories, maintaining only weekday ones. Another basic issue is that the mobility is not the same in every geographical area. Every area has its own type of mobility with certain characteristics depending on the surface, the topology and the

[1] http://www.octotelematics.com/it.

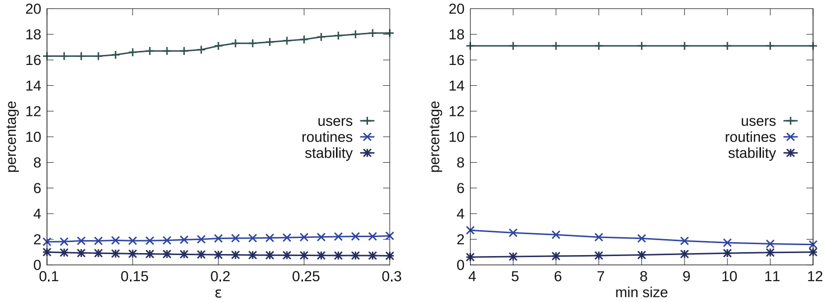


Fig. 4. Profile test ε (left) and min size (right).

number of inhabitants. In order to consider this fact, it was made a geographical filter to split the dataset in provinces by considering for each province all the trajectories that pass through its administrative borders. In this paper we present the results obtained for Pisa, Florence, Siena and Grosseto province.

In order to obtain sound routines we perform some test to set the best parameter to extract reliable mobility profile. Figure 3-right depicts an example of profile extracted in Pisa province modeling the users' systematic movements. The distance function used in the clustering step is *Route Relative Synch* described in [13]. The clustering algorithm used is Optics [3], a density-based algorithm. We study Optics parameters on a subset of $1,000$ users in Pisa province. Thus, we vary ε in the range $[0.1, 0.3]$ with step 0.01, Fig. 4-left. The bigger ε is, the more different trajectories are allowed to be clustered together. The threshold *min size*, the minimum number of trajectories that must be in a cluster considered valid, is varied in the range $[4, 12]$, Fig. 4-right. The aspects we consider to tune the values are: *(a)* the dataset coverage, *(b)* the profile distribution per user, and *(c)* the profile stability. From this empirical study we decide to use middle values because the plots obtained do not lead to a clear setting. Anyway, in each plot, after the middle values the curves change more rapidly than before them. We choose ε equal to 0.2, it expresses 80 % of similarity between two trajectories and, a reliable value for *min size* is 8 since a routine is a movement repeated a sufficient number of time during a month.

Figure 5-left shows the number of routines per users in Pisa province where each user has almost one or two routines, which, should correspond to the commute to and from work. Indeed, note that the average number of routines per profile is 2, this is probably due to the home-work-home pattern. In Fig. 5-right the temporal distribution of the trajectories and routines is shown. Here, we can see how the profile set has a working-like trend, highlighting the three peeks during the early morning 5–6, lunchtime 11–12, and late afternoon 17–18. This confirms the previous assumption: mobility profiles are reliable to model systematic movement and thus can be exploited to retrieve systematically visited places.

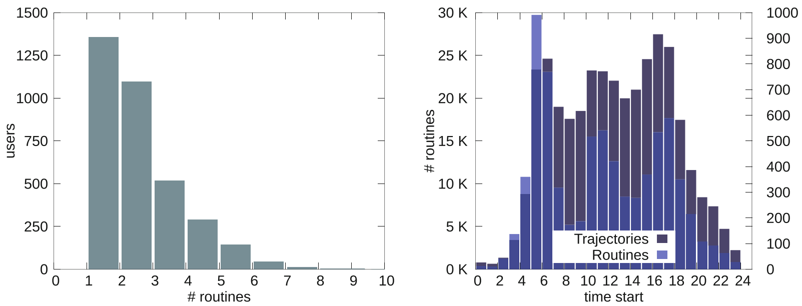


Fig. 5. Routine per user (left), trajectories and routines time start (right) distributions.

5.2 Mobility POIs Extraction Analysis

In the following we analyze the mobility POIs extraction method. Two main issues are considered to build reliable POIs: *(a)* a significant number of POIs must be visited by at least two users otherwise they will be useful only as an individual information in a urban collective scenario, *(b)* POIs shape cannot degenerate, i.e. they cannot be too big, nor too long, nor sausage-shaped. Only two parameters must be considered in POIs extraction process: ε and ε', and, since ε' depends on ε, we study only ε. We tested POIs extraction using the routines of $1,000$ profiled users in Pisa province with $\varepsilon \in [20, 100]$ and step 10. In this case ε in Optics represents the meters of distance between two individual POIs to be considered close. We recall that every POI is important for someone because it is generated by a routine. In order to guarantee both *(a)* and *(b)* we perform an accurate analysis. Thus we study the number of POIs extracted and the average number of users in a POI depicted in Fig. 6-left. We notice that the number of POIs extracted rapidly decreases while the number of POIs with more than one user grows slowly. On the contrary, the average number of user in a POI increases linearly. Moreover, we examined the maximum area and diameter for the POIs extracted, reported in Fig. 6-right. From these lines trend we observe that the maximum values, accordingly to the median and average ones (here not shown), rapidly rise leading to some degenerate POI that collects conceptually different places. Thus, by looking together at these plots, a reasonable value suggested for ε appears to be 50 m. Consequently, we set $\varepsilon' = 45$ to have a remarkable buffer even for individual POIs. In fact, this combination of parameters leads to a good number of POIs neither too big nor too small visited on average by at least two users.

For each province, we obtain a POIs distribution per profiled users telling us that the bigger subset of profiled users stop from 1 to 5 POIs. As it is shown in Table 1, the average number of profiled users per POI in every province ranges from 2 to 4 meaning that, on the whole, a collective points is nearly always visited by at least two users. This happens because, many places (probably home) are visited only by one user, while other social POIs like hospitals and shopping centers are visited by many users. For the home-work-home pattern, the majority of the users visit at least two places. Moreover, still from Table 1,

Table 1. Tuscany mobility POIs statistics. The public source for surface, inhabitants and density is http://en.wikipedia.org/wiki/Tuscany.

Province	Pisa	Florence	Siena	Grosseto
POIs	9,760	12,848	7,299	6,567
Users	20,898	41,724	27,242	14,036
Users profiled	21.05 %	11.82 %	15.13 %	33.24 %
Avg users per POI	2.14	3.25	3.73	2.14
Routines	7,383	9,801	6,458	7,281
Surface (km^2)	2,448	3,514	3,821	4,504
Inhabitants	409,251	983,073	268,706	225,142
Density (inh./km^2)	167.2	279.8	81.9	50.0

we note that for every province the number of POIs extracted is not correlated neither with the number of routines, nor with the number of users profiled, nor with the surface. On the contrary, it seems to better correlate with the number of inhabitants.

As final analysis it interesting to observe which are the most visited POIs in every province. Thus we counted how many trajectories present in the initial provincial datasets start and end in every POIs. It emerges that for each provincial dataset there are few POIs visited by many people and many POIs visited by few car drivers, following a long tailed power low distribution. In Fig. 7, depicting semi-log normalized number of visits distributions, we can notice how, despite the difference in number of POIs extracted, all the distributions are quite close. This indicates that the whole mobility is similar with respect to our POIs for the provinces analyzed. This obviously happens because there are some POIs with a role of prevalence, that is more visited, with respect to the others. They are very fascinating places because these POIs are visited both by *systematic drivers* working there and by *occasional drivers*. As an example we can think to the following. Doctors and salesman, working in hospitals and shopping centers

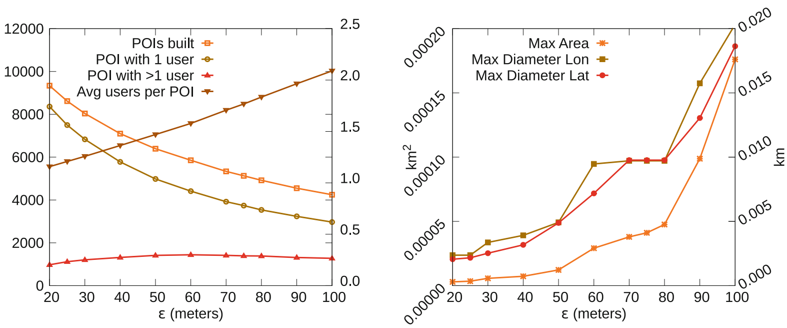


Fig. 6. POI construction test parameter ε (left) POIs numbers, (right) POIs shape.

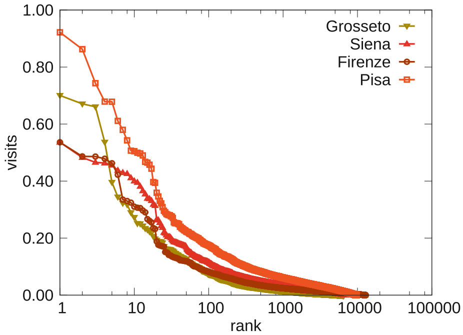


Fig. 7. POIs number of visits distribution for each provincial dataset.

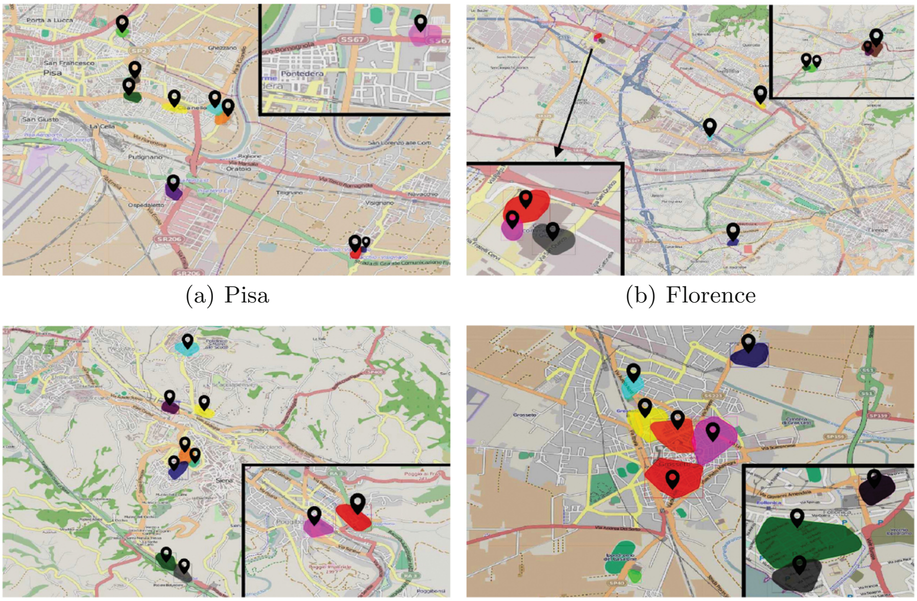
respectively, stop there systematically, while patients and customers just visit these places when they need. The former category, *systematic drivers*, surely belong some routines that start and end there, that is the reason why they have been extracted as POIs. On the other hand, the latter category, *occasional drivers*, belong just several trajectories starting or ending there. However, due to the fact that places like hospitals and shopping centers are attractors for many people, there are many trajectories starting and ending in these places, augmenting in this way the visitors count. Conversely, the great majority of less visited places are POIs for at least a driver by definition, and thus they correspond to homes or to not very frequented working places. Figure 9 shows the ten most visited POIs in Pisa, Florence, Siena and Grosseto. As suggested above, they are mainly big shopping centers, hospitals and car parks close to locations visited very often by many people. We can notice how for every province there are some of these popular POIs out of the main town corresponding to car parks close to big malls. In Fig. 8 is depicted a zoom on the four most visited POIs in Pisa province. As one can see, the POIs areas bound perfectly the car parks close to the real point of interest. This demonstrate the good quality and precision of the mobility POIs extraction method proposed.

5.3 Mobility POIs as Mobility Summary

An important result emerges as a side effect from the POIs extraction process: *the mobility POIs, and thus the mobility profiles, are a good representation of*

(a) 1^{st} - IperCoop Mall (b) 2^{nd} - Pisanova Mall (c) 3^{rd} - Cisanello Hospital (d) 4^{th} - Carrefour Mall

Fig. 8. The fourth most visited POIs in Pisa.



(a) Pisa (b) Florence (c) Siena (d) Grosseto

Fig. 9. Ten most visited POIs in Pisa, Florence, Siena and Grosseto.

the overall mobility. Taking into account that this process starts from the routines and not from all the trajectories, it is interesting to notice that, for every provincial dataset, about 80 % of the trajectories start or end into the POIs extracted. Detailed statistics about coverage are reported in Table 2. Figure 10 shows all the trajectories starting or ending in a little sample of POIs in Pisa. As you can see, the map is almost completely covered by the red lines representing the trajectories. This is a signal that these POIs have an high importance in the overall mobility because they can capture nearly all the route traveled in the considered geographical area. This fact visually reinforces the hypothesis that mobility profiles, that is systematic trajectories, are a good representation of all the mobility. Consequently, mobility POIs are good to capture human mobility. As this assumption appears true, then it is a great simplification to use routines instead of all the trajectories to analyze human mobility. Another confirmation of the strength brought by mobility profiles emerges from the visual inspection of the starting and ending points of the trajectories not starting nor ending in any POIs. It comes out that these places are not really interesting because they do not correspond to important locations but they are almost all private houses which are also occasionally visited by their owners.

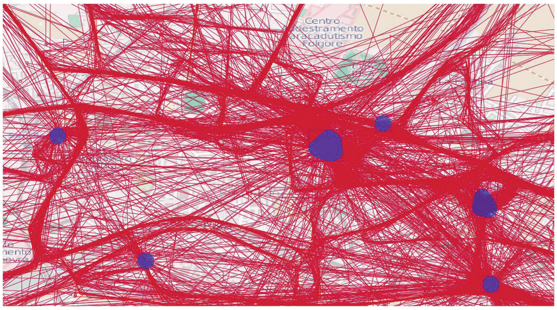


Fig. 10. Trajectories stopping in a POIs sample.

Table 2. Tuscany mobility POIs coverage.

Province	Pisa	Florence	Siena	Grosseto
POIs	9,760	12,848	7,299	6,567
Trajectories	476,267	1,358,596	478,424	661,116
POIs coverage	80.32 %	76.44 %	79.80 %	84.54 %

6 Mobility POIs Applications

There is wide set of mobility applications which need a data driven POIs extraction method to perform different tasks and solve distinct problems. In the following we illustrate a broad range of mobility task where reliable and functional POIs are mandatory in order to obtain good results.

A deep study in mobility data relates to mobility flows and patterns. A mobility pattern represents the regularity of a set of users moving from a place to another, that is, from a POI to another. Thus, in order to deal with worth patterns, interesting POIs must be considered. As in [10], these patterns can be used to solve *mobility prediction* tasks. A mobility prediction is a statement about the place someone will be in the future, often but not always based on experience or knowledge. Prediction tasks use previously extracted movement patterns, which are a concise representation of behaviors of moving users as sequences of places frequently visited with a typical travel time. A decision tree, is built and evaluated with a formal training and test process. Consequently, to produce proper predictions, it is essential to use a sound set of POIs as those extracted by our method.

Another typical application to use our POIs is a *recommendation system*. A recommendation system is a subclass of information filtering system that seeks to predict the “rating” or “preference” that user would give to an item. Recommendation systems have become extremely common in recent years, and are applied in a variety of applications such as mobility. As an example, a recommendation system can exploit the correlation between geographical locations in the space of human behavior, that is, POIs correlation, to suggest new POIs to visit. In [16], for example, by taking into account users travel experience and the subsequent locations visited, the authors learn the location correlation from a large number of user-generated GPS trajectories. Then, by using the POIs correlation, they conduct a personalized location recommendation system, which is evaluated on the basis of a real-world GPS dataset.

Finally, some possible applications are related to *complex network analysis studies* in which the POIs are the nodes of the networks. An example can be found in [11] where the authors analyze the urban mobility and the POIs trying to featuring the places in a city according to how people move among them. Then, they build a POIs network by connecting POIs where trajectories pass. From such a network they extract the communities finding group of places highly connected by people mobility. As another example, a possible mobility data driven analysis could consist in building the bipartite graph of drivers and POIs to investigate the relationship between how the movements of people are affected by the POIs, and how the places themselves are characterized and connected to the mobility of people.

7 Conclusion

One of the most fascinating challenges of our time is to study the global interconnected society, especially, to understand the human mobility. The analysis of movement data and locations of interest has been recently promoted by the wide diffusion of new techniques and systems for monitoring, collecting and storing positional data. In this paper we have shown a novel approach to extract people real POIs from mobility GPS data. We have seen that the procedure is efficient because it does not need all the trajectories present in the data but just a representative abstraction. Moreover, we have observed that the places extracted with the proposed method capture both famous collective POIs and individually important POIs. Finally, as a positive side effect of this study, we have shown that the mobility POIs extracted do not lose in generality even though generated only from systematic movements. A possible future work related to the POIs extraction method consists in adjusting the radius used by the clustering algorithm with respect to the population density of the area in which the POIs are retrieved. Another possible improvements consists in extending the geographical and systematical information given by our POIs with the static and semantic knowledge contained in classical points of interest. That is, we could extend the informative power of our POIs by matching them with common points of interest information saying for example that they are bars, museum, hospital and so on.

Acknowledgements. This work has been partially supported by the European Commission under the FET-Open Project n. FP7-ICT-284715, ICON, and by the European Commission under the SMARTCITIES Project n. FP7-ICT-609042, PETRA.

References

1. Adrienko, N., Adrienko, G.: Spatial generalization and aggregation of massive movement data. IEEE Trans. Vis. Comput. Graph. **17**(2), 205–219 (2011)
2. Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., Wrobel, S.: From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In: 2011 IEEE Conference on VAST. IEEE (2011)
3. Ankerst, M., Breunig, M.M., Kriegel, H.-P., Sander, J.: Optics: Ordering points to identify the clustering structure. In: ACM SIGMOD Record, vol. 28. ACM (1999)
4. Coscia, M., Rinzivillo, S., Giannotti, F., Pedreschi, D.: Optimal spatial resolution for the analysis of human mobility. In: 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE (2012)
5. Giannotti, F., Nanni, M., Pedreschi, D., Pinelli, F., Renso, C., Rinzivillo, S., Trasarti, R.: Unveiling the complexity of human mobility by querying and mining massive trajectory data. VLDB J. Int. J. Very Large Data Bases **20**(5), 695–719 (2011)
6. Giannotti, F., Nanni, M., Pinelli, F., Pedreschi, D.: Trajectory pattern mining. In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2007)
7. Hillier, B., Penn, A., Hanson, J., Grajewski, T., Xu, J.: Natural movement-or, configuration and attraction in urban pedestrian movement. Environ. Plann. B **20**(1), 29–66 (1993)
8. Kim, M., Kotz, D., Kim, S.: Extracting a mobility model from real user traces. In: INFOCOM, vol. 6 (2006)
9. Kostakos, V., Juntunen, T., Goncalves, J., Hosio, S., Ojala, T.: Where am i? location archetype keyword extraction from urban mobility patterns. PloS one **8**(5), e6398 (2013)
10. Monreale, A., Pinelli, F., Trasarti, R., Giannotti, F.: Wherenext: a location predictor on trajectory pattern mining. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2009)
11. Ramalho Brilhante, I., Berlingerio, M., Trasarti, R., Renso, C., de Macedo, J.A.F., Casanova, M.A.: Cometogether: discovering communities of places in mobility data. In: 2012 IEEE 13th International Conference on Mobile Data Management (MDM). IEEE (2012)
12. Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., Strogatz, S.H.: Redrawing the map of great britain from a network of human interactions. PloS One **5**(12), e14248 (2010)
13. Trasarti, R., Pinelli, F., Nanni, M., Giannotti, F.: Mining mobility user profiles for car pooling. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2011)
14. Wang, D., Pedreschi, D., Song, C., Giannotti, F., Barabasi, A.-L.: Human mobility, social ties, and link prediction. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2011)
15. Zheng, Y., Li, Q., Chen, Y., Xie, X., Ma, W.-Y.: Understanding mobility based on gps data. In: Proceedings of the 10th International Conference on Ubiquitous Computing. ACM (2008)
16. Zheng, Y., Xie, X.: Learning location correlation from gps trajectories. In: 2010 Eleventh International Conference on Mobile Data Management (MDM). IEEE (2010)