Assessing the Attractiveness of Places with Movement Data

Andre S. Furtado[1], Renato Fileto[1], Chiara Renso[2]

[1] PPGCC/INE - CTC, Federal University of Santa Catarina, Florianópolis-SC, Brazil
{asalvaro, fileto}@inf.ufsc.br
[2] KDD LAB, ISTI-CNR, Pisa, Italy
chiara.renso@isti.cnr.it

Abstract. Attractiveness of places has been studied by several sciences, giving rise to distinct ways for assessing it. However, the attractiveness evaluation methods currently available lack versatility to analyze diverse attractiveness phenomena in different kinds of places and spatial scales. This article describes a novel method, called M-Attract, to assess the attractiveness of interesting places, based on movement data. M-Attract examines trajectory episodes (e.g., stop, pass) that happen in places and their encompassing regions to compute their attractiveness. It is more flexible than state-of-the-art methods, with respect to land parcels, parameters, and measures used for attractiveness assessment. The proposed method has been extensively evaluated in experiments with real data, which demonstrate its contributions to analyze attractiveness of places and identify relevant phenomena in the geographic space.

Categories and Subject Descriptors: H.2.8 [**Database Management**]: Database Applications—*Spatial databases and GIS*; H.2.m [**Database Management**]: Miscellaneous

Keywords: moving objects trajectories, places attractiveness, spatio-temporal information analysis.

1. INTRODUCTION

Attractiveness quantifies how much something is able to attract the attention and influence the decisions of one or more individuals [Uchino et al. 2005]. It can help to explain a variety of spatial-temporal phenomena. Furthermore, methods to properly estimate attractiveness of places are important tools to build applications for several domains, such as traffic, tourism, and retail market analysis.

The attractiveness of geographic places has been studied for decades, by disciplines like geography and economy. Several theories have been proposed to quantify the attractive force and delimit the region of influence of a place, including the Gravitational Attractiveness Model [Reilly 1931] and the Theory of Central Places [Christaller 1933]. Since this pioneering work, a myriad of proposals have been presented to assess the attractiveness of places, in fields like urban planning, transport, marketing, business, migration and tourism. These work use a variety of data to derive attractiveness, including the population in each region, distances between given regions and a target region, surveys based on voting, trajectories of moving objects such as taxis, and time periods when the moves occur, among other. However, these proposals lack versatility with respect to the categories of places they can consider, and the measures used to assess their attractiveness.

Recently, the widespread use of mobile devices (e.g., cell phones, GPS, RFID) enabled collecting large volumes of raw trajectories, i.e., time ordered sequences of spatial-temporal positions of moving objects. It has pushed the demand for mechanisms to extract useful information and knowledge from trajectory data, and can help reduce the burden for collecting travel survey data. Furthermore, trajectory data can provide more detailed spatial-temporal information about the routes, visited

Work supported by CAPES, CNPq (grant 478634/2011-0), EU-IRSES-SEEK project (grant 295179), EU-FET DATASIM (grant 270833) and CNR-CNPQ bilateral project.

places, goals, and behaviors of a variety of moving objects.

Trajectories occur around places in the geographic space. Consequently, several kinds of relations between trajectories and places can be extracted by processing raw trajectories integrated with geographic data. Spaccapietra et al. [2008] defines a semantic trajectory as a set of relevant places visited by the moving object. According to this viewpoint, a trajectory can be regarded as a sequence of relevant episodes that occur in a set of places. Formally, an episode is a maximal segment of a trajectory that comply to a given predicate (e.g., is inside a place, is close to somewhere, is stopped) [Mountain and Raper 2001]. Several techniques have been proposed to extract episodes from raw trajectories. These techniques usually identify the episodes based on the movement pattern (e.g., acceleration change, direction change) or by investigating spatio-temporal and semantic relations between trajectories and places [Parent et al. 2013].

This article is an enhanced version of a previous work [Furtado et al. 2012] that proposes the M-Attract (*Movement-based Attractiveness*) method to assess the attractiveness of places based on raw trajectories. M-Attract is a method that defines different measures of places' attractiveness based on the number of episodes trajectories and certain episodes related to these places and surrounding regions. M-Attract is more flexible than other state-of-art methods with respect to the kinds of individuals, places and geographic scales that can be considered. This article have the following improvements on top of the previous work: (i) the text have been reviewed and expanded; (ii) some definitions and formulas of the proposed method have been improved; (iii) the methods has been validated in more extensive experiments that use a much bigger dataset and more varied configurations of parameters; and (iv) new figures and maps have been produced and incorporated to this article for better explaining the method and show its benefits for places attractiveness assessment and its applications. With the results of these experiments we were able to show that the proposed attractiveness measures allow the identification of different aspects of the attractiveness phenomena, and the analysis of their spatial distribution in maps.

The rest of this article is organized as follows. Section 2 discusses related work. Section 3 provides definitions necessary to understand the proposal. Section 4 presents the proposed method for attractiveness assessment. Section 5 reports experiments and their results. Finally, Section 6 enumerates contributions and directions for future work.

2. RELATED WORK

Traditionally, the attractiveness of places have been calculated from survey data, geographical features, and population distribution. For instance, the attractiveness measure of points of interest (PoIs) proposed by Huang et al. [2010] considers static factors (e.g., the size of commercial places, the distance to their customers' homes) and dynamic factors (e.g., restaurants are more attractive at mealtimes).

The use of trajectories data has just started to be investigated for assessing attractiveness of places [Giannotti et al. 2007; Giannotti et al. 2011; Wei et al. 2010; Yue et al. 2009; Yue et al. 2011]. The seminal work of Giannotti et al. [2007] presents an algorithm for discovering regions of interest based on their popularity, which is defined as the number of distinct moving objects that pass up to a certain distance threshold of these regions (to compensate trajectory points inaccuracies), during a given time period. Several analysis of large volumes of trajectories, based on notions like presence and density of trajectories, are presented by Giannotti et al. [2011]. These work build regions of interest from a grid-based partition of the space into rectangular cells, by aggregating adjacent cells whose measures of trajectories concentration around them are considered similar according to chosen criteria, or high enough to include the cell in a region of interest. They do not calculate attractiveness of predefined regions of interest (e.g., cities, neighborhoods) that can be taken from legacy spatial databases.

The framework for pattern-aware trajectories mining proposed by Wei et al. [2010] uses the density-based algorithm introduced by Giannotti et al. [2007] to extract regions that are passed by at least a certain number of trajectories. They propose an algorithm that exploits the concept of random walk to derive attractiveness scores of these regions. Then, they derive trajectories' attractiveness from these scores. A trajectory is considered more attractive if it visits more regions with high attractiveness.

The work presented by Yue et al. [2009] and Yue et al. [2011] are both based on the analysis of taxi trajectories. Yue et al. [2009] build clusters that groups spatial-temporal similar pick-up and drop-off points of trajectories, and measures the attractiveness of the clusters based on the time-dependent flows between clusters. Yue et al. [2011] assesses the attractiveness of shopping centers, by using data about them (gloss leasable area, number of nearby shopping malls, available parking space, etc.) and trajectory data (number of taxis within their area of influence in different time periods).

The method proposed in this article is more versatile than the previous ones, for the following reasons: (i) it works in several scales using different categories of places, that can be mined by using methods such as those proposed by Giannotti et al. [2007], or taken from legacy databases including popular geographic systems that gather information via crowdsourcing, such as OpenStreetMap[1] and WikiMapia[2]; (ii) it considers real trajectory data from individuals, that can be automatically collected; (iii) it includes a variety of attractiveness measures that can consider episodes in places and/or some of their encompassing regions, calculated with parameters to define time thresholds for considering stops and sizes of buffer around places.

3. PRELIMINARY DEFINITIONS

The goal of M-Attract is to assess how much places of interest are attractive, based on trajectory episodes that occur in their surroundings. This section describes the places, subregions and region of interest and the trajectory episodes considered by the method.

3.1 Regions and Places of Interest

The M-attract method works in a chosen analysis scope, determined by a region, subregions and places of interest[3]. According to the scale of analysis, that can vary in different application domains, the same land parcel can be seen as a region or as a place (e.g., a shopping mall can be seen as a place or a region, depending on the interest in individual stores inside it).

Definition 3.1. A **region of interest** is the totality of the analyzed space. It completely covers all the subregions, places, and trajectories taken into account.

The region of interest (r) determines the spatial scope of the analysis. Depending on the application domain, r can be chosen in a different scale or spatial hierarchy level (if a hierarchy is available). For example, to analyze airspace trajectories r can cover (a considerable portion of) the whole world, and to analyze urban movement r can be just a city or a metropolitan area.

Definition 3.2. **Subregions of interest** are non-overlapping portions of the region of interest that are relevant for the attractiveness analysis.

Many subregions of interest can be considered in an analysis. If the region r of interest is a city, each subregion s may be a a neighborhood, for example.

[1] `http://www.openstreetmap.org`
[2] `http://wikimapia.org`
[3] In this article, we consider that a region, a subregion or a place is represented by a single simple polygon, just for simplicity. However, our proposal can be generalized to work with multi-polygons and polygons with roles.

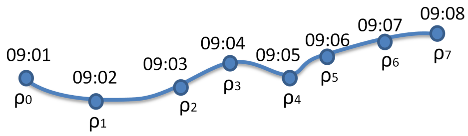

Fig. 1. Example of a raw trajectory.

Definition 3.3. **Places of interest** are non-overlapping portions of the subregions of interest considered in the analysis.

A place of interest (p) is an atomic unity of analysis. Places inside a city zone or neighborhood may be, for example, commercial establishments, public services or tourist places, among others. The classes of places of interest considered in an analysis depend on the application domain.

3.2 Moving Objects' Trajectories

The attractiveness of places can be estimated by the trajectories of moving objects around these places. A raw trajectory can be defined as follows [Alvares et al. 2007].

Definition 3.4. A **Raw trajectory** is a time ordered sequence of spatio-temporal points $t = (\rho_0, \rho_1, ..., \rho_n)(n > 0)$, with each point in the form $\rho_i = (tid, (x, y)_i, time_i)$, where tid is the trajectory identifier, $(x, y)_i$ is a pair of geographic coordinates and $time_i$ is a instant.

The spatio-temporal points of a raw trajectory correspond to sequential observations (samples) of the moving object's position along time. These points can be collected by using technologies such as GPS or GSM. Figure 1 illustrates an example of a raw trajectory.

3.3 Categories of Trajectory Episodes considered in M-Attract

Episodes such as stops [Parent et al. 2013] can be inferred from moving objects dynamic attributes such as speed and acceleration, or the continuous period spent inside or close to places of interest. Some episodes are useful to determine the attractiveness of places. M-Attract estimates the attractiveness of places by counting episodes of the following categories (with $0 \leq i \leq k \leq j < |t|, \xi$).

Definition 3.5 ***stop***. Given a trajectory $t = \{\rho_0, ..., \rho_n\}(n = (|t| - 1) > 0)$, a place p, and the parameters $\xi \geq 0$ and $\delta > 0$, an episode *stop* is a quadruple (eid, tid, pid, t_{ij}), where eid, tid and pid are respectively the episode, the trajectory and the place p identifiers, and t_{ij} is the maximal subsequence of t that complies to Expression 1:

$$\forall \rho_k \in t_{ij} : (x, y)_k \ intersects \ (buffer(geom(pid), \xi)) \wedge (time_j - time_i) \geq \delta \tag{1}$$

Definition 3.6 ***pass***. Given a trajectory $t = \{\rho_0, ..., \rho_n\}(n = (|t| - 1) > 0)$, a place p and $\xi \geq 0$, an episode *pass* is a quadruple (eid, tid, pid, t_{ij}), where eid, tid and pid are respectively the episode, trajectory and place identifiers and t_{ij} is the maximal subsequence of t that complies to Expression 2:

$$\forall \rho_k \in t_{ij} : (x, y)_k \ intersects \ (buffer(geom(pid), \xi)) \tag{2}$$

Figures 2 and 3 illustrate these two categories of episodes. Each episode refers to a trajectory segment (i.e., a subsequence of spatial-temporal observation points) satisfying the respective condition (namely, Definitions 3.5 and 3.6) with respect to a place p. The buffer operator is used in Expressions 1 and 2 to allow a certain degree of uncertainty in face of data accuracy and interpretation issues (e.g., a car with GPS can be parked at a certain distance of place visited by its passengers).

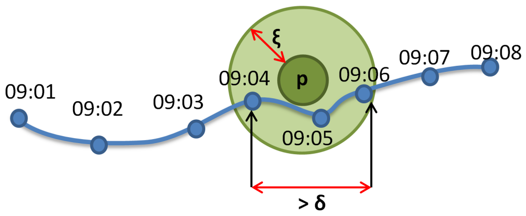

Fig. 2. Episode *stop*

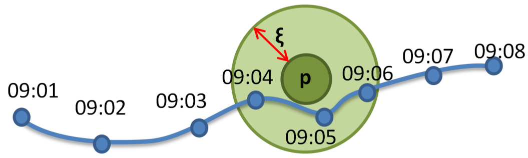

Fig. 3. Episode *pass*

We have chosen these categories of trajectory episodes to develop the M-Attract method because they carry useful information for analyzing the attractiveness of places, though being easy to understand and allowing efficient algorithms to discover such episodes in large collections of raw trajectories and geographic places of interest. Based on the Definitions 3.5 and 3.6, we specified two functions to detect episodes: $stop(t, p, \xi, \delta) : T \times P \times \mathbb{Q}_+{}^2 \to E$ and $pass(t, p, \xi) : T \times P \times \mathbb{Q}_+ \to E$, both functions receive the parameters mentioned in the Definitions 3.5 and 3.6 and return a set of episodes E.

4. THE M-ATTRACT MEASURES OF ATTRACTIVENESS

Let P be a collection of places as described in Definition 3.3, and T be a collection of raw trajectories as described in Definition 3.4. Given a place $p \in P$, the number of episodes, as those described in Definitions 3.5 and 3.6, can give some hint of p's attractiveness. However, for doing deep attractiveness analysis and capturing some subtle attractiveness phenomena, we need to consider not only these basic measures for each place, but also measures for the interesting region where the place is located. This means that we do not want to count only the number of episodes in the places, which is a measure of popularity. We must quantify how much the place attracts the movement of people traveling in the nearby area. This is formalized in the attractiveness measures defined below.

All the proposed measures are based on the number of episodes in places. The parameters buffer size ξ and minimum staying time to characterize a stop ξ may be dependent on the place p being considered. Thus, in the following we denote these parameters as ξ_p and δ_p, respectively. For simplicity and generality, we avoid to mention these parameters in the left-hand side of the following formulas. Furthermore, we sum the numbers of episodes for the places contained each subregion and the whole analyzed region, to make metrics for the respective land parcels (place, subregion and region) that are additive across the space hierarchy considered. This ensures that the proposed measures, stated by Equations 3 to 6, always give real numbers in the interval $[0, 1]$, if the respective denominator is greater than 1. Otherwise the numerator is also 0 and the value of the measure is 0 by convention.

4.1 Stopping Capacity of Places

The following two measures allow the assessment of the stopping capacity of a place p, with respect to trajectories from a set T that pass close to p or stop in any place p' contained in the subregion s that contains p, respectively.

Absolute Stopping Capacity (ASC). Proportion of $pass(t, p, \xi_p)$ that also yield $stop(t, p, \xi_p, \delta_p)$, for a place p, buffer size $\xi_p \geq 0$, minimum staying time $\delta_p > 0$, and trajectory set T (Equation 3).

$$ASC(p, T) = \left(\sum_{t \in T} count(stop(t, p, \xi_p, \delta_p)) \right) \div \left(\sum_{t \in T} count(pass(t, p, \xi_p)) \right) \tag{3}$$

High ASC means that a high percentage of people pass close to place p actually visit p. It can happen for example when tp has a good advertisement to attract people, who are around. Another case of high ASC is when people move to the subregion specially to visit p, which is isolated in the area or other places have low attractiveness.

Relative Stopping Capacity (RSC). Ratio between the number of stops at a given place p and the number of stops in all places p' contained in a given subregion s that contains p, for buffer sizes $\xi_p, \xi_{p'} \geq 0$, minimum staying times $\delta_p, \delta_{p'} > 0$, respectively, and a trajectory set T(Equation 4).

$$RSC(p,s,T,P) = \left(\sum_{t \in T} count(stop(t,p,\xi_p,\delta_p)) \right) \div \left(\sum_{t \in T,\;\; p' \in P}^{s\ contains\ p'} count(stop(t,p',\xi_{p'},\delta_{p'})) \right) \quad (4)$$

RSC measures the attractiveness of a place p compared to other places located in the same subregion s. High RSC for place p means that it is frequently visited, while the other places in s are not.

4.2 Relative Density of Trajectory Episodes in Subregions

The results of some preliminary experiments suggested a need to consider the relative density of passing and stopping episodes in the subregion s containing a place of interest p, with respect to the respective episodes in the whole region r considered for analysis. Thus, we developed the following episodes density measure for subregions of interest.

Relative Passing and Stopping (RPS). Ratio between the number of *pass* in places p within s and in places p' within r, multiplied by the ratio between the number of *stop* in places of p within s and in of in places p' within r, for the trajectories set T (Equation 5).

$$RPS(s,r,T,P) = \left(\frac{\sum_{t \in T,\ p' \in P}^{s\ contains\ p'} count(pass(t,p',\xi_{p''}))}{\sum_{t \in T,\ p'' \in P}^{r\ contains\ p''} count(pass(t,p'',\xi_{p''}))} \right) * \left(\frac{\sum_{t \in T,\ p' \in P}^{s\ contains\ p'} count(stop(t,p',\xi_{p'},\delta_{p'}))}{\sum_{t \in T,\ p'' \in P}^{r\ contains\ p''} count(stop(t,p'',\xi_{p''},\delta_{p''}))} \right) \quad (5)$$

4.3 Attractiveness of Places

Finally, using the measures defined above, we propose the following attractiveness measure for a place of interest p located in subregion of interest s.

Strict Attractiveness (SA). Product of the absolute stopping capacity of a place p, the relative stopping capacity of p with respect to a subregion s containing p, and the relative passing and stopping of s (Equation 6).

$$SA(p,s,r,T,P) = ASC(p,T) * RSC(p,s,T,P) * RPS(s,r,T,P) \quad (6)$$

This measure enables the appraisal of strict attractiveness phenomena, as it is high only when all the measures in the product are high, for the place of interest p and a subregion s that contains p (e.g., a shop with high ASC and high RSC, located in a busy neighborhood, i.e., having high RPS).

4.4 Algorithm for Calculating the Proposed Measures

Algorithm 1 computes the M-Attract measures. Its inputs are a region r considered for analysis, a set S of subregions of interest contained in r, a set of places P in which each $p \in P$ has a pointer to its respective buffer size ξ_p and minimum staying time δ_p to extract the episodes necessary to calculate their attractiveness measures, and a set T of trajectories that occur inside r. The outputs (pM, sM, and rM) hold the calculated measures for each place of interest $p \in P$, each subregion of interest in $s \in S$, and the region of analysis r, respectively.

Algorithm 1. Compute M-Attract Measures

Require: r, S, P, T

Ensure: $pM[sizeOf(P)], sM[sizeOf(S)], rM$

1: $countEpisodes(r, S, P, T, \&pM, \&rM, \&sM)$;
2: **for all** $s \in S$ **do**
3: **if** $(sM[s].totalStop > 0)$ **then** $sM[s].RPS = \frac{sM[s].totalPass}{rM.totalPass} * \frac{sM[s].totalStop}{rM.totalStop}$;
4: **for all** $p \in P | s\ contains\ p$ **do**
5: **if** $(pM[p].totalPass)$ **then** $pM[p].ASC = pM[p].totalStop \div pM[p].totalPass$;
6: **if** $(sM[s].totalStops)$ **then** $pM[p].RSC = pM[p].totalStop \div sM[s].totalStop$;
7: $pM[p].SA = pM[p].ASC * pM[p].RSC * sM[s].RPS$;
8: **end for**
9: **end for**

First (line 1), $countEpisodes(r, S, P, T, \&pM, \&rM, \&sM)$ processes the whole analysis region, its regions of interest, and their places of interest, along with the trajectories to find the total numbers of episodes *stop* and *pass* in each land parcel. These numbers are stored in the vectors pM, rM, and sM, respectively, to calculate the proposed measures. Then (lines 2 to 11), the algorithm calculates the M-Attract measures, according to the formulas presented in Equations 3 to 6.

We have been using a method to implement the procedure *countEpisodes* that extracts each kind of episode separately. It is based mainly in a generalization of the SMoT algorithm [Alvares et al. 2007]. However, we are working on efficient methods for extracting all these episodes at once. Due to scope and space limitations of this article, we plan to present those methods in future work.

5. EXPERIMENTS

The dataset used in the experiments are legacy geographic data taken from Wikimapia, OpenStreetMap, and GADM[4]. The region considered for analysis was the city of Firenze, in the province of Tuscany, Italy. We considered 14 subregions of the city, and 3125 places (buildings) located inside these subregions, that have a variety of categories in the OpenStreetMap's database. The experiments also used more than 15 thousand trajectories (with over 2 million spatio-temporal points) of residents in Tuscany, collected between 8^{th} and 14^{th} May 2011.

These data were stored in PostgreSQL and managed with PostGIS to implement the algorithm described in Section 4.4. It has run in a i5-3317M 1.7 Ghz processor, with 8Gb of RAM 1066MHz and a 128Gb SSD. It took 93 minutes to process the whole dataset and extract the proposed measures.

In the reported experiments we have used standardized values of buffer size ($\xi = 10$ meters) and minimum staying time to consider a stop ($\delta = 600$ seconds) for extracting episodes in all places. These parameters were chosen based on the kind of moving objects that generated the trajectories. The buffer size of 10 meters is an approximation for parking cars at some distance from the visited place. The time threshold of 600 seconds avoid counting unintentional short stops (e.g., traffic lights). Using these parameters we detected 9,709 episodes *stop* and 377,524 episodes *pass*.

5.1 Results and Discussion

This section reports the insights that the M-Attract measures of attractiveness enabled in our case study. Maps and tables presented in this section show the spatial distribution of these measures for places of interest in different neighborhoods of Firenze. The information is presented using graduated symbols (or colors) according to 10 ranges of measures determined by the the Jenk's method (natural breaks) - only values greater than 0 are shown.

[4] `http://www.gadm.org`

Table I. Top 5 places by *stop* amounts.

Place	stop	pass	ASC
Centro*Gavinana	325	561	0.5793
Esselunga Centro Commerciale	136	501	0.2714
Centro C. Ponte a Greve	130	228	0.5701
Esselunga	119	499	0.2384
Santa Maria Novella	106	949	0.1116

Table II. Top 5 *pass* amounts.

Place	stop	pass	ASC
Piazza Vittorio Veneto	0	3560	0
Supermarket Via Pisana	36	2589	0.0139
Fortezza da Basso	43	1645	0.0261
Beyfin Gas	55	1402	0.0392
Apartments (Via Franc. Berni)	6	1306	0.0045

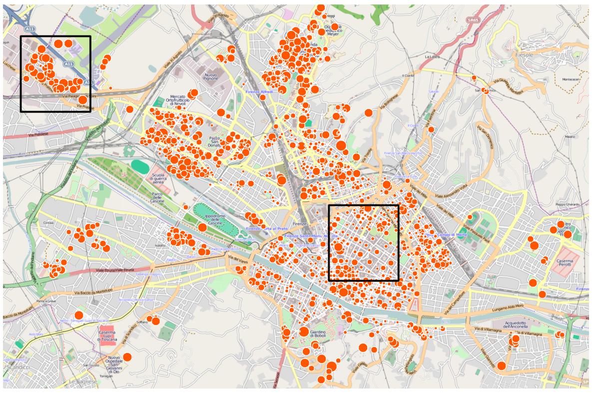

Fig. 4. ASC in places of Firenze (1:50000)

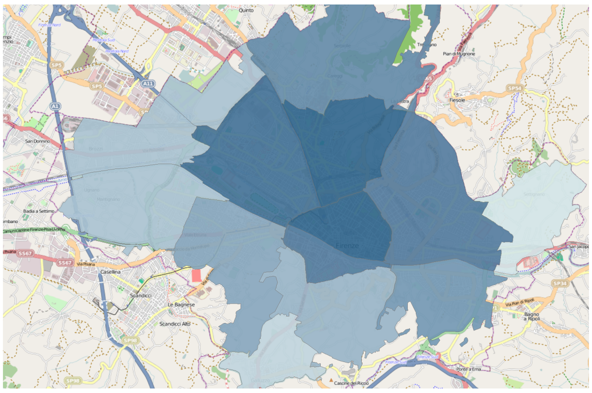

Fig. 5. RPS in neighborhoods of Firenze (1:80000)

Tables I and II list the 10 places with the highest numbers of *stop* and *pass* episodes, respectively. They show that these measures are not enough to explain the attractiveness of places. Some places have a relatively high number of *stop*, but relatively low number of *pass*, making the ratio between these basic measures high. It frequently happens with supermarkets and shopping centers (e.g., Centro C. Ponte a Greve 130/228). Conversely, this ratio is lower for some busy places or places situated in busy neighborhoods (e.g., Santa Maria Novella 106/949). Furthermore, some places have a high number of *pass*, but few *stop* (e.g., Piazza Vittorio Veneto, near the bridge della Vitoria). We call this ratio, formally described in Equation 3, Absolute Stopping Capacity (ASC). It helps to distinguish highly attractive places (e.g., shopping malls, supermarkets) from passage places (e.g., gas stations).

Figure 4 shows that the concentration of high values of ASC is higher in peripheral areas (usually residential) than in central areas (usually commercial) even with the peripheral areas having less episodes *stop*. It happens because usually the ASC also is high for places with relative low number of visits (e.g., homes), located in low movement regions (e.g., residential areas), because a high proportion of moving objects that *pass* near these places, also *stop* in them (several residential buildings has only 1 *stop* and *pass* resulting in a ASC value of 1). The factors RPS and RSC (Equations 5 and 4, respectively) help to solve this problem in the measure SA (Equation 6).

The Relative Passing-Stopping (RPS) of subregions is the proportion of the number of *passes* and *stops* in each subregion s, compared to their total number in the analyzed region r. It differentiates the places according to the movement of subregions where they are located. The distribution of RPS in Firenze neighborhoods is shown in Figure 5 (darker colors represent higher values). As expect the central regions has higher values than the peripheral areas. Although the region in the North also has high value, it happens because the University of Firenze is located there. The Relative Stopping Capacity (RSC) is low for places with low numbers of *stops* in subregions with relatively high numbers of this episode (e.g., a desert alley in a central neighborhood). It differentiates these places strict attractiveness (SA) from that of other places in the same subregion.

Figures 6 and 7 show circles representing the ASC and the number of *stops*, respectively, in a residential area in the west of Firenze. While the ASC values are high, the number of *stops* is very low, with the exception of a supermarket located in the north of this area, which also have a high

Fig. 6. *ASC* in a residential area (1:5000)

Fig. 7. *stops* in a residential area (1:5000)

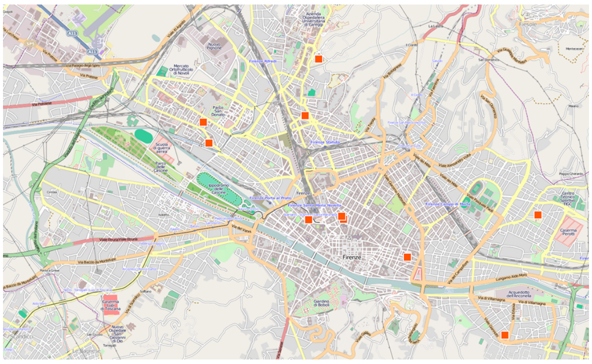

Fig. 8. Top 10 places by SA (1:50000)

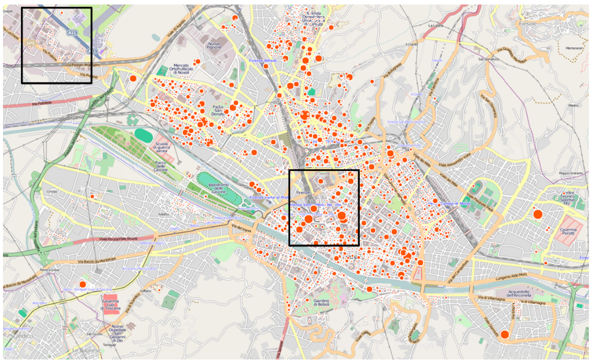

Fig. 9. *SA* in places of Firenze (1:50000)

Table III. Top 10 places by SA

Place Name	#stop	#pass	SA
Coop (Via Galluzzi)	104	276	0.00115
Mercarto Centrale	45	63	0.00063
Esselunga	119	499	0.00045
Coop	61	133	0.00041
Facolta di Architettura / Parcheggi S. Ambrogio	101	558	0.00036
Centro*Gavinana	325	561	0.00033
Parking Area (Via Franc. Baracca)	45	111	0.00028
Santa Maria Novella	106	949	0.00023
Piazza del Mercato Centrale / Parking Area	31	84	0.00022
Ospedale C.T.O.	17	52	0.00016

number of *stops* (e.g., Esselunga Centro Commerciale).

Finally, Figures 8 and 9 illustrate the effectiveness of the measure Strict Attractiveness (SA). Figures 8 shows the 10 places with highest SA, and Figures 9 shows the distribution of SA in places of Firenze. Although high values of SA are more common in the central region (and regions around), there are places with high SA, most of them shopping malls or supermarkets, spread across different areas of the city. The interested reader can found details of the 10 places with highest SA in Table III.

6. CONCLUSIONS AND FUTURE WORK

This article introduces the M-Attract method to assess the attractiveness of places based on collections of moving objects trajectories around these places. M-Attract counts trajectory episodes to compute a family of measures to support analysis of attractiveness phenomena. The main advantages of M-Attract compared to state-of-art methods are: (i) flexibility to work with different kinds of places and regions in varying scales; (ii) parameters to tune the trajectory episodes extraction rules according to the domain, dataset and application at hand (e.g., different parameters can be used when working with cars and people's trajectories); (iii) attractiveness measures with gradually stricter conditions, which combine the number of trajectory episodes in places and regions containing these places; and (iv) the use of real dynamic data of moving objects, which yields more precision to M-Attract than methods that rely on grouped and/or estimated static data (e.g., total population or area). M-Attract enables the assessment of diverse attractiveness phenomena using raw trajectory data, and detects some useful patterns in the spatial distribution of the attractive places.

Our planned future work include: (i) develop efficient algorithms to detect trajectory episodes and compute attractiveness measures on large data collections; (ii) investigate attractiveness measures that can capture temporal aspects (e.g., a Stadium may be attractive only during some events) and consider, among other issues, the duration of the stops (instead of simply counting episodes); (iii) evaluate the effectiveness of M-Attract with other datasets; (iv) apply the M-Attract measures to semantically enrich geographical datasets and trajectory collections for analysis and mining purposes; and (v) employ the proposed measures to detect and solve data privacy issues (e.g., high ASC and few stops usually happen in residential areas, which are usually sensitive to privacy breaches).

REFERENCES

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., and Vaisman, A. A Model for Enriching Trajectories with Semantic Geographical Information. In *Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*. Seattle, USA, pp. 22:1–22:8, 2007.

Christaller, W. *Die Zentralen Orte in Suddeutschland*. Gustav Fischer, 1933.

Furtado, A. S., Fileto, R., and Renso, C. M-attract: assessing places attractiveness based on moving objects trajectories data. In *Brazilian Symposium on Geoinformatics*. Campos do Jordão, BR, pp. 84–95, 2012.

Giannotti, F., Nanni, M., Pedreschi, D., Pinelli, F., Renso, C., Rinzivillo, S., and Trasarti, R. Unveiling the Complexity of Human Mobility by Querying and Mining Massive Trajectory Data. *The VLDB Journal* 20 (5): 695–719, 2011.

Giannotti, F., Nanni, M., Pinelli, F., and Pedreschi, D. Trajectory Pattern Mining. In *Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining*. San Jose, USA, pp. 330–339, 2007.

Huang, L., Li, Q., and Yue, Y. Activity Identification from GPS Trajectories Using Spatial Temporal POIs' Attractiveness. In *Proceedings of the ACM SIGSPATIAL Workshop on Location Based Social Networks*. San Jose, USA, pp. 27–30, 2010.

Mountain, D. and Raper, J. Modelling Human Spatio-temporal Behaviour: a challenge for location based services. In *Proceedings of the International Conference on GeoComputation*. Brisbane, Australia, pp. 24–26, 2001.

Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M. L., Gkoulalas-divanis, A., Macedo, J., Pelekis, N., Theodoridis, Y., and Yan, Z. Semantic Trajectories Modeling and Analysis. *ACM Computing Surveys (to appear)* 45 (4), 2013.

Reilly, W. *The Law of Retail Gravitation*. Reilly, 1931.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., and Vangenot, C. A Conceptual View on Trajectories. *Data and Knowledge Engineering* 65 (1): 126–146, 2008.

Uchino, A., Furihata, T., Tanaka, N., and Takahashi, Y. Some Contribution toward Spatial Urban Dynamics. In *Proceedings of the System Dynamics Conference*. Boston, USA, pp. 1–7, 2005.

Wei, L.-Y., Peng, W.-C., Chen, B.-C., and Lin, T.-W. PATS: a framework of pattern-aware trajectory search. In *Proceedings of the IEEE International Conference on Mobile Data Management*. Kansas, USA, pp. 372–377, 2010.

Yue, Y., Wang, H. d., Hu, B., and Li, Q. q. Identifying Shopping Center Attractiveness Using Taxi Trajectory Data. In *Proceedings of the International Workshop on Trajectory Data Mining and Analysis*. Beijing, CH, pp. 31–36, 2011.

Yue, Y., Zhuang, Y., Li, Q., and Mao, Q. Mining Time-dependent Attractive Areas and Movement Patterns from Taxi Trajectory Data. In *International Conference on Geoinformatics*. Fairfax, USA, pp. 1–6, 2009.