M-Attract: Assessing the Attractiveness of Places by using Moving Objects Trajectories Data

Andre Salvaro Furtado[1 2], Renato Fileto[1], Chiara Renso[3]

[1]PPGCC, Federal University of Santa Catarina (UFSC)
PO BOX 476, 88040-900, Florianópolis-SC, BRAZIL

[2]Geography Department (DG), Santa Catarina State University (UDESC)
Av. Madre Benvenuta, 2007 - Itacorubi, 88035-001, Florianópolis-SC, BRAZIL

[3]KDD LAB, ISTI-CNR, Via Moruzzi 1, 56100, Pisa, ITALY

asalvaro,fileto@inf.ufsc.br,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 interesting attractiveness of places, based on moving objects trajectories. M-Attract examines trajectory episodes (e.g., stop at, pass by) 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. M-Attract has been evaluated in experiments with real data, which demonstrate its contributions to analyze attractiveness of places.*

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 these 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 works use a variety of data to derive attractiveness, including 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) enabled collecting of large volumes of raw trajectories, i.e., sequences of spatial-temporal positions of moving objects. It has pushed the demand for mechanisms to extract useful knowledge from this data. The use of automatic collected trajectory data to derive knowledge about movement in the geographic space can reduce the burden for collecting travel survey data. Furthermore, it can provide more detailed spatial-temporal information about the routes, visited 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 these places can be extracted by processing raw trajectories integrated with geographic data. Spaccapietra [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 spatial-temporal intersections between trajectories and places [Parent et al. 2012].

This article proposes the M-Attract (*Movement-based Attractiveness*) method to assess the attractiveness of places based on raw trajectories. The specific contribution of this method is three-fold: (i) M-attract defines different notions of attractiveness based on the analysis of the trajectories of people moving around the analyzed places; (ii) the notion of attractiveness is based not only on the effective visits to the places but also on the people movements in the geographical context where the places are located in; (iii) all the attractiveness measures we propose are formally defined by properly combining three kinds of trajectory episodes. These measures are defined with gradually higher strictness, in the sense that high values of stricter measures are only achieved by places satisfying more conditions, with respect to trajectory episodes inside them and the region in which they are located. The proposed method is more flexible than state-of-the art ones as it uses parameters for the identification of epsiodes in places and their surrounding regions.

M-Attract has been evaluated in a case study, using local residents private car trajectories in the city of Milan, and geographic data about places and regions of interest collected from several data sources. The results of experiments show that the proposed attractiveness measures allow the identification of several attractiveness phenomena, and the analysis of their spatial distribution in maps.

The rest of this paper 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 in [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 seminar 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 around these regions (up to a certain distance threshold, to compensate possible inaccuracies in trajectory sample points), during a given time period. Several analysis of large volumes of trajectories, based on notions like presence and density of trajectories, are presented in [Giannotti et al. 2011]. These works 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 in [Wei et al. 2010] uses the density-based algorithm introduced in [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 the attractiveness of regions. A trajectory is considered more attractive if it visits more regions with high attractiveness.

The works presented in [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] assess 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 proposed method 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 in [Giannotti et al. 2007], or taken from legacy databases including popular geographic crowdsourcing systems like OpenStreetMap[1] and Mappedia[2]; (ii) it considers real trajectory data from individuals, that can be automatically collected; (iii) 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 land parcels 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

[1] `http://www.openstreetmap.org`

[2] `http://wikimapia.org`

[3] In this article, we consider that a region, a subregion or a place can be represented by a single simple polygon, for simplicity and to avoid further discussions, as the article is subject to size limitations. However, our proposal can be generalized to work with multi-polygons and polygons with roles.

different domains of the application the same land parcel can be seen as a region or as a place - that in our definition is the atomic unity of analysis (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. Depending on the application domain, r can be chosen in a different scale or spatial hierarchy level (if a hierarchy is available). For example, the r to analyze airspace trajectories can cover all the world or a considerable portion of it, the r to analyze long trips can include some countries or provinces, and the r to analyze urban movement can be just a city.

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, for example, subregions s may be city zones or neighborhoods.

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

Places of interest (ρ) 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** τ is a timely ordered sequence of observation points of the form (x, y, t), where x and y refer to the position of the moving object in an instant t.*

The spatial-temporal points of a raw trajectory correspond to sequential observations of the moving object's position along time. These points can be collected by using technologies such as GPS or GSM. Figure 1 shows in its left hand side a representation of the Milan city region (comune), with some subregions of interest (neighborhoods) inside this region, and places of interest inside their respective subregions. The right hand side of Figure 1 shows a set of local residents private car trajectories in this region.

3.3. Categories of Trajectory Episodes considered in M-Attract

Trajectories of moving objects can be used to investigate relations between these objects and other entities in the geographical space. Relevant trajectory episodes, such as stops [Parent et al. 2012], can be inferred from moving objects dynamic attributes such as speed and acceleration, or the continuous period spent inside or close to land parcels of interest. Some of these episodes are useful to determine the attractiveness of places. In this article, we estimate the attractiveness of places by using the following categories of episodes.

Definition 3.5 (**stopAt**(τ, ρ, ξ, δ)). *A trajectory τ is said to $stopAt$ a land parcel ρ when τ continually stays in the buffer of size $\xi \geq 0$ around ρ for at least an amount of time $\delta > 0$.*

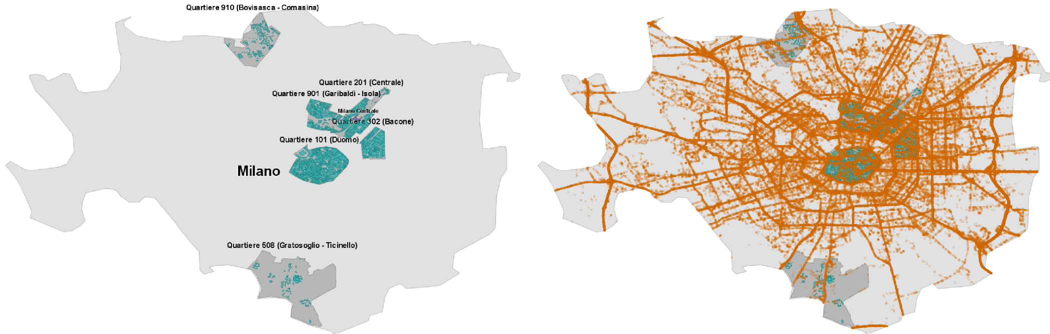

Figure 1. Left: Milan city region, some subregions (neighborhoods), and places of interest; Right: trajectories of private cars inside Milan city.

Definition 3.6 (**passBy**(τ, ρ, ξ)). *A trajectory τ is said to $passBy$ a land parcel ρ when at least one observation point of τ is inside the buffer of size $\xi \geq 0$ enclosing ρ.*

Definition 3.7 (**passIn**(τ, ρ)). *A trajectory τ is said to $passIn$ a land parcel ρ when at least one observation point of τ is inside ρ.*

Figure 2 illustrates these three categories of episodes. Each episode is a trajectory segment (i.e., a subsequence of spatial-temporal observation points) satisfying the respective condition (namely, Definition 3.5, 3.6 or 3.7) with respect to a land parcel ρ. The operator buffer is used in Definitions 3.5 and 3.6 to allow a certain degree of uncertainty for the respective episodes in face of data accuracy and/or interpretation issues (e.g., a car equipped with GPS for collecting trajectories can be parked at a certain distance of place to allow its passengers to visit that place).

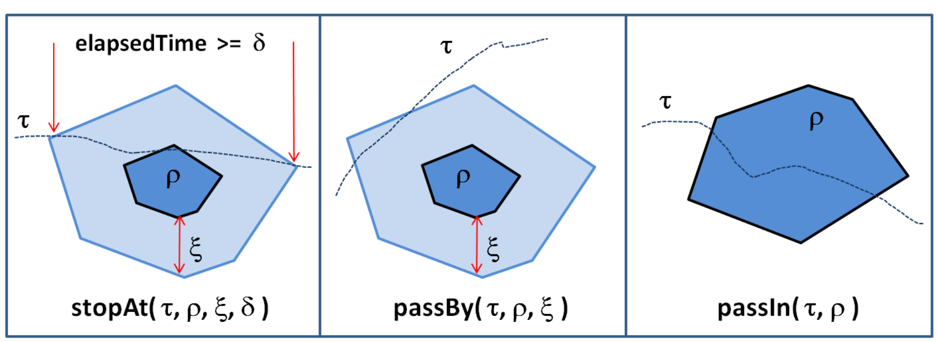

Figure 2. Categories of episodes considered in the proposed method.

We have chosen these three 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.

4. M-Attract Measures

Let Φ be a collection of places as described in Definition 3.3, and Γ be a collection of raw trajectories as described in Definition 3.4. Given a place $\rho \in \Phi$, the number of episodes, as those described in Definitions 3.5 and 3.6, can give some hint of ρ'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 ρ being considered. Thus, in the following we denote these parameters as ξ_ρ and δ_ρ, 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 that are additive across the space hierarchy considered. This ensures that the proposed measures, stated by Equations 1 to 4, 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 ρ, with respect to trajectories from a set Γ that pass close to ρ or stop in any place ρ' contained in the subregion s that contains ρ, respectively.

Absolute Stopping Capacity (ASC) : proportion of $passBy(\tau, \rho, \xi_\rho)$ episodes that also yield $stopAt(\tau, \rho, \xi_\rho, \delta_\rho)$, for a given place ρ, its associated buffer size $\xi_\rho \geq 0$, its minimum staying time $\delta_\rho > 0$, and a trajectory set Γ, as stated by Equation 1.
High ASC intuitively means that a high percentage of people moving in the subregion actually visit the place. This can happen for example when the place have a good advertisement thus attracting people, who was there for other reasons, to stop. Another case of high ASC is when people moves to the subregion on purpose to visit the place and this may mean that the place is isolated in the area or other places have low attractiveness.

$$ASC(\rho, \Gamma) = \frac{\sum\limits_{\tau \in \Gamma} Count(stopAt(\tau, \rho, \xi_\rho, \delta_\rho))}{\sum\limits_{\tau \in \Gamma} Count(passBy(\tau, \rho, \xi_\rho))} \tag{1}$$

Relative Stopping Capacity (RSC) : ratio between the number of stops at a given place ρ and the number of stops in all places ρ' contained in a given subregion s that contains ρ, for their respective buffer size $\xi_\rho, \xi_{\rho'} \geq 0$ their respective minimum staying times $\delta_\rho, \delta_{\rho'} > 0$, and a trajectory set Γ, as stated by Equation 2.
RSC gives a measure of the stop capacity of a place compared to other places in the subregion. High RSC for a place means that it is highly visited and it is located in a subregion with other places which are rarely visited.

$$RSC(\rho, s, \Gamma, \Phi) = \frac{\sum\limits_{\tau \in \Gamma} Count(stopAt(\tau, \rho, \xi_\rho, \delta_\rho))}{\sum\limits_{\substack{s \text{ contains } \rho' \\ \tau \in \Gamma,\ \rho' \in \Phi}} Count(stopAt(\tau, \rho', \xi_{\rho'}, \delta_{\rho'}))} \tag{2}$$

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 ρ, 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 total number of $passIn$ referring to places in subregion s and to places in the region r multiplied by the relative number of $stopAt$ referring to places contained in s and to places contained in the whole analyzed region r, for trajectories set Γ (Equation 3).

$$RPS(s,r,\Gamma,\Phi) = \frac{\sum_{\tau\in\Gamma} Count(passIn(\tau,s))}{\sum_{\tau\in\Gamma} Count(passIn(\tau,r))} * \frac{\sum_{\tau\in\Gamma,\ \rho'\in\Phi}^{s\ contains\ \rho'} Count(stopAt(\tau,\rho',\xi_{\rho'},\delta_{\rho'}))}{\sum_{\tau\in\Gamma,\ \rho''\in\Phi}^{r\ contains\ \rho''} Count(stopAt(\tau,\rho'',\xi_{\rho''},\delta_{\rho''}))} \tag{3}$$

4.3. Attractiveness of Places

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

Strict Attractiveness (SA) : product of the absolute stopping capacity of a place ρ, the relative stopping capacity of ρ with respect to a subregion s containing ρ, and the relative passing and stopping of s (Equation 4).

$$SA(\rho,s,r,\Gamma,\Phi) = ASC(\rho,\Gamma) * RSC(\rho,s,\Gamma,\Phi) * RPS(s,r,\Gamma,\Phi) \tag{4}$$

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 ρ and a subregion s that contains ρ (e.g., a commercial center high ASC and high RSC with respect to a busy neighborhood, i.e., a neighborhood with high RPS).

4.4. Algorithm for Calculating the Proposed Measures

Algorithm 1 computes the proposed M-Attract measures. Its inputs are a region r considered for analysis, a set S of subregions of interest contained in r, a set P of records where each $p \in P$ has a pointer to a place of interest $\rho \in \Phi$ with the respective buffer size ξ_ρ and minimum staying time δ_ρ to extract the trajectory episodes necessary to calculate their attractiveness measures, and a set Γ of trajectories that occur inside r. The outputs (pM, sM, and rM) hold the calculated measures for each place of interest $p.\rho | p \in P$, each subregion of interest in $s \in S$, and the region of analysis r, respectively.

First (line 1), the total number of episodes $stopAt$, $passBy$ and $passIn$ in each land parcel are extracted by calling $CountEpisodes(r, S, P, \Gamma, \&pM, \&rM, \&sM)$. It processes the land parcels and trajectories to found the episodes necessary to calculate the proposed measures and stores the number of each kind of episode found in each place of interest, each region of interest, and the whole analysis region, in the vectors pM, rM, and sM, respectively. Then (lines 2 to 11), the algorithm calculates the M-Attract measures, according to the formulas presented in Equations 1 to 4.

Algorithm 1. Compute M-Attract Measures

INPUT: r, S, P, Γ

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

$CountEpisodes(r, S, P, \Gamma, \&pM, \&rM, \&sM)$;
for each $s \in S$ **do**
  if $(sM[s].totalStops > 0)$ **then**
    $sM[s].RPS = \frac{sM[s].totalPassIn}{rM.totalPassIn} * \frac{sM[s].totalStops}{rM.totalStops}$;
  for each $p \in P | s\ contains\ p.\rho$ **do**
    if $(pM[p].totalPassBy)$ **then**
      $pM[p].ASC = \frac{pM[p].totalStopAt}{pM[p].totalPassBy}$;
    if $(sM[s].totalStops)$ **then**
      $pM[p].RSC = \frac{pM[p].totalStops}{sM[s].totalStops}$;
    $pM[p].SA = pM[p].ASC * pM[p].RSC * sM[s].RPS$;
  end for
end for

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 Milan, Italy. We have selected 40 subregions of the city (central, transition and peripheral areas), and 16044 places (buildings) inside these subregions with a variety of categories from OpenStreetMap's database. The experiments also used more than 10 thousand trajectories of local resident's private cars in Milan, collected between 5^{th} and 7^{th} April 2007.

These data were stored in PostgreSQL and managed with PostGIS to implement the algorithm described in Section 4.4. It has run in a i7-620M 2.66Ghz processor, with 4Gb of RAM 1066MHz and a 320Gb Hard Drive 7200RPM. It took 4 hours to process the whole dataset to extract the proposed measures to assess the attractiveness of places.

In the reported experiments we have used standardized values of buffer size ($\xi = 30$ meters) and minimum time to consider a stop ($\delta = 120$ seconds) to extract episodes from all places. These parameters were chosen based on the kind of individuals that trajectories were collected. The buffer size of 30 meters is an approximation for parking cars at some distance from the visited place. The time threshold of 120 seconds avoid counting unintentional short stops (e.g., traffic lights):

- 16280 $stopAt$, in 5561 distinct trajectories, and 5360 distinct places of interest;
- 232801 $passBy$, in 8246 distinct trajectories, and 14467 places of interest;
- 42145 $passIn$, in 9439 distinct trajectories, and 40 distinct subregions of interest.

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

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 shows the spatial distribution of these measures in places of interest in different neighborhoods of Milan. The size of the circle at each place is proportional to the respective measure for that place.

Tables 1 and 2 list the 10 places with the highest numbers of $stopAt$ and $passBy$ episodes, respectively. They show that these measures are not enough to explain the attractiveness of places. Some places have a relatively high number of $stopAt$, but relatively low number of $passBy$, making the ratio between these basic measures high. It frequently happen with supermarkets and shopping centers (e.g., Bicocca Village 70/83). Conversely, this ratio is lower for some busy places or places situated in busy neighborhoods (e.g., Milano Centrale 58/261). Furthermore, some places have a high number of $passBy$, but few $stopAt$ (e.g., Cascina Gobba 12/300, near A51 highway). We call this ratio, formally described in Equation 1, Absolute Stopping Capacity (ASC). It helps to distinguish highly attractive places (e.g., shopping malls, supermarkets) from passage places (e.g., parking lots, train stations). However, the ASC is sometimes high also for places with relative low number of visits (e.g., homes), located in low movement regions (e.g., residential areas) (see Table 3), because a high proportion of moving objects that $passBy$ these places, also $stopAt$ them. The factors RPS and RSC (Equations 3 and 2, respectively) help to solve this problem in the measure SA (Equation 4).

Place Name	StopAt	PassBy	ASC
Metropoli	154	177	0.8700
Esselunga di Via Ripamonti	80	109	0.7339
Bicocca Village	70	83	0.8433
Milano Centrale	58	261	0.2222
Centro Commerciale Bonola / Parking Lot Via Antonio Cechov	53	111	0.4774
Centro Commerciale Piazza Lodi	47	130	0.3615
Galleria Manzoni	45	109	0.4128
Mango Italia	43	95	0.4526
Lounge Milano / Hollywood	41	128	0.3203
Eselunga di Via Lorenteggio / Parcheggio Soterraneo	41	66	0.6212

Table 1. Top 10 $stopAt$ amounts.

Place Name	StopAt	PassBy	ASC
Cascina Gobba	12	300	0.04
Unes	6	300	0.02
Parking Viale Enrico Forlanini	8	299	0.0267
Forno Antico	7	287	0.0243
Intesa Sanpaolo	14	283	0.0494
Aikido Di Fujimoto Yoji	4	280	0.0142
Europarco Srl Noleggio Furgoni	0	272	0
Parking - Viale Mugello	2	268	0.0074
Parking - Viale Corsica	8	263	0.0304
Milano Centrale	58	261	0.2222

Table 2. Top 10 $passBy$ amounts.

Place Name	StopAt	PassBy	ASC
Apartments (Via P. Fiuggi, 19)	5	5	1
Starhotels Tourist	6	6	1
Apartments (Viale dell'Aviazione, 62-72)	1	1	1
Apartments (Via Donna Prassede,2)	6	6	1
Houses (Via Privata Faiti, 1-9)	1	1	1
Apartments (Via Val Maira)	1	1	1
Apartments (Via Luigi Bertelli)	6	6	1
Asilo Nido	2	2	1
Apartments (Via San Mirocle)	6	6	1
House (Via Gaetano Crespi)	1	1	1

Table 3. Top 10 ASC.

Place Name	StopAt	PassBy	SA
Metropoli	154	177	0.00198
Bicocca Village	70	83	0.00098
Esselunga di Via Ripamonti	80	109	0.00097
Esselunga di Via Rubattino	38	81	0.00082
Esselunga - Missaglia	40	43	0.00062
Mediaworld	24	48	0.00055
Mango Italia	43	95	0.00041
Galleria Manzoni	45	109	0.00039
Esselunga di Via Novara	34	51	0.00038
Milano Centrale	58	261	0.00036

Table 4. Top 10 SA.

Figure 3 shows the distribution of the number of $stopAts$ and $passBy$ episodes, which are more concentrated in central neighborhoods than in peripheral ones. Figure 4 (left) shows that the concentration of high values of ASC is higher in peripheral areas. By comparing this distribution with that of the number of $stopAt$, it is possible to distinguish the patterns found in commercial areas from those of residential areas. It can be observed in more detail in Figure 6, that plot these measures for a central commercial area (Duomo) and a peripheral residential area (Gratosoglio - Ticinello).

Figure 3. Distribution of $stopAt$ (left) and $passBy$ (right) in places of Milan.

The Relative Stopping Capacity (RSC) of places decreases for places with low number of $stopAt$ in subregion with relatively high numbers of this episode (e.g., a desert alley in a central neighborhood). It differentiates these places attractiveness from that of other places in the same subregion. The Relative Passing-Stopping (RPS) of subregions is the proportion of the number of $passIn$ and $stopAt$ 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 Milan neighborhood is shown in Figure 5 (darker colors represent higher values).

Finally, Figure 7 illustrates the effectiveness of the measure Strict Attractiveness (SA). Its left side shows the 10 places with highest SA, and its left side shows the distribution of SA in places of 40 Milan neighborhoods. Although high values of SA are concentrated in the city center, 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 4.

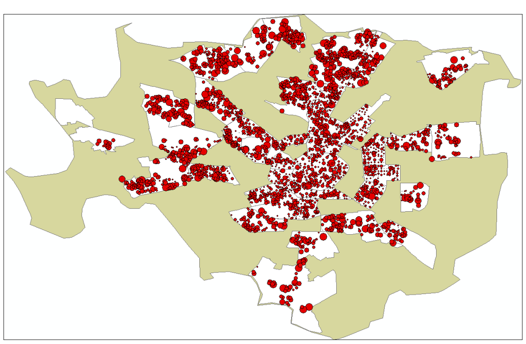

Figure 4. ASC in places of Milan.

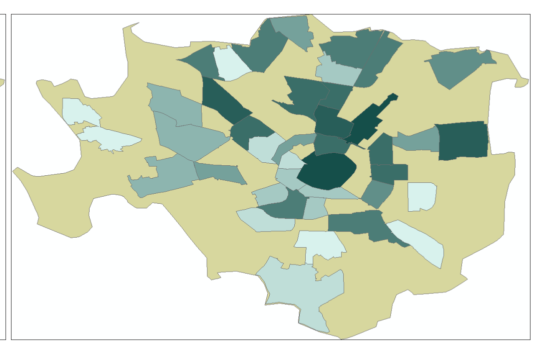

Figure 5. RPS in neighborhoods.

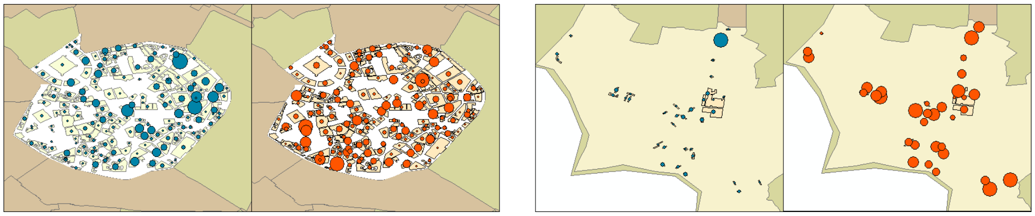

Figure 6. $stopAt$ **versus** ASC **in Duomo (left), and Gratosoglio - Ticinello (right).**

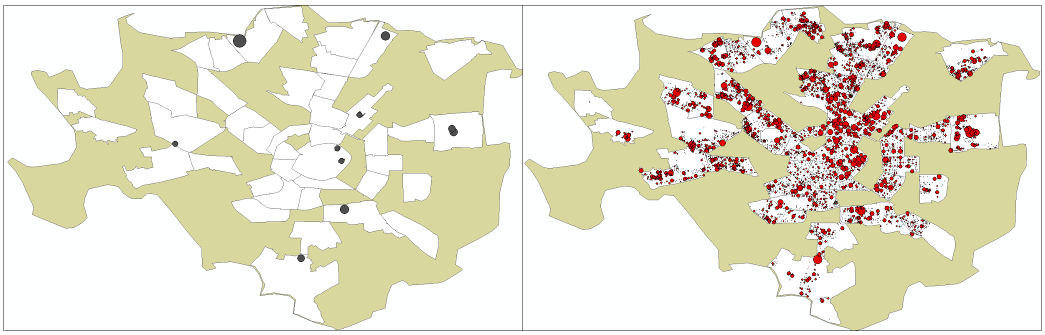

Figure 7. Top 10 attractive places (left) and SA **(right) in places of Milan.**

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 empirically defined measures to support analysis of attractiveness phenomena. The main advantages of this method 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 individuals giving more precision 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, detecting some useful patterns in a set of places spatial distribution from raw trajectory data.

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 Sports Stadium can be attractive only when an event is happening) and consider among other variables the duration of the stops (instead of simple counting episodes); (iii) evaluate the effectiveness of M-Attract with other datasets; and (iv) apply the M-Attract measures to semantically enrich geographical datasets and trajectory collections for searching and mining purposes.

Acknowledgments

Work supported by CAPES, CNPq (grant 478634/2011-0), and EU-IRSES-SEEK project (grant 295179). Thanks to Vania Borgony, for the on time criticism and incentive.

References

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., and Vaisman, A. (2007). A model for enriching trajectories with semantic geographical information. In *Proc. of the 15th annual ACM Intl. Symp. on Advances in GIS*, ACM-GIS, pages 22:1–22:8, New York, NY, USA. ACM.

Christaller, W. (1933). *Central places in Southern Germany (in German).*

Giannotti, F., Nanni, M., Pedreschi, D., Pinelli, F., Renso, C., Rinzivillo, S., and Trasarti, R. (2011). Unveiling the complexity of human mobility by querying and mining massive trajectory data. *The VLDB Journal*, 20(5):695–719.

Giannotti, F., Nanni, M., Pinelli, F., and Pedreschi, D. (2007). Trajectory pattern mining. In *Proc. of the 13th ACM Intl. Conf. on Knowledge Discovery and Data Mining*, SIGKDD, pages 330–339, New York, NY, USA. ACM.

Huang, L., Li, Q., and Yue, Y. (2010). Activity identification from GPS trajectories using spatial temporal POIs' attractiveness. *Proc. of the 2nd ACM SIGSPATIAL Intl. Workshop on Location Based Social Networks - LBSN '10*, page 27.

Mountain, D. and Raper, J. (2001). Modelling human spatio-temporal behaviour: a challenge for location based services. In *Proc. of the 6th Intl. Conf. on GeoComputation*, pages 24–26, Brisbane, Australia.

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. (2012). Semantic trajectories modeling and analysis. *ACM Computing Surveys (to appear).*

Reilly, W. (1931). *The law of retail gravitation*. W.J. Reilly.

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

Uchino, A., Furihata, T., Tanaka, N., and Takahashi, Y. (2005). Some contribution toward Spatial Urban Dynamics (From relative attractiveness point of view). In *Proc. of the System Dynamics Conference*.

Wei, L.-Y., Peng, W.-C., Chen, B.-C., and Lin, T.-W. (2010). Pats: A framework of pattern-aware trajectory search. *Proc. of 11th IEEE Intl. Conf. on Mobile Data Management*, pages 372–377.

Yue, Y., Wang, H. d., Hu, B., and Li, Q. q. (2011). Identifying shopping center attractiveness using taxi trajectory data. In *Proc. of Intl. Workshop on Trajectory Data Mining and Analysis*, TDMA, pages 31–36, New York, NY, USA. ACM.

Yue, Y., Zhuang, Y., Li, Q., and Mao, Q. (2009). Mining time-dependent attractive areas and movement patterns from taxi trajectory data. In *Geoinformatics, 2009 17th Intl. Conf. on*, pages 1–6.