Looking inside the Stops of Trajectories of Moving Objects

Bruno N. Moreno¹, Valéria C. Times¹, Chiara Renso², Vania Bogorny³

¹ Centro de Informática, UFPE, Brazil
² KDDLAB, ISTI-CNR, Pisa, Italy
³ Departamento de Informática e Estatística, UFSC, Brazil

{bnm,vct}@cin.ufpe.br, chiara.renso@isti.cnr.it, vania@inf.ufsc.br

Abstract. Trajectory data are normally generated as sample points, which are very difficult to understand and to analyze because they are often collected with no semantic information. Several studies have been developed for trajectory data analysis. Recently, a new model was designed to reason over trajectories as stops and moves, where stops are the important parts of trajectories. Based on this work, different methods have been developed to instantiate this model, based on different characteristics like speed and direction, aiming to give more semantics to trajectories. In this work we go one step forward to existing works that compute stops of trajectories. We evaluate the behavior of a trajectory considering first its geometric properties like velocity and direction change, and then, based on this analysis we propose to use domain knowledge that describes some characteristics of the application domain to infer the goal of the stops. To validate the proposed method we present some experiments over real trajectory data.

Keywords
Moving objects, automatic semantic trajectory annotation, movement behavior, stops and moves, trajectory patterns

1. INTRODUCTION

The price reduction of mobile devices like mobile phones, GPS, and RFID has significantly increased their use for several objectives. This has generated large amounts of data that can be explored for several application domains, like traffic management, animal migration, human behavior in a shopping mall, etc.

The mobile devices leave behind spatio-temporal traces that characterize the trajectory of a moving object. Trajectory data are normally generated as sample points, which are very difficult to understand and to analyze because they are often collected with no semantic information. It is even more difficult to extract implicit and previously unknown patterns from this data.

Several works have been developed for trajectory data analysis. One group of works has developed methods to generate patterns focusing on the geometrical properties of trajectories and defining types of trajectory patterns like convergence, encounter, flock, leadership, etc [12]. Another group of works has focused on the analysis and mining of trajectory sample points, basically considering time and space. Some examples include the extraction of clusters of trajectories located in dense regions...
[15], groups of trajectories that move between regions in the same time interval [8], patterns of trajectories with similar shapes [11], or with similar distances [18].

More recently some works started focusing on the analysis of trajectories from a semantic point of view, trying to add context information. Guc et. al [9], for instance, developed a method where the user manually annotates the trajectories with the interesting points.

In 2008, Spaccapietra proposed the first data model looking at trajectories from the conceptual point of view [20]. In this approach, a trajectory is a set of important places called stops. From this starting point, different works have been proposed to instantiate the model of stops, like [1], [12], and [17]. Alvares [1] proposed an approach that identifies the important parts of trajectories considering as context information a set of geographic information available for the region where the trajectories were collected. In [12] an algorithm is proposed to instantiate stops based on the variation of the direction of the trajectory. The work of Palma [17] computes the important places (stops) of trajectories by finding the regions where the velocity is lower than the average speed of the trajectory.

Figure 1 shows three examples of trajectories. In Figure 1.1, a trajectory sample point is presented. In Figure 1.2 a semantic trajectory is presented for a tourism application, and in Figure 1.3 a semantic trajectory is presented for a traffic management application. In Figure 1.2 stops were computed by the method IB-SMoT, proposed by [1], where the only semantic information considered is the geographic location. In this work, a trajectory must intersect a previously defined geographic location for a minimal amount of time. In Figure 1.3, the method CB-SMoT was used to find low speed stops, among which, some are known geographic places (e.g. airport), but two are clusters that do not intersect any geographic location given a priori.

Figure 1 - Examples of trajectories

One problem of the method CB-SMOT is that it finds the regions with low velocity, but if these regions do not intersect a known geographic location, these places are labeled as unknown stops, as shown in Figure 1.3. Besides this problem, how could we distinguish for instance, in two animal stops if they are either feeding or resting? How can we distinguish stops at an open shopping in downtown or in a big shopping mall if people are shopping, eating at a restaurant, watching a movie, or working?
In this work we propose to go one step forward to existing works that have focused on the generation of stops as the identification of important parts of trajectories. We propose a novel method to look inside the stop analyzing the behavior of the moving object to infer the goal of the stop. This method performs basically two main steps: first, it evaluates the behavior of the trajectory considering the geometric properties like velocity and direction change, and second, it makes use of domain knowledge to infer the goal of the stop based on the pattern of speed and direction change.

The remainder of this paper is organized as follows. Section 2 presents some related works and the main contribution. Section 3 presents the basic definitions of the proposed method and an algorithm to infer the goal of the stop. Section 4 gives experimental results and Section 5 concludes the paper and suggests future work.

2. Related Works and Contribution

Existing works for adding more semantic information to trajectories can be classified in two groups: the works like [3] that annotate, normally manually, information to trajectories, and the works that follow the model of stops and moves [19]. We follow the second approach, which allows us to automatically discover context information to enrich trajectories, independently of application domain, since this approach is more generic.

The method presented in this article gives meaning to certain points of a trajectory, which correspond to the important places and are called stops. Stops are application dependent, and are automatically generated by a method that is the most appropriate for the domain.

In this section, we summarize three methods to find stops in trajectories that are closely related to our work and that deal with single trajectories: IB-SMoT, CB-SMoT and DB-SMoT. IB-SMoT [1] generates groups of trajectory points based on the intersection of these points with geographic objects defined as relevant to the application domain. This intersection must meet a minimum time threshold, such that the subtrajectory should continuously intersect the geographic object for the minimum time. These intersected places can be hotels, schools, etc. The main problem of this method is that for several applications there might be no geographic information.

CB-SMOT [15] is a clustering method based on the variation of trajectory speed. This method has basically two steps: first it evaluates each trajectory and generates clusters formed by substratajectories in which the speed is lower than a given threshold, called \( \text{avgSpeed} \), for a minimal amount of time (\( \text{minTime} \)). For example, if the average speed of the trajectory is 100 km/h and the \( \text{avgSpeed} \) is specified as 50 km/h, all substratajectories in which the speed is lower than 50 km/h for at least \( \text{minTime} \), will be labeled as unknown stop. In a second step, the method tests the intersection of the clusters with a set of user defined candidate stops, which are geographic objects relevant to the application. All clusters (unknown stops) that intersect the geographic objects for a minimal amount of time will be labeled with the name of the geographic object, otherwise they remain as unknown stops.

The algorithm DB-SMoT [13] is also a clustering method, but clusters on single trajectories are generated based on the variation of the trajectory direction. This method
is interesting in specific domains where the direction variation has a greater impact than speed. Clusters are generated for subtrajectories where the direction variation is lower than a given threshold $minDir$ and for a minimal amount of time $minTime$.

Existing methods to find stops are unable to either discover the behavior of the moving object or the goal of the trajectory or a stop. One main reason is because only objective measures are used, like speed, direction and time. The main contribution of our proposal is to go one step forward by looking inside the stop, considering not only objective measures, but semantic information stored into a knowledge base about the domain to infer the behavior.

In summary, the main contributions of the paper are (I) an algorithm that is generic enough to enrich trajectory important places with semantic information based on the behavior of the moving object, in different application domains; (II) the use of domain knowledge to interpret and understand traces of moving objects in order to use this information for decision making processes in applications like urban planning, animal migration, marketing, etc.

3. The Proposed Method
In this section we first present some new concepts and definitions that may be useful to understand the proposed approach and then we present the algorithm that makes use of these concepts.

3.1 Basic Definitions
According to [20], a trajectory is the user defined record of the evolution of the position (perceived as a point) of an object that is moving in space during a given time interval in order to achieve a given goal.

**Definition 3.1 (Trajectory Context):** Trajectory Context is a set of conditions and influences used to identify why a mobile object has stopped while it is moving in space during a given time interval.

Context information provides the ability to discriminate what is important or not at any given time [20]. Context information can be geographic (where and when the object has gone) or behavioral (how and why the object executed the movement) or about recognition (who is the mobile object or what has moved) [5] [16] [22]. The context information about trajectories allows the movement of the mobile object to be tracked and understood in order to enable the planning of future actions for certain types of events or situations, or even find groups of objects with similar behavior.

In this paper, context information is seen as the knowledge that indicates the reason or the purpose of the movement of the mobile object because in this work, as indicated in Definition 3.1, we are interested in defining why a mobile object has stopped. The proposed method investigates the goal of the trajectory by analyzing its stops. The definition of stop given in [17] is therefore extended to include the goal of a stop at a given time and location.

**Definition 3.2 (Contextualized Stop):** A contextualized stop represents an important place of a trajectory in which the mobile object has been for a minimal amount of time and for a given reason.
The proposed method makes use of contextual information about the stops and spatio-temporal data on the movement of mobile objects to infer the reason why the mobile object has performed a given stop. To perform this type of inference, we consider that each stop has a sub-trajectory. In order to identify different sub-trajectories inside stops we apply again the clustering algorithm used to compute stops (based on direction or speed variation), in order to lead to new stops. These new stops are defined as contextualized substops.

**Definition 3.3 (Contextualized Substop):** A contextualized substop is a stop of a subtrajectory such that:

(i) Its goal is derived from a set of rules; and

(ii) It is part of the goal of the contextualized stop that represents the subtrajectory.

The proposed method accesses a knowledge base represented by a set of rules and checks if the sub-trajectory of each substop satisfies one or more rules. Each rule of the knowledge base represents a possible goal. The set of goals inferred by the method for all substops of the subtrajectory summarize the purpose of the contextualized stop. Figure 2 shows an example of a knowledge base describing the behavior of pedestrians inside a shopping center. The first rule expresses that when a moving object stays for at least 2 hours without moving (speed zero and direction variation zero), the object is at a cinema. The third rule expresses that if the moving object stays for at least 8 hours, with speed lower than 0.5 km/h and little direction variation no more than 10 degrees, the goal of the stop is working.

<table>
<thead>
<tr>
<th>minTime</th>
<th>maxSpeed</th>
<th>maxDirection</th>
<th>goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 hours</td>
<td>0</td>
<td>0</td>
<td>cinema</td>
</tr>
<tr>
<td>1 hour</td>
<td>1.5 km/h</td>
<td>20 degrees</td>
<td>shopping</td>
</tr>
<tr>
<td>8 hours</td>
<td>0.5 km/h</td>
<td>10 degrees</td>
<td>working</td>
</tr>
</tbody>
</table>

**Figure 2. Example of a knowledge base**

The objective of this work is not limited to detect the interesting places visited by a moving object, but what he/she was doing at these places. Figure 3 gives a more clear idea of this work. Each stop has a purpose, a goal, for a given moving object (person, animal, etc.), and this stop occurs at some place. This means that the goal of a stop, is one more aspect of the stop concept. The goal of the stop can be an activity that is assumed to be done by the moving object at a specific place. The specific activity depends on the application and can be found in a knowledge base, such as an ontology. The same occurs for a substop, it has a goal and occurs at some place. For instance, the most generic goal of a stop at a shopping center can be for entertainment, and what we intent to discover is what the purpose/goal of each substop is (e.g. watching a movie, eating at a restaurant, shopping).
3.2 The Proposed algorithm

Listing 1 presents the algorithm that illustrates the proposed approach. The algorithm receives as input a set of stops $S$ computed by any of the existing methods in the literature, that the user chooses as the most interesting for the application domain, and a knowledge base about the domain, $kbase$.

```plaintext
1 INPUT:
2 $S$ //set of stops
3 kBase //Knowledge Base
4
5 OUTPUT: B // a set of contextualized substops
6
7 METHOD:
8    subStops -> CB-SMoT($S$);
9  subStops -> subStops + DB-SMoT($S$);
10  IF (subStops!= {})
11    B -> executeInference(subStops, kBase);
12 END
```

Listing 1 – Computing Substops psudo-code

The output of the method is a set of contextualized substops. In a first step the algorithm computes substops based on both speed variation and direction variation. For this, it calls the algorithm CB-SMoT (line 8) to compute low speed clusters/substops for all stops $S$, generating substops. In a second step, substops are computed over the same set of stops $S$ with the method DB-SMoT (line 9), which finds clusters/substops where the direction of the moving object has changed. Of course both methods require input parameters as $minDir$, $avg$ and $minTime$, but we consider this as part of the methods CB-SMOT and DB-SMoT that are well known in the literature. These parameters can also be obtained by the methods from the knowledge base.

If a list on non-empty set of substops is generated, then these substops will be analysed using a knowledge base, through the method named executeInference (line 11), that is explained in Listing 2. If a stop has no substops, the stop itself is inserted in the set of substops.
The inference method presented in Listing 2 receives as an input a set of substops and the knowledge base. The output is a set of contextualized stops/substops C. For each stop s in substops (line 9) the method recovers the duration of the stop (timeStop) and the speed and direction variation of the substop (lines 11 and 12), previously computed by the methods CB-SMOT and DB-SMOT. Then for each rule/row in the knowledge base (line 13) the speed and direction variation of the stop are then compared with the maximum speed and maximum direction variation stored in the rule (line 17). If the speed variation of the substop is lower or equal to the speed variation of the rule and the direction variation is lower or equal to the direction variation of the rule, than the time is tested.

If the stop duration is equal or greater to the minimal time defined in the rule (line 18), than the goal of this substop is found (line 29) in the knowledge base and a contextualized substop is inserted in the set of contextualized substops (line 20).

Listing 2 – executeInference pseudo-code

In the following section we present some initial results obtained with the proposed method.

4. EXPERIMENTS AND EVALUATION

We performed experiments with two different datasets: a bird dataset and a dataset of pedestrians in a park, as explained in the following sections.

4.1 Bird dataset

A first experiment was performed over the ceconia bird’s trajectory dataset. These data were collected during the migration of Ciconias, being most birds fitted with geographical positioning devices. The acquired data were transmitted to a group of
researchers\(^1\) who gave a name to each bird. The whole dataset has only 1886 records. As a first experiment, we chose the trajectory of the bird *Prinzesschen*, which has more points. From the total of 1886 records, this bird has 528 points.

Four stops were generated for this trajectory, and considered as input for the proposed method, as shown in Figure 4(1). For two stops (stop 2 and 3), one substop was generated, as shown in Figure 4(2). After the substops generation, the next step is to infer the objective of each substop. Using a fictitious knowledge base, shown in Figure 5, we discovered the goals of two substeps, as shown in Figure 4(2). It is important to note that the information contained on this base was prepared without the aid of an expert in the area of birds, such as an ornithologist. However, assuming that such information will be defined by a domain expert, this does not affect the experimental results.

![Figure 4 - Trajectory stops and their respective sub-stops](image1)

<table>
<thead>
<tr>
<th>minTime</th>
<th>maxSpeed</th>
<th>maxDirection</th>
<th>goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 hours</td>
<td>0 km/h</td>
<td>0 degrees</td>
<td>Resting</td>
</tr>
<tr>
<td>2 hours</td>
<td>6.5 km/h</td>
<td>30 degrees</td>
<td>Feeding</td>
</tr>
<tr>
<td>3 hours</td>
<td>10 km/h</td>
<td>60 degrees</td>
<td>Hunting</td>
</tr>
</tbody>
</table>

**Figure 5. Fictitious knowledge base**

The bird trajectory dataset is not very interesting for evaluating the proposed method because there are only a few points, with long gaps in time. Therefore, we evaluated the method with another real dataset as shown in the following section.

4.2 Pedestrian dataset

A second experiment was performed over a pedestrians dataset generated at a park in The Netherlands [21]. A set of people were equipped with a GPS device and each person was asked with the activity that he/she would do at the park. Among these activities, some were walking, running, walk the dog, picnic, etc. These data, differently from the birds dataset has the points very close in time, in an average of about every 10 seconds. Based on the metadata send with the data, we created a knowledge base with some activities that the pedestrians would perform in the park, as shown in Figure 6. Basically, what characterizes the behavior of the moving object is the speed variation and direction change. For instance, two activities have as minimal time 15 minutes (walking and cycling), but we suppose that if a person is walking the maximal speed is 7km/h, while cycling would be at 36 km/h. The knowledge base is fictitious, but the objective is to show that the method is able to give more meaning to trajectories considering prior knowledge.

<table>
<thead>
<tr>
<th>minTime (min)</th>
<th>maxSpeed (km/h)</th>
<th>maxDirection</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>7</td>
<td>50</td>
<td>walking</td>
</tr>
<tr>
<td>15</td>
<td>36</td>
<td>80</td>
<td>cycling</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>90</td>
<td>dog letting</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>45</td>
<td>photo</td>
</tr>
<tr>
<td>60</td>
<td>2</td>
<td>20</td>
<td>picnic</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>20</td>
<td>relaxing</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>50</td>
<td>running</td>
</tr>
</tbody>
</table>

Figure 6. Knowledge base of possible activities in a park

This experiment was performed over 246 trajectories. We generated stops with the method CB-SMoT, in order to give the input of our method. A total of 148 stops were generated, considering 30 minutes as the minimal time and the speed should be half of the average speed of the trajectory. So the input of our method were 148 stops. Among these stops, a total of 494 substops were generated using the knowledge base parameters to generate them. Among the 494 substops, 160 had their objective inferred.

Because of space limitations, we show the result of only two contextualized trajectories, which are shown in Figures 7 and 8. Figure 7 left shows an example of a trajectory that had one big stop as input and 5 substops were generated (Figure 7 right). From the 5 substops, 4 have been contextualized, in taking pictures (photo) and walking.

Figure 8 left shows an experiment on a single trajectory that also had one big stop as input and 3 small substops were generated with our method, among which one was contextualized (Figure 8 right). Additional experimental results are available at http://cin.ufpe.br/~bnm/lookingInsideStops.
4.3 Discussion

It is obvious that from the semantic point of view the results of the experiments would have to be evaluated by a specialist in the application domain in order to semantically validate them. The knowledge bases would also have to be build by a domain expert. However, the experiments presented in this paper have the objective to show the effectiveness of the method, that the proposed algorithm is able to infer the activities/goals of individual trajectories through the analysis of the behavior of the moving object and the use of domain information, what is novel in this research field.

So far there are no similar methods that infer the goal of a trajectory through the analysis of its direction and speed variation, and its interpretation using domain knowledge. Existing works use only the spatial intersection of stops with geographic information that for several applications is not available or may not help to infer the goal of trajectories for decision making processes.
5. CONCLUSIONS AND FUTURE WORKS

Several studies have focused on the semantic properties of trajectories. Most of them propose different objective measures to instantiate the well known model of stops and moves. The novelty of this paper is the analysis of speed and direction variation, for a certain time, and the use of context information stored in a knowledge base to infer new and more knowledge about important places of trajectories (stops). With this discovery we are able to infer the behavior of the moving object and understand the goals of his/her trajectory. Among other objectives, this method can be used for two main reasons: to discover the meaning of an unknown stop or to discover the activity of the moving object inside a known stop. The possibility of using a knowledge base on the analysis of stopping points of a moving object brings great benefits to trajectory data mining. Semantic trajectories can be used to discover common group behavior patterns.

The main drawback of this work is that for one stop we may infer more than one goal. For instance, if a moving object is at a restaurant in shopping mall or at a cinema, both stops have speed zero and direction variation zero. Future ongoing works include the discovery, based on a set of trajectories, the probability of stop goals that satisfy more than on rule.

ACKNOWLEDGMENTS

The authors thank to MODAP project, CNR (Italy) CNPq, FAPESC (CP 005/2009) and UFSC that partially provided support for this research. Thanks also to Wageninghen for the trajectory data.

6. REFERENCES


