

Original software publication

Stop@: A framework for scalable and noise-resistant stop-move segmentation of large datasets of trajectories in outdoor and indoor spaces

Fatima Hachem^{*}, Maria Luisa Damiani

University of Milan, Via Giovanni Celoria 18, Milan, Italy

ARTICLE INFO

Keywords:

Human and animal mobility
Trajectories
Stop-move detection

ABSTRACT

Capturing the mobility behavior of moving entities from their traces is a prominent theme in mobility data science. *Stop@* supports behavior analysis by providing a generic framework for the mining of stop-move patterns in spatial trajectories across animal and human mobility scenarios. The framework is built around a stop detection method, successfully used in diverse applications in animal ecology. The method has been recently validated against accurate ground truth stops collected in a museum, proving to be effective and robust, also for the study of human mobility. *Stop@* provides a rich set of functionalities to facilitate the stop-move analysis, including the parallel processing of large datasets of trajectories collected outdoor and indoor.

Code metadata

Current code version

Permanent link to code/repository used for this code version

Permanent link to Reproducible Capsule

Legal Code License

Code versioning system used

Software code languages, tools, and services used

Compilation requirements, operating environments & dependencies

If available Link to developer documentation/manual

Support email for questions

v1.0.3

<https://github.com/ElsevierSoftwareX/SOFTX-D-24-00156>

GNU General Public License v3.0 (GPL)

none

Python

Pandas, json, multiprocessing, statistics, numpy, utm, geopandas, contextily

fatme.hachem@unimi.it

1. Motivation and significance

Capturing the behavior of mobile entities from the history of their movements is key in a variety of fields such as marketing, computational sociology, urban mobility planning or animal ecology. A major challenge posed by these scenarios, typically targeting large populations of moving entities, is how to automate behavior analysis through the identification of patterns in datasets of *trajectories*. *Stop-move* is a prominent example of such kind of pattern, where a *stop* is an interval-based event, specifying a period of permanence in some region, and a *move* is the transition from one stop to another. Globally, stops and moves constitute a trajectory *segmentation*. Stops are content-rich, for example, can represent the staying of animals in their home range and the visits to museum exhibits. Yet, stops are complex to detect, especially in trajectories affected by some form of noise, including the noise deliberately induced by entities, for example, through an

extemporaneous movement. A number of stop detection methods based on different paradigms exist in the literature [1]. However, none of them have been systematically validated against a ground truth, thus there is no quantitative evidence of how good these methods are. Moreover, the methods are applied at the level of a single trajectory, while the analysis often regards large datasets of trajectories and thus is computationally demanding. Further, the question of how general these methods are and how well they fit in different application scenarios, is overlooked.

In this article, we present the *Stop@* framework [2] for the flexible and scalable detection of stop-move patterns in trajectory datasets, centered on the *SeqScan* algorithm. Different from the stop detection methods included in recent mobility software libraries, such as *MovingPandas* [3] and *Scikit-mobility* [4], which employ the distance-based paradigm, *SeqScan* leverages point density. Initially, it was designed for

^{*} Corresponding author.

E-mail addresses: fatme.hachem@unimi.it (Fatima Hachem), maria.damiani@unimi.it (Maria Luisa Damiani).

<https://doi.org/10.1016/j.softx.2024.101815>

Received 31 March 2024; Received in revised form 28 June 2024; Accepted 30 June 2024

Available online 17 July 2024

2352-7110/© 2024 Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the analysis of the migratory behavior of a group of animals wearing a GPS-equipped collar (i.e., roe deer, [5]), and has been progressively generalized [6] and applied to the study of both animal and human mobility [7–10]. In this sense, SeqScan is unique in bridging animal and human mobility analysis. Moreover, a methodology for the validation against ground truth stops has been put in place in a museum where visitors have been tracked with an accuracy of few decimeters using Ultra-wideband system (UWB) [9]. Compared to popular state-of-the-art methods, such as SPD [11], SeqScan proves to be highly efficient, and less sensitive to parameter variations.

The goal of Stop@ is to support stop-move analysis over large datasets of trajectories. It provides a rich set of functionalities, including pre and post-processing of stops, and parallel processing of large datasets of trajectories collected outdoor and indoor. The architecture of Stop@ is presented in Section 2, while Section 3 illustrates two examples. The final Sections discuss the impact of Stop@, reporting final considerations.

2. Software description

The software is designed to play a role in the mobility data analysis by extracting stops from trajectories. Before delving deeper, it is essential to establish the definitions of trajectory and stop, both conceptually and practically in terms of how they are represented in the software's input and output.

Trajectory. A trajectory T is a sequence of timestamped points, portraying the movement of an entity over time.

$$T = [p_1, p_2, \dots, p_n], p_i = (loc_i, t_i)$$

Where loc_i denotes the position at time t_i , expressed by either geographic or planar coordinates.

Practically, a trajectory is the input of our software, it is represented as a CSV file containing at least the fields related to the location and time. Additionally, when dealing with multiple trajectories within a single file, an extra field representing the trajectory's id becomes necessary for proper identification.

Stop. A stop is a segment, or subsequence within the trajectory, where the moving object stays for at least some defined time.

$$s \subseteq T, s = [p_i, p_{i+1}, \dots, p_j]$$

This subsequence of points forms a density-based cluster, potentially encompassing periods of absence from the cluster.

Practically, a stop is represented in 2 modes, the first is detailed, and the second is summarized:

1. A sequence of points classified as: *stop* or *excursion*, along with the corresponding stop identifier. A point is labeled *stop* when it belongs to the cluster, and *excursion* when the point temporally falls in the stop time interval but is spatially distant from the corresponding cluster, indicating a period of absence.
2. The stop identifier with a temporal interval denoting its start and end timestamps.

SeqScan. SeqScan is a cluster-based segmentation technique introduced in [5] and has undergone a thorough evaluation in [6]. It splits a trajectory in a series of segments representing the *stops* and *moves*: stops are temporally non-overlapping. Any two consecutive stops may be separated by a *move* segment indicating the transition. SeqScan requires three parameters: ϵ and n , the spatial density related parameters (similarly to DBSCAN [12]), and δ , which denotes the minimum required presence within a cluster to qualify as a *stop*. We need to highlight that the presence and duration are two distinct notions: Presence refers to the effective time the entity spends within the location of the stops, excluding excursion periods. Duration, on the other hand, encompasses the entire time period from the start to the end of the stop. Fig. 1 illustrates an example of a SeqScan stop, and shows the difference between presence and duration.

2.1. Software architecture and functionalities

The core structure of Stop@ is composed of three blocks (Fig. 2):

- **Pre-segmentation:** The purpose is to provide insights into the input trajectories. This involves statistical summaries about the trajectory's: (i) point count, (ii) total duration, (iii) minimum, maximum, median, average, and standard deviation of step length and duration. Step length refers to the distance between consecutive points, while step duration denotes the temporal gap between them. Additionally, a 3D trajectory plotter is employed to visualize the trajectory in spatial-temporal dimensions. These pre-segmentation tools help in better understanding the movement, and in determining the parameters ranges to be used for the SeqScan algorithm in the segmentation phase.
- **Segmentation:** This is the core component of the software, where the segmentation method (SeqScan) is applied to the trajectories. Technically, the output consists of two components: point classification and symbolic trajectories. Consequently, two CSV files are generated. The first representing the point classification maintains a structure similar to the input file, but it includes four additional fields in order to classify each point into one of these cases: (i) belonging to a stop, more specifically to a SeqScan cluster, along with the corresponding stop ID, (ii) being an excursion point, along with the corresponding stop ID, (iii) belonging to a *move* segment, indicating a transition from one stop to another, (iv) it is a noise point. Table 1 explains in details the role and the possible components for each field. It is worth mentioning that the selection of these fields was made with future developers' usage in mind. While both *cluster* and *class* indicate whether the point belongs to a stop or not, *cluster* holds numerical values (the stop's ID if the point is clustered, or -1 otherwise), whereas *class* contains textual values (STOP_i, where i represents the stop's ID, or MOVE otherwise). The second file represents the symbolic trajectory, consisting of a sequence of the extracted stops. Each stop is represented by its ID, start and end time and its centroid coordinates. Examples of the input and output files will be provided in the next section.
- **Post-segmentation:** This step aims to provide further information by statistically describing the resultant symbolic trajectories. These statistics include the number of stops along with the minimum, maximum, median, average, and standard deviation of their durations. Additionally, it includes the average presence/duration (P/D) of the stops. The P/D , introduced as a quality metric in [7], takes a value between (0, 1] to quantitatively measure the temporal compactness of the stops. A higher value indicates that the moving entity remains more consistently within the stop with fewer excursions. The output symbolic trajectory is also described by another functionality, the *Symbolic trajectory plotter*, which shows the stops in space. When working with GPS coordinates, the corresponding piece of the map is used in the plot. The stops are represented as points (disk symbols) using their centroid coordinates, where the size of each disk is proportional to the duration of the stop. Thus, the stop with the longest duration will have the largest size, and so on. The labels of the stops are also shown. This visualization allows the user to grasp information about the location of the stops on the map, their duration, and their ordering. In addition to the statistics related to the stops, the post-segmentation phase includes another functionality that generates statistics related to the *move* parts as well.

2.2. Two modes: single v.s bulk

In various research contexts or studies, the mobility dataset may be extensive, involving multiple moving objects and hence, multiple

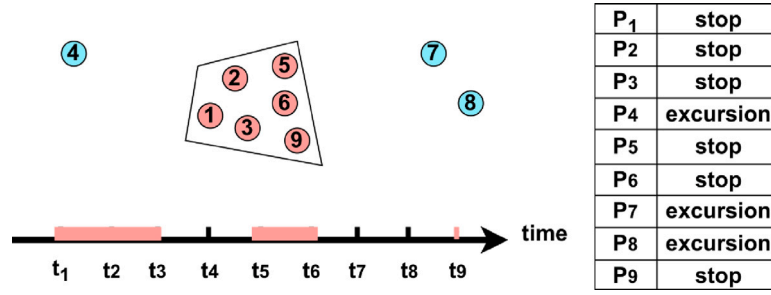


Fig. 1. Example of a SeqScan stop. The points are represented in red (stop point) and blue (excursion) dots simulating their spatial location. The presence and absence periods are depicted on the timeline. The symbolic representation of this stop is: $stop_i[t_1, t_9]$, while the point classification is shown in the table to the right. The duration D of this stop is $[t_1, t_9]$ (8 time units), while the presence P is $[t_1, t_3] \cup [t_5, t_6] \cup t_9$ (3 time units). $P/D = 3/8 = 0.375$.

Table 1
Output of SeqScan: meanings of point classification fields and their values.

cluster	class	type	details	meaning
i	STOP_i	cluster	cluster #i	The point is a stop point, belonging to cluster of id = i
-1	MOVE	excursion	of cluster #i	The point is an excursion during the STOP_i
-1	MOVE	transition	from cluster #i	The point belongs to the 'move' segment between STOP_i and STOP_{i+1}
-1	MOVE	noise	before/after clustering	The point does not belong to any cluster, and is situated before the start of the first cluster or after the last one's end

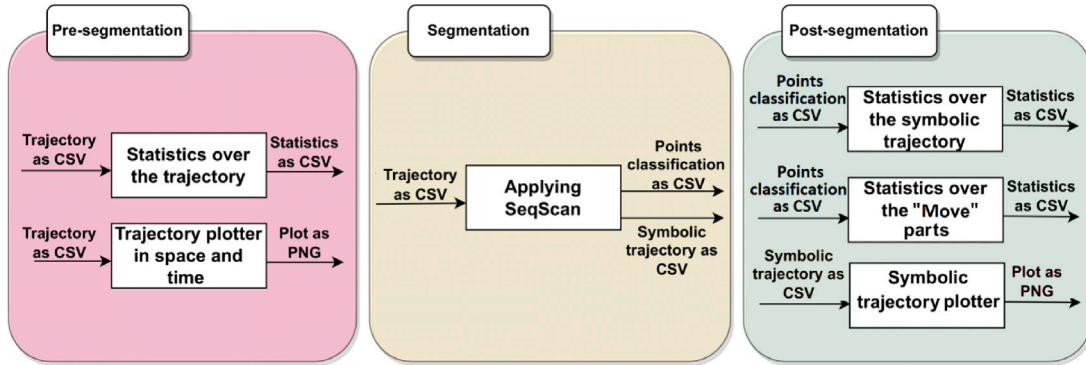


Fig. 2. The three primary components of the software and their respective functionalities. Input and output for each function are depicted by the arrows.

trajectories. Our software is designed to deal with multiple trajectories at once, a mode we refer to as the *bulk mode*.

In the single mode, each functionality expects a single CSV file as input, containing the data of a single trajectory. Correspondingly, output is a single CSV file as well.

When confronted with the need to process multiple trajectories, the software offers two options for representing the dataset:

1. A single CSV file containing data for multiple trajectories, distinguished by the trajectory identifier field.
2. A directory comprising multiple files, each corresponding to a single trajectory.

The output structure mirrors the input format. In the former case, the output consists of a single CSV file encompassing data for multiple trajectories. In the latter case, the output is a folder containing a file for each trajectory.

There is an exception concerning the trajectory plotter functionality in bulk mode: the output always comprises multiple PNG files, with each file representing a distinct trajectory. Detailed descriptions are available in Table 2.

2.3. Parallelism

When handling a bulk of trajectories to be segmented and analyzed, the execution time may become a concern. To tackle this issue, we have incorporated multiprocessing capabilities into our software for bulk operations. The multiprocessing utilizes the processing capabilities of multiple cores present in the computer's hardware simultaneously. This allows for tasks to be divided and executed in parallel across multiple cores, increasing overall processing efficiency and speed. The users have the flexibility to specify the number of cores they wish to engage in the software run. For instance, when segmenting N trajectories using M cores, with $N > M$ the software processes up to M trajectories simultaneously until all N trajectories are segmented. In more details, the SeqScan functionality in bulk mode:

- When presented with a single CSV file containing multiple trajectories, the system first parses the file, extracting a list of trajectories. Each trajectory is then assigned to an available core, where the SeqScan algorithm is executed. The result is appended to the single CSV file in output (Fig. 3).
- Alternatively, when dealing with a directory of trajectory files, the system iterates through each file, assigning them to distinct cores

Table 2

Two modes are available for every functionality: single and bulk.

(a) Pre-segmentation functionalities			
Functionality	Mode	Input	Output
Statistics over the trajectory	Single	1 CSV - 1 trajectory	1 CSV - 1 data row
	Bulk	1 CSV - N trajectories	1 CSV - N data rows
		1 Directory of N CSVs	1 Directory of N CSV s, each has 1 data row
Trajectory plotter	Single	1 CSV - one trajectory	1 PNG
	Bulk	1 CSV - N trajectories	N PNGs
		1 Directory of N CSVs	1 Directory of N PNGs
(b) Segmentation functionalities			
Functionality	Mode	Input	Output
SeqScan	Single	1 CSV - 1 trajectory	1 CSV-points classification, 1 CSV-symbolic trajectory
	Bulk	1 CSV - N trajectories	1 CSV - N trajectories' points classification 1 CSV - N symbolic trajectories
		1 Directory of N CSVs	1 Directory of N CSVs: each for 1 trajectory points classification 1 Directory of N CSVs: each for 1 symbolic trajectory
(c) Post-segmentation functionalities			
Functionality	Mode	Input	Output
statistics over the symbolic trajectory	Single	1 CSV -points classification	1 CSV-1 data row
	Bulk	1 CSV - N trajectories' points classification	1 CSV - N data rows
		1 Directory of N CSVs	1 Directory of N CSVs: each has 1 data row
statistics over the <i>Move</i> parts	Single	1 CSV -points classification	1 CSV-1 data row
	Bulk	1 CSV - N trajectories' points classification	1 CSV - N data rows
		1 Directory of N CSVs	1 Directory of N CSVs: each has 1 data row
Symbolic trajectory plotter	Single	1 CSV - one trajectory	1 PNG
	Bulk	1 CSV - N trajectories	N PNGs
		1 Directory of N CSVs	1 Directory of N PNGs

for processing. Each trajectory is segmented by SeqScan, and the results are exported to separate files within the output directory (Fig. 4).

Same procedures are also applied on all other functionalities.

3. Illustrative examples

In this section, two examples are presented. The first example involves a dataset capturing the movement of individuals in outdoor settings, focusing particularly on the city of Beijing, China. This dataset is accessible through the Microsoft GeoLife website [13], offering GPS trajectories represented by geographical coordinates to illustrate outdoor mobility patterns.

The second example involves a dataset of moving individuals within the ATC shopping center in Osaka, Japan. The dataset utilized in the examples is available online via the website [14], with detailed information on data collection provided in [15]. It depicts indoor mobility, with coordinate data provided in Cartesian space and measured in millimeters.

All datasets used in this paper are available online in the project directory, and comprehensive instructions for calling any functionality in any scenario are provided in detail within the README document.

3.1. Outdoor mobility data — GeoLife dataset

The GPS trajectories of this dataset are represented by a sequence of timestamped points. For the user convenience we provide a CSV file containing the trajectory of user 005 on the day 24/10/2008 (20081024041230). This trajectory contains 4298 points and spans a duration of ~ 11 h 46 min. A snippet of the file is presented in Fig. 5a. The config file needs to be updated specifying that the coordinates are geographical, setting the columns names accordingly and determining *minutes* as the time unit (Fig. 5b). This time unit is taken into consideration when providing the δ parameters for the SeqScan algorithm, and in the duration-related statistics. The trajectory plot in space and time is shown in Fig. 5c, it reveals potential stops during the person's movement. The statistics over the trajectory (Fig. 5d) indicates irregular sampling with possible missing points, as step durations vary from 0 to 250 min, with a mean and standard deviation of 0.16 min (10 s) and 3.8 min, respectively. The average step length is 8 m. These information are useful to select the parameters for SeqScan. We run the technique with $\epsilon = 8$ m, $n = 30$ points and $\delta = 15$ min. Two stops are extracted and represented by the symbolic trajectory (Fig. 5f), with point classification details illustrated in Fig. 5e. The latter demonstrates an example where points are classified as transitions from the first stop to the second, followed by clustered points belonging to the second stop, with excursions also observed during the second stop.

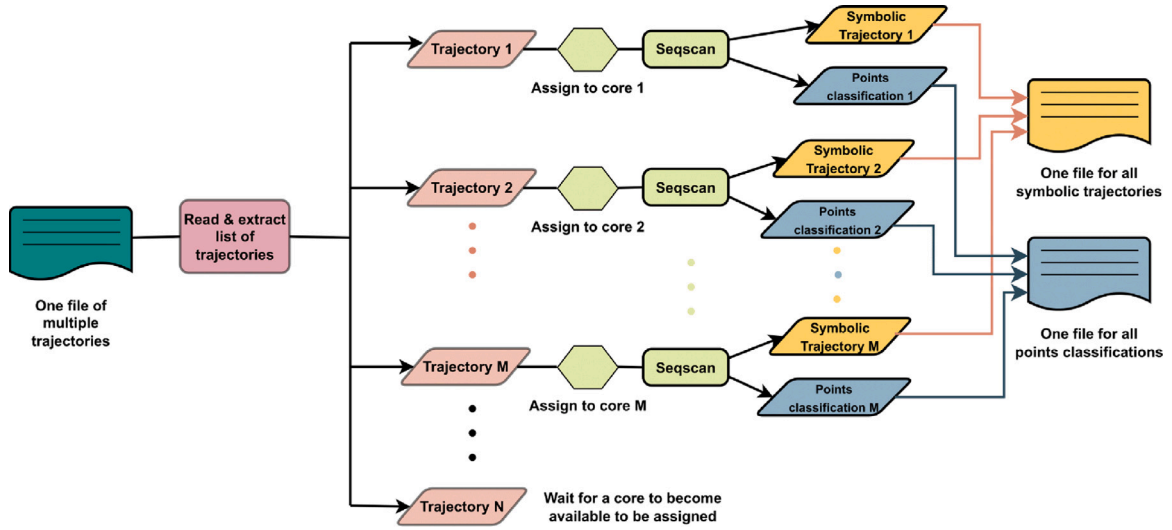


Fig. 3. Process flow in bulk mode with input as single file of multiple trajectories.

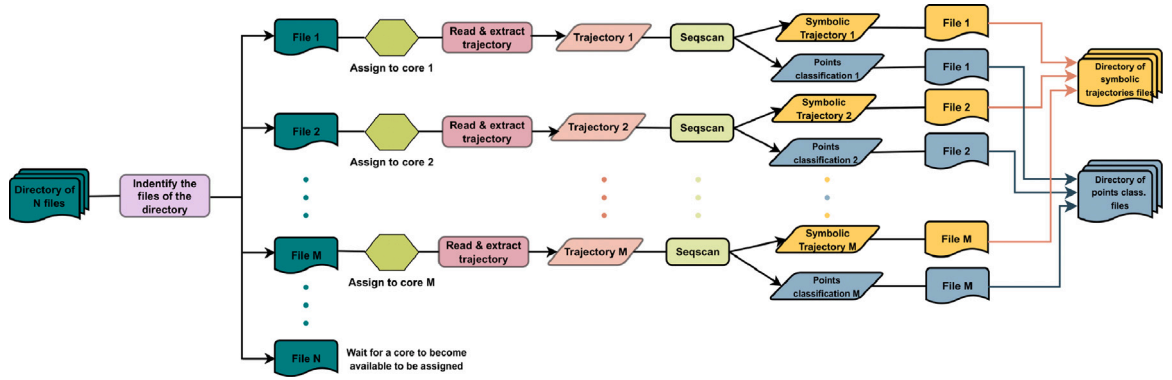


Fig. 4. Process flow in bulk mode with input as directory of multiple trajectories.

The occurrence of excursions affected the P/D ratio, which averaged 0.57 as indicated by post-segmentation statistics (Fig. 5g). This frequent occurrence of excursions is also confirmed in the statistics related to the *Move* parts, where the total number of move segments is 57 (Fig. 5h).

3.2. Indoor mobility data — ATC shopping center

For user convenience and quick reproducibility, a small subset of the original data is made available online in our code directory: in total, 7 trajectories are used. They are represented in 2 ways, in order to cover all the possible scenarios of single/bulk modes with/without parallelism: as a directory of 7 separate files, each containing a single trajectory, and as a unique file containing all 7 trajectories. Due to space constraints, we will focus on discussing bulk mode using a single CSV file. For detailed step-by-step instructions and examples of bulk mode with a directory of files, please refer to the code directory online. We execute SeqScan over the list of trajectories in the input file, using the parameters ($\epsilon = 1000$ mm, $n = 5$ points and $\delta = 10$ s). This yields 2 CSV files in output: one for the points classification and the other for the symbolic trajectories (Fig. 6a,b). Upon conducting statistical analysis of the symbolic trajectories (Fig. 6c), we observe variations in the number of stops per trajectory, ranging from 1 to 4 stops. Additionally, the average duration varies across trajectories, ranging from 10.5 s to 628 s. All trajectories exhibit a P/D ratio of 1, indicating that individuals did not make any excursions while within the stops.

4. Impact

The Stop@ framework will benefit the broad and interdisciplinary community working on mobility data science. In particular, it will facilitate the study of stop-based behaviors in diverse scenarios related to animal and human mobility, e.g., the study of migratory and exploratory behavior of animals, the study of museum visitors, the shopping behavior of customers.

Also, this pattern serves as a fundamental building block for more complex mobility patterns, such as periodic behaviors e.g. [8] and collective patterns. The availability of stop representation at multiple levels of details enables both in-depth investigations and a more compact representation. While the distributed processing option enables the efficient segmentation and analysis of large datasets of trajectories.

5. Conclusions

Stop@ is a Python framework offering tools for analysis, plotting, and segmenting trajectories to extract valuable insights. Designed for user convenience, it offers possibility to work in single or bulk modes, flexibility in handling indoor or outdoor data, and customizable configuration options. Additionally, its support for parallel processing enhances efficiency. With straightforward function calls and the comprehensive examples provided, users can quickly leverage its capabilities. Future enhancements may include additional pre and post-segmentation tools to further enhance usability and functionality.

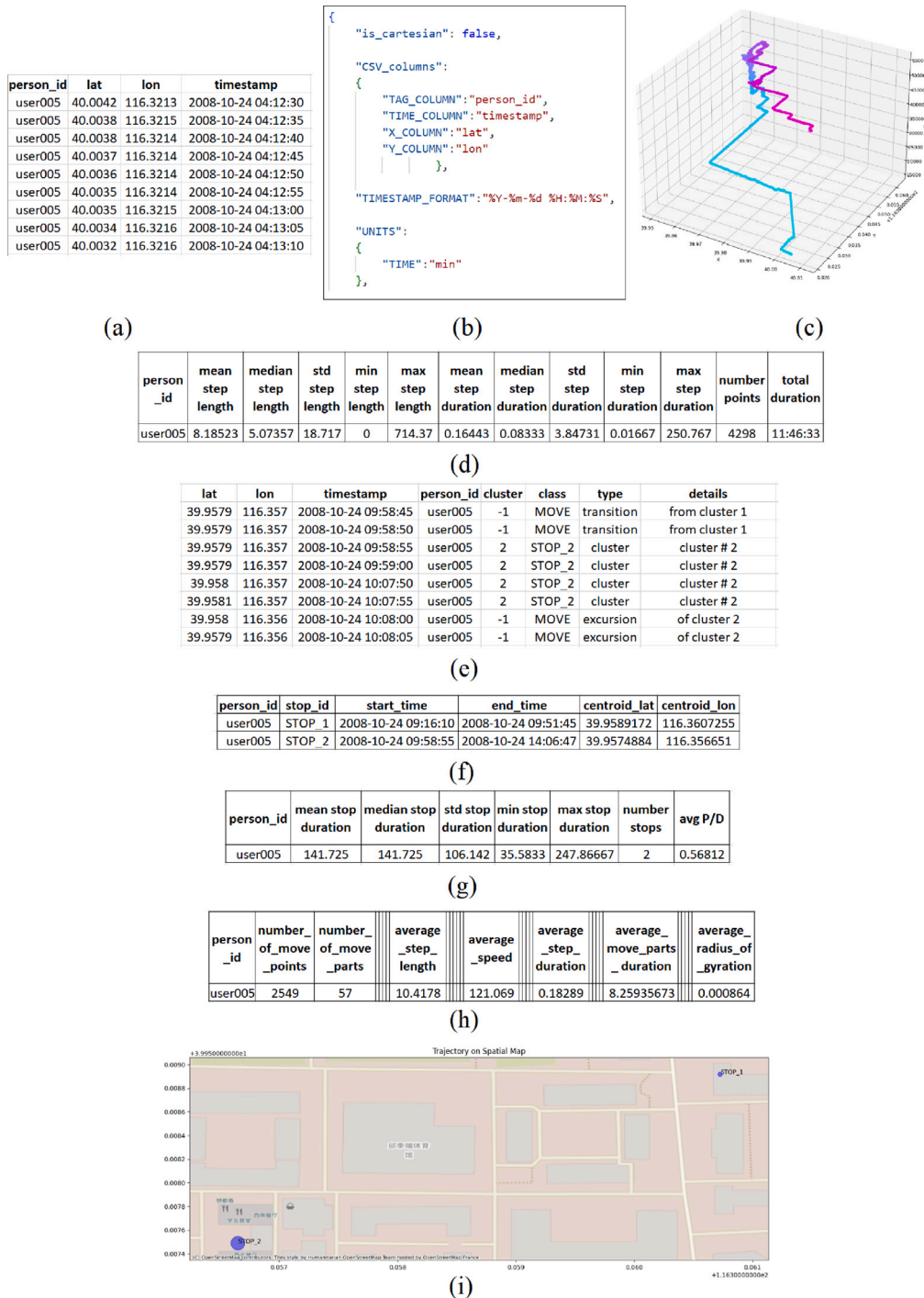


Fig. 5. GeoLife example: (a) part of the input trajectory, (b) The config file modified accordingly, (c) The trajectory plot, (d) the statistics of the trajectory, step length in meters and duration in minutes, (e) SeqScan output: part of the points classifications file, (f) SeqScan output: symbolic trajectory, (g) Statistics over the symbolic trajectory, duration in minutes, (h) Some pieces of the statistics over the *Move* parts, (i) The symbolic trajectory plot.

avg_x	avg_y	timestamp_sec	person_id	cluster	class	type	details
-10350	-8564	2013-01-09 03:35:19	10340900	1	STOP_1	cluster	cluster # 1
-10311	-8635.5	2013-01-09 03:35:20	10340900	1	STOP_1	cluster	cluster # 1
-10284	-8632.1	2013-01-09 03:35:21	10340900	1	STOP_1	cluster	cluster # 1

avg_x	avg_y	timestamp_sec	person_id	cluster	class	type	details
-215.43	3901.21	2013-01-09 03:01:33	10005800	1	STOP_1	cluster	cluster # 1
-671.64	4312.64	2013-01-09 03:01:34	10005800	-1	MOVE	transition	from cluster 1
-887.07	4830.81	2013-01-09 03:01:35	10005800	-1	MOVE	transition	from cluster 1
-1203.1	5191.54	2013-01-09 03:01:36	10005800	-1	MOVE	transition	from cluster 1
-2072	5283.25	2013-01-09 03:01:37	10005800	-1	MOVE	transition	from cluster 1
-2652.6	5673.32	2013-01-09 03:01:38	10005800	-1	MOVE	transition	from cluster 1
-3187.4	5991.89	2013-01-09 03:01:39	10005800	2	STOP_2	cluster	cluster # 2

person_id	stop_id	start_time	end_time	centroid_x	centroid_y
10340900	STOP_1	2013-01-09 03:35:03	2013-01-09 03:36:25	-11027.664	-9040.71
10340900	STOP_2	2013-01-09 03:36:27	2013-01-09 03:36:47	-12565.68	-5355.8299
10340900	STOP_3	2013-01-09 03:37:12	2013-01-09 03:37:40	-3669.9392	8409.71295
10340900	STOP_4	2013-01-09 03:37:48	2013-01-09 03:38:04	-11875.251	7586.81701
10383900	STOP_1	2013-01-09 03:38:49	2013-01-09 03:38:59	-6734.8568	5096.03821
10383900	STOP_2	2013-01-09 03:39:00	2013-01-09 03:39:11	-4532.9243	2185.49101
10394301	STOP_1	2013-01-09 03:39:43	2013-01-09 03:40:07	17105.925	-13534.428
10394301	STOP_2	2013-01-09 03:40:08	2013-01-09 03:40:20	18092.8221	-10389.282
10410100	STOP_1	2013-01-09 03:41:07	2013-01-09 03:41:31	-33228.952	-55.46531
10410100	STOP_2	2013-01-09 03:41:33	2013-01-09 03:41:51	-35760.409	-4128.851
10001800	STOP_1	2013-01-09 03:00:42	2013-01-09 03:00:53	-30812.768	-2703.1019
10005800	STOP_1	2013-01-09 03:01:15	2013-01-09 03:01:33	1500.92293	3169.97316
10005800	STOP_2	2013-01-09 03:01:39	2013-01-09 03:01:50	-4352.3992	6581.89001
10005800	STOP_3	2013-01-09 03:02:20	2013-01-09 03:03:16	-17893.244	-10334.448
10312900	STOP_1	2013-01-09 03:31:33	2013-01-09 03:31:45	9584.58508	-3979.8731

person_id	mean stop duration	median stop duration	std stop duration	min stop duration	max stop duration	number stops	avg P/D
10340900	36.5	24	26.622	16	82	4	1
10383900	10.5	10.5	0.5	10	11	2	1
10394301	18	18	6	12	24	2	1
10001800	11	11	0	11	11	1	1
10410100	21	21	3	18	24	2	1
10312900	12	12	0	12	12	1	1
10005800	628.333	18	868.1	11	1856	3	1

Fig. 6. ATC example: (a) Portions of the points classification file, generated by SeqScan, relative to all the trajectories, distinguished by their corresponding ID field. (b) The symbolic trajectories file, output of SeqScan, (c) the statistics over the 7 symbolic trajectories.

CRedit authorship contribution statement

Fatima Hachem: Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Conceptualization.
Maria Luisa Damiani: Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Original data are available online, links are cited. Our online code repository includes parts of these data as well (utilized in the examples)

Acknowledgments

This work was supported in part by project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Italian MUR. Neither the European Union nor the Italian MUR can be held responsible for them.

References

- [1] Damiani ML, Hachem F. Segmentation techniques for the summarization of individual mobility data. *WIREs Data Mining Knowl Discover* 2017;7(6):e1214. <http://dx.doi.org/10.1002/widm.1214>.
- [2] Hachem F, Damiani ML, Issa H. Stop@: A framework for scalable and noise-resistant stop-move trajectory segmentation (v1.0.3). 2024. <http://dx.doi.org/10.5281/zenodo.11662207>.

- [3] Graser A, Bell R, Boiko G, Parnell A, et al. MovingPandas. 2023. <http://dx.doi.org/10.5281/zenodo.10302333>.
- [4] Pappalardo L, Simini F, Barlacchi G, Pellungrini R. Scikit-mobility: A Python library for the analysis, generation and risk assessment of mobility data. 2019. [arXiv:1907.07062](http://arxiv.org/abs/1907.07062).
- [5] Damiani ML, Issa H, Cagnacci F. Extracting stay regions with uncertain boundaries from GPS trajectories: a case study in animal ecology. In: *Proc. of the 22nd ACM SIGSPATIAL international conference on advances in geographic information systems*. 2014. <http://dx.doi.org/10.1145/2666310.2666417>.
- [6] Damiani M, Hachem F, Issa H, Ranc N, Moorcroft P, Cagnacci F. Cluster-based trajectory segmentation with local noise. *Data Min Knowl Discov* 2018;32(4):1017–55. <http://dx.doi.org/10.1007/s10618-018-0561-2>.
- [7] Damiani M, Issa H, Fotino G, Heurich M, Cagnacci F. Introducing 'presence' and 'stationarity index' to study partial migration patterns: an application of a spatio-temporal clustering technique. *Int J Geogr Inf Sci* 2016;30(5):907–28. <http://dx.doi.org/10.1080/13658816.2015.1070267>.
- [8] Hachem F, Damiani M. Periodic stops discovery through density-based trajectory segmentation. In: *Proceedings of the 26th ACM SIGSPATIAL international conference on advances in geographic information systems*. Seattle, WA, USA: ACM; 2018. <http://dx.doi.org/10.1145/3274895.3274946>.
- [9] Hachem F, Vecchia D, Damiani ML, Picco GP. Fine-grained stop-move detection in UWB-based trajectories. In: *2022 IEEE international conference on pervasive computing and communications*. 2022. <http://dx.doi.org/10.1109/PerCom53586.2022.9762404>.
- [10] Lanszki J, Schally G, Heltai M, Ranc N. Golden jackal expansion in europe: First telemetry evidence of a natal dispersal. *Mammalian Biol* 2018;88:81–4. <http://dx.doi.org/10.1016/j.mambio.2017.11.011>.
- [11] Zheng Y, Zhang L, Ma Z, Xie X, Ma W-Y. Recommending friends and locations based on individual location history. *ACM Trans Web* 2011;5(1):1–44. <http://dx.doi.org/10.1145/1921591.1921596>.
- [12] Ester M, Kriegel H, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of the 2nd international conference on knowledge discovery and data mining*. Portland, OR, USA: AAAI Press; 1996.
- [13] GeoLife GPS Trajectories Dataset, URL <https://www.microsoft.com/en-us/download/details.aspx?id=52367>.
- [14] Datasets from the ATC shopping center, URL https://dil.atr.jp/crest2010_HRI/ATC_dataset/.
- [15] Bršćić D, Kanda T, Ikeda T, Miyashita T. Person tracking in large public spaces using 3-D range sensors. *IEEE Trans Hum-Mach Syst* 2013;43(6):522–34. <http://dx.doi.org/10.1109/THMS.2013.2283945>.