

Ying Zhang, Karsten Klein*, Oliver Deussen, Theodor Gutschlag, and Sabine Storandt

Robust visualization of trajectory data

<https://doi.org/10.1515/itit-2022-0036>

Received May 23, 2022; revised August 10, 2022; accepted August 12, 2022

Abstract: The analysis of movement trajectories plays a central role in many application areas, such as traffic management, sports analysis, and collective behavior research, where large and complex trajectory data sets are routinely collected these days. While automated analysis methods are available to extract characteristics of trajectories such as statistics on the geometry, movement patterns, and locations that might be associated with important events, human inspection is still required to interpret the results, derive parameters for the analysis, compare trajectories and patterns, and to further interpret the impact factors that influence trajectory shapes and their underlying movement processes. Every step in the acquisition and analysis pipeline might introduce artifacts or alterate trajectory features, which might bias the human interpretation or confound the automated analysis. Thus, visualization methods as well as the visualizations themselves need to take into account the corresponding factors in order to allow sound interpretation without adding or removing important trajectory features or putting a large strain on the analyst. In this paper, we provide an overview of the challenges arising in robust trajectory visualization tasks. We then discuss several methods that contribute to improved visualizations. In particular, we present practical algorithms for simplifying trajectory sets that take semantic and uncertainty information directly into account. Furthermore, we describe a complementary approach that allows to visualize the uncertainty along with the trajectories.

Keywords: Robustness, Quantification, Trajectory Analysis, Visualization

ACM CCS: Applied computing → Life and medical sciences → Computational biology, Theory of compu-

***Corresponding author: Karsten Klein**, University of Konstanz, Department of Computer and Information Science, D-78457, Konstanz, Germany, e-mail: karsten.klein@uni-konstanz.de

Ying Zhang, Oliver Deussen, Theodor Gutschlag, Sabine Storandt, University of Konstanz, Department of Computer and Information Science, D-78457, Konstanz, Germany, e-mails: ying.zhang@uni-konstanz.de, oliver.deussen@uni-konstanz.de, theodor.gutschlag@uni-konstanz.de, sabine.storandt@uni-konstanz.de

tation → Randomness, geometry and discrete structures → Computational geometry

1 Introduction

Improvements in sensor technology and satellite imagery over the last years facilitate the collection of large amounts of movement data, including vehicle traffic and animal or human movement [8, 38, 36, 37]. Movement data can be either collected from sensor tags attached to a moving object, e. g. GPS sensors and inertial measurements units (IMUs), or by processing of video or imagery data [3, 4]. The underlying backbone of the data, which also is a major factor in the analysis process, is usually a series of locations, e. g. derived from GPS fixes delivered by sensor tags or by tracking humans or animals in video footage. The latter is in particular of interest in confined spaces such as building infrastructure, cages or aviaries, where movement is restricted to a small area and can be easily captured with stationary cameras. These locations constitute a movement trajectory, which will usually be a simplified model of the original movement. In a wide range of application areas, such as urban planning, disaster management, engineering, sociology, and animal ecology, the collected data is used to serve important purposes such as the analysis of disease spreading, animal behavior, design requirements, for the development of corresponding models, and for decision making. The corresponding analysis pipelines often include automated analysis methods, e. g. based on geometric measures of the trajectory and on algorithms, but also visual exploration or a combination of the two. Visual exploration is particularly important when the interpretation of the movement has to take into account the context of the environment in which the movement occurred [11, 10], and when there are potential influence factors which cannot be easily assessed, quantified, and integrated into an automated approach. This is often the case in behavior analysis, e. g. for animal decision making, where impact factors such as topography, food supply, and social interaction can play a crucial role. Several large databases collect movement data sets and make them available to interested analysts and the public, e. g. the Movebank [9] animal movement database, with strongly increasing number and volumes of data sets in recent years.

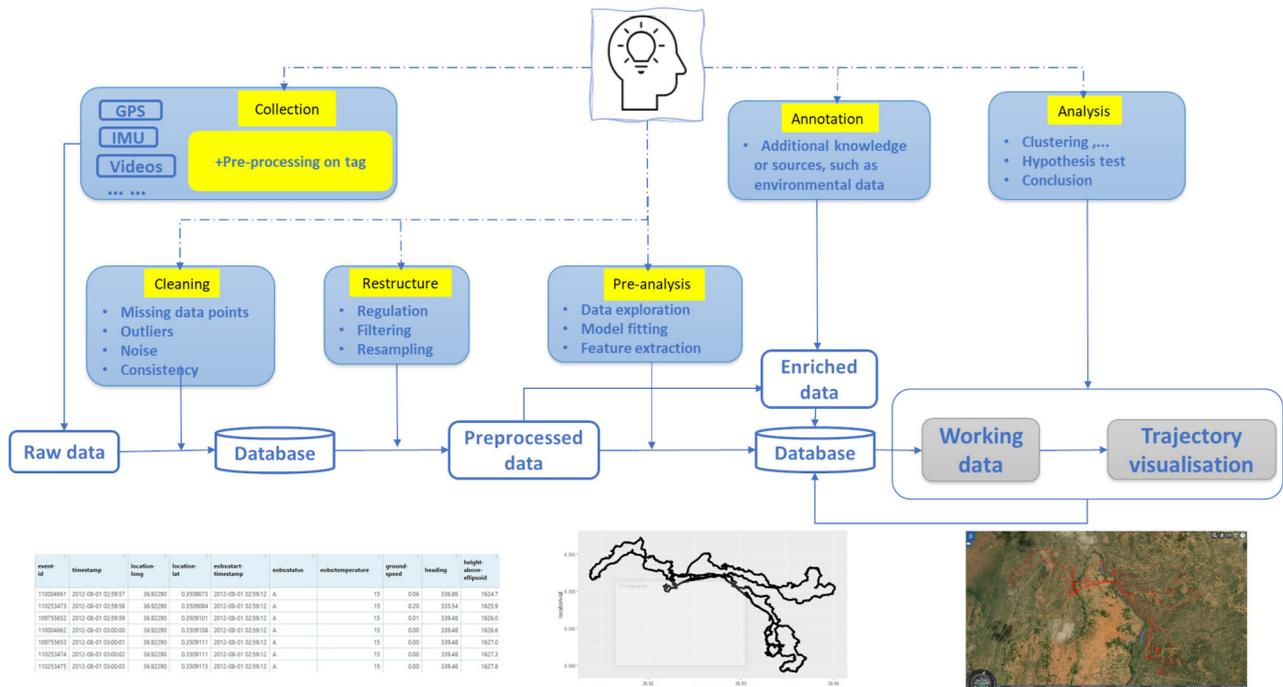


Figure 1: Simplified data collection and analysis pipeline – all components might contribute to loss or alteration of information and features required for the analysis. The analyst can be involved in all parts of the pipeline processes (indicated by the dashed arrows), e. g. by performing analysis tasks or data cleaning. The feedback from the different processing steps might help an analyst to build up deeper insight into the nature of the data, but there might be also different roles distinguished for human intervention, such as collector, curator, and analyst. The three pictures at the bottom, from left to right, show a snippet of typical input data, baboon movement (taken from [28, 27]), a corresponding simple data plot using R, and a trajectory visualization on a map created using the Teamwise tool [10] (based on the cesium platform and using Bing maps aerial imagery, © Microsoft Corporation).

Movement data from measured movement, and subsequent derived data, is often subject to quality issues regarding biases, noise, incompleteness, imprecision, and inconsistency, commonly associated with data veracity [5], and leading to uncertainty for analysis and visualization [1, 7], i. e. missing knowledge or imprecision and inaccuracy with respect to the original data. This uncertainty can stem from multiple causes, such as:

- Missing data, e. g. due to battery, storage, or transmission issues
- Inaccurate/imprecise measurements (high error rate), e. g. by low sensor or image processing accuracy
- Low spatial or temporal resolution
- Cleaning and preprocessing of the data, e. g. on the sensor tag
- Filtering and processing through automated methods, as well as their parameterization

These causes can occur at basically any step of the collection and analysis pipeline, see Figure 1. For some of the arising uncertainty, known limitations of the movement allow to assume certain restrictions, and to poten-

tially discard improper data points, e. g. based on a maximum speed or the straight flight direction of large scale bird migration over the sea. Still, a significant amount of uncertainty will remain and might affect most of the data aspects used for analysis. This needs to be taken into account in further analysis processing and has to be reflected in the visual representation in order to avoid misinterpretation by a human analyst. While the creation of effective data visualizations already requires careful considerations without data quality issues, uncertainty adds another challenge. A visualization might be properly representing the facts in the measured data, and nothing else, a property termed expressiveness [2], but cannot properly represent exactly the facts in the underlying ground truth, as it is not fully captured by the measurement. Trajectory visualization approaches therefore need to exhibit a robustness with respect to the data quality in order to support proper interpretation by the analyst.

An important step in the processing is the simplification of the preprocessed data, e. g. to make further analysis and visual representation more scalable. This simplification might not only be used to reduce the volume of a data

set, but also can help to remove artifacts that were introduced in the collection or preprocessing steps. An important goal at this stage is to avoid the removal of characteristic patterns or features in the movement that are relevant for later interpretation. Thus, it is crucial to consider the robustness in the decision of removing data points. Section 2 gives an overview on exploratory trajectory visualization and robustness issues with an example use case. Sections 3 and 4 describe how uncertainty information can be taken into account to improve robustness during the processing of trajectories and the subsequent visualization of results.

2 Exploratory analysis of trajectory data

The sheer volume, scale, and complexity of trajectory data that is nowadays collected, and the associated uncertainty, create challenges for analysis and interpretation. Exploratory analysis is often an important step that is performed with the help of a rich toolbox of analysis methods and statistics, but also using visualization to allow the analyst to get an overview on the structure and compare the trajectory with known classes. Such a visualization can help to get a first impression on the data at hand, using the analyst's knowledge to obtain an initial assessment of, for example, animal behavior, human mobility, or commuting patterns. In many application cases, additional annotations need to be integrated into the visualization, for example environmental conditions or features of the surrounding infrastructure. Often, trajectories are not investigated separately, but combined into a set or a graph, where common locations across trajectories form the nodes. Such a combination of multiple trajectories, e. g. to investigate the interplay or to create a summary such as a representative trajectory by clustering [25], can add further quality issues into the analysis process. Similarly, further associated data structures can be created, e. g. networks that model relations between the moving entities, such as social interaction or similar movement characteristics. A typical example in animal behavior analysis is leader-follower analysis, where the temporal correlation of movement patterns across individuals is used to derive roles within a group of animals [31]. A further example are dynamic networks that are derived by modeling edges between subjects based on criteria such as distance (e. g. a threshold under which a disease can be transmitted) or similar movement patterns (to detect learning or roles in an interaction) [26].

2.1 Acquisition, processing, and analysis tasks

The raw data that is collected usually either is provided as a series of GPS fixes, i. e. composed of longitude and latitude values, and sometimes altitude values, or as x , y , z coordinates, e. g. from an inertial measurement unit (IMU) or video analysis in a local or global coordinate system. Data sets can be collected for small time periods, such as several minutes to analyze behavior in a controlled setting, or over large periods of time, up to several years, e. g. to analyze development or migration and to compare seasons and impact of environmental changes. Depending on the analysis goal and the available technology, the temporal resolution can range from a few data points per day to 1 Hz or even higher resolutions. The resulting data sets thus might range from only a few hundred data points to several tens of gigabyte in size with tens of millions of data points. The latter case in practice still requires reduction before further analysis or trajectory visualization can be performed, as scalability to such volume is currently not supported throughout the full analysis pipeline. Still, quality issues can occur on all levels of scale. Software on sensor tags might already preprocess the data, e. g. for error correction, and in particular IMUs need careful calibration and corresponding data might suffer from drift over longer periods of time [33]. Further preprocessing is required to ensure sufficient data quality for different analysis purposes, ranging from simple summary calculations to statistical methods and models, and visualization is recommended for all preprocessing steps, e. g. to ensure plausibility [32]. In addition, due the parameterization of the methods involved, an interactive visual analytics loop can greatly support the analysis and the understanding of raw data and the impact of preprocessing [34]. Visual analytics combines automated analysis techniques with interactive visualizations for effective understanding, reasoning and decision making of complex data sets. [35]. For larger projects, the analysis (see Figure 1) is often performed in a circular fashion, going from analysis of an initial data set either back to adjust cleaning, restructuring or pre-analysis, or to the collection and analysis of further data for extension or comparison reasons, e. g. new seasons or events. The location series then can be interpreted as the movement trajectory of the entity. In order to derive information from these trajectories, e. g. on human mobility or animal behaviour [37, 36, 38], a variety of analyses can be performed using the trajectory data, such as home range estimation, step selection, path selection, path construction, clustering, segmentation, as well as corridor construction and estimation.

Those analyses are then used to help solve movement-related tasks [38] and to answer movement-related research questions [39]:

- Analysis of space use, e. g. home ranges of animals or activity spaces of humans.
- Behaviour identification via quantifying moving patterns, such as resting, travelling etc. [14], and often supported by a segmentation of the trajectory to identify points of changing behavior.
- Interaction and collective behavior in the case of multiple trajectories and entities.
- Investigation of the processes that underlie movement patterns, such as identifying influence factors of variation in movement rates by the environment, or quantification of how sex and reproductive status influence the duration of, and transition among, different behavioural modes.
- Prediction and modeling of movement.

2.2 Robustness requirements

The data quality issues can have a detrimental effect on the analysis for each of these tasks and questions: The extent to which space use is correctly identified, events and patterns can be detected, and further characterizations such as behaviour categories can be derived, is strongly affected by the coverage of movement features in the data that are characteristic for the targeted pattern or behavior. These could include turning points or spatio-temporal clusters, but also more complex features required in the categorization of movement or behavior. Detection methods and algorithms could be sensible to the absence or sparsity of such features, and thus might fail to detect events and patterns, or provide false results based on misinterpretation of the data. An analysis of the influence of location errors, as induced by different tracking technologies, on parameter estimation and subsequent biological pattern analysis for animal movement showed a significant decline in the ability to detect patterns [19]. In case of low resolution or missing data, interpolation methods can be used to fill the gap, but potentially even increase uncertainty or provide incorrect information. For example, dead reckoning can be used to calculate position information and fill up the gap by employing IMU data, but uncertainty in the resulting movement data can be further increased due to the imprecision and inaccuracy of measurement [19]. Sparsity of required movement features, and thus high uncertainty, might be caused simply by low temporal resolution in the acquisition, i. e. low sampling rate of sensors (or in a more general setting low resolution of the

data) [18]. Models are developed with the aim to reduce the uncertainty of trajectories [17], such as Fuzzy C-Means (FCM), which is an clustering algorithm with noise, and further variants [16]. A second cause can be later removal or adjustment of data points in the preprocessing step, e. g. by probability sampling. In fact, some established methods might require resolutions lower than what is routinely collected these days, thus making data reduction steps mandatory [13]. In special cases, such as human mobility analysis, the restrictions of the built environments infrastructure can largely decrease the amount of data required for location estimation with high probability [12]. In contrast, in less restricted cases, such as animal social network analysis in the wild, different sampling rates can change the perceived structure of a network [20], and the simplified, binary method to construct a social network can lead to wrong interpretations about the social structure and wrong inferences about the position of individuals within the network [21].

Statistic frameworks and models are widely used in animal behavior analysis, such as maximum-likelihood estimation (MLE), a method to estimate distribution parameters based on an observed sample, or Bayesian inference which is based on Bayes' theorem, where the probability for a hypothesis is updated when more evidence or information becomes available. The results of such methods based on the distributions and corresponding probability introduce uncertainty factors for the further analysis, yet those factors are usually not shown in the subsequent trajectory visualization. Adding these factors might help to increase analysis quality and trust. Section 4 gives examples on how visualizations can be enriched with uncertainty information to provide helpful input for the analysis.

Similar to the automated analysis, human interpretation of data might be misled by uncertainty in the data or the visualization. The uncertainty stemming from the causes listed in Section 1 needs to be taken into account in the further processing, visualization, and interpretation in order avoid wrong conclusions. Both automated analysis methods and visualisation metaphors can be sensitive to changes in the input data characteristics. For example, methods for event or pattern detection in movement time-series might either miss events or assign wrong categories, e. g. for behaviour classes, to time windows when the temporal resolution of measured locations is not sufficient or the locations are off from the real location by a certain amount.

2.3 Example use case

We demonstrate robustness issues in animal behavior research, where animal movement tracks are analyzed to derive characteristics of animal decision making, social interaction, and influence of environmental factors. To this end, we show example trajectories that despite their small scale already exhibit features that allow to highlight visual analysis issues.

Figure 2 shows two trajectory visualizations based on a goose migration data set (#12 from [30, 29]), one showing the full data set with a duration from end of January to August and a mean temporal resolution of around eighteen hours, and the other with a temporal resolution of one day. Even with this small change in resolution a major feature, a zig-zag movement pattern, is missing. Putting the movement into the context of the environment in which it happened is a common procedure for interpretation. Figure 3 shows the same trajectory projected on a globe visualization. It shows the zig-zag pattern happening at the Gulf of Riga, with a potentially important role for evidence of behavioral patterns. Such feature removals appear consistently for many different types of movement patterns when data reduction is performed, e.g. when following or circumflying natural resources such as rivers or mountains. Figure 4 shows three movement trajectories, where two are sampled from the original trajectory using the RadeHabitatLT package [6] with hourly and daily sampling resolutions. The red trajectory is generated using all data points in the file, the green with hourly sampling, and the blue trajectory using one data point each day. To generate the trajectory, the package requires a reference time data point, with the closest timepoint to the reference data point being picked daily or hourly sampling.

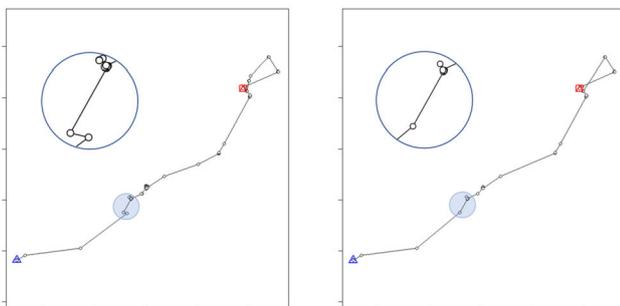


Figure 2: Goose migration trajectory. Left: using all data points from the data set. Right: using daily sampling rate. The zig-zag movement feature is removed (inside blue disc, see blue circle for a zoom-in).



Figure 3: Travel trajectory of a goose projected on a globe map using the Teamwise framework [10] (based on the cesium platform and using ESRI National geographic imagery, © ESRI). This visualization helps to take into account aspects of the surrounding environments which can help to interpret movement decisions, e.g. by showing topographic information.



Figure 4: Sampling and data reduction effects: The red trajectory represents the full data set, the green one represents the a reduced data set with hourly rate, the blue one represents the daily rate trajectory.

Analyzing the movement using the blue line between data points is based on the hypothesis that a location is not only directly accessible from the previous one, but that also the direct way is taken in this case. However, there might be mountains between two locations that hinder direct access by the investigated animal, or rivers, roads, or lakes that it follows. Further causes of deviations could be bad weather or interaction with other animals. As a result, the animal will probably take a detour to reach the next location. In this case, the lowest sampling however still resembles closely the global shape of the movement, allowing the analyst to investigate the overall migration route. It however omits several curves that might contain important environmental features for animal decision making, most prominently the one leading from Voronezh to the region around Kharkiv (bottom left). Figure 5 shows a baboon movement with trajectory visualizations based on two temporal resolutions. The hourly resolution strongly changes the characteristics of important features such as



Figure 5: Close-up of the baboon movement visualization from Figure 1, showing river and road structures that represent orientation or obstacles in animal movement. The red line shows the original temporal resolution, the green line hourly sampling.

river (dark green band from top left to bottom center) crossing points and orientation along roads (light orange band roughly following the river shape on the right hand side). The analysis based on time spend at locations, such as home range analysis or first passage time, without taking into account the environmental factors might draw an inaccurate conclusion based on the direct line trajectory [15].

3 Simplification of trajectory data

Simplification is an essential building block of many trajectory processing frameworks, as it allows to provide a cleaner view of the essential trajectory shape by reducing noise and clutter. Hence visual analytic tasks might be easier carried out on a simplified representation. Furthermore, simplification allows to reduce the data volume that needs to be stored and visualized, which is especially important when dealing with large data sets or interactive visualizations. In this section, we will discuss conventional and uncertainty-aware simplification of trajectories and provide experimental results on artificial and real-world data.

3.1 Polyline simplification

Applying trajectory simplification, one needs to be careful that the simplification process does not erase important patterns and that the overall movement shape is sufficiently preserved. Therefore, the simplification is typically governed by a distance measure $d(T, T')$ which determines how similar the input trajectory T and the simplified trajectory T' are. The user then can specify an error thresh-

old $\varepsilon > 0$, and the goal of the simplification process is to compute T' of smallest size, such that $d(T, T') \leq \varepsilon$. As the simplification is only concerned with the spatial aspects of the trajectories and not the temporal information, we will refer to trajectories as polylines in the remainder of the chapter. A polyline P is then simply a sequence of points p_1, \dots, p_n with $p_i \in \mathbb{R}^2$ and induced straight line segments between consecutive points. The size of a polyline equals the number of points n . Commonly used distance measures for polylines are the *Hausdorff* distance d_H and the *Fréchet* distance d_F . The latter is often deemed superior for movement analysis as it takes the order of the points along the polyline into account, while d_H ignores this aspect. An optimal polyline simplification under d_H can be computed in $\mathcal{O}(n^2)$ [23] and under d_F in $\mathcal{O}(n^2 \log n)$ [24].

3.2 Uncertainty-aware simplification

In contrast to previous approaches which assume precise knowledge of the point locations, Buchin et al. [22] recently introduced the problem of uncertain polyline simplification, where the input is not a sequence of points but a sequence of regions. The regions are assumed to contain the true location. The goal of the simplification process is to compute a polyline P' such that $d(P, P') \leq \varepsilon$ for all possible realizations of P given the input regions. We will focus here on the simple model in which all regions are disks. This complies with the fact that imprecision of GPS measurements can be expressed with an error radius r . As the error might differ depending on the absolute location, the received signal strength of the satellites and other aspects, the radius is not assumed to be homogeneous. Hence, the input is a sequence of disks $\mathcal{D} = D_1, \dots, D_n$, where each disk $D_i = (c_i, r_i)$ is a tuple of a center point c_i and an error radius $r_i \in \mathbb{R}^+$. A polyline $P = p_1, \dots, p_n$ is called a realization with respect to \mathcal{D} if $p_i \in D_i$. See Figure 6 for an illustration of this concept.

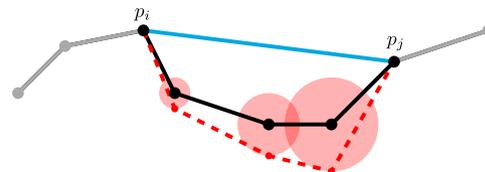


Figure 6: The blue line segment $\overline{p_i p_j}$ may only be part of the simplified polyline if it has a distance of at most ε to the subpolyline between p_i and p_j (marked black). Taking uncertainties into account (red disks), it is only valid if its distance towards any subpolyline realization within the given ordered disks is sufficiently small, including the one shown as dashed red line.

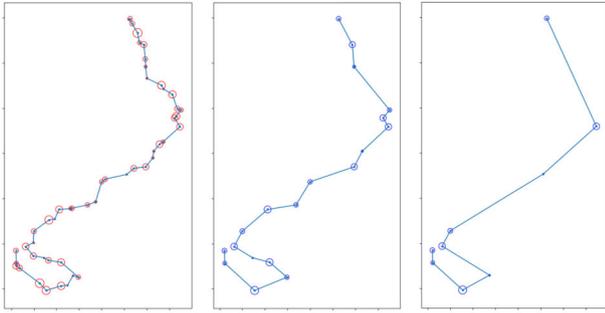


Figure 7: Left: Input sequence of 50 disks with an exemplary poly-line realization through their center points. Middle and right: Optimal simplifications for $\epsilon = 2$ and $\epsilon = 3$, respectively.

An example of a possible input disk sequence and the envisioned outcome is shown in Figure 7. The runtime to find the best simplified polyline in this setting is in $\mathcal{O}(n^3)$ for both d_H and d_F [22]. But so far, no practical feasibility study was conducted.

We first conducted an experimental study on generated disk sequences. Using artificial data allows to control aspects as the number of disks n , the maximum radius r_{max} , and the shape of the trajectory. Creating sequences with up to $n = 1000$ disks, and varying the value of ϵ , we observe that the running time in practice is quadratic in n and not cubic as predicted by the worst-case theoretical analysis. As a result, the running time of the algorithm is not much larger than that of the simplification algorithm that does not take uncertainty into account. In our experiments, the maximum increase in running time when considering disks was 25 %.

Based on the encouraging results of the feasibility study, we also conducted experiments on real-world data, particularly on the 18 trajectories from geese data set described in Section 2.3, which are depicted in Figure 8.

The data sets does not contain information about the uncertainty of the single locations. Hence we estimated error radii. We intentionally overestimated GPS uncertainty to make uncertainty-aware simplification – which needs



Figure 8: Migration routes of 18 geese.

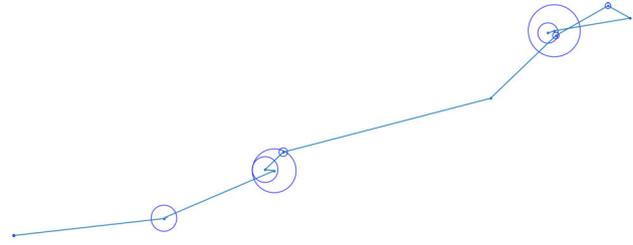


Figure 9: Trajectory 3 from Figure 8 simplified with $\epsilon = 2$. Circles indicate the uncertainty regions of the remaining points.

to consider all possible realization – more difficult. Nevertheless, the simplification capability is quite pronounced even for small values of ϵ . For $\epsilon = 1$, the average trajectory size was reduced to roughly 35%, for $\epsilon = 2$ to around 5%, allowing for a cleaner visualization while still maintaining the overall shape. For larger values of ϵ , only very few data points remain and the trajectory becomes too oversimplified for visual analysis. But for sensible choices of ϵ , we get the desired granularity that allows for concise but meaningful visualization, see Figure 9 for an example.

3.3 Context-aware bundle simplification

When studying animal or human movement data, one is typically not only interested in a single trajectory but a collection of trajectories from different individuals, such as from a flock of birds or travels along a road network. If now each trajectory is simplified independently of the others, we might erase relevant interactions between the individuals, for example, groups of animals meeting at a location. Furthermore, as discussed for the example of baboon movement (see Figure 5) environmental elements as rivers might guide or influence their behaviour. Again, if all trajectory data points close to the river would be simplified away, the movement context might be lost for the viewer. Figure 10 (left and middle) shows an example where the

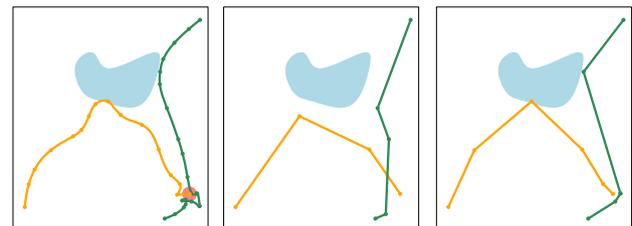


Figure 10: Left image: Two animal trajectories (orange and green), that both visit a waterhole (blue) and then come close to each other (red disk). Middle: Simplification without context consideration. Right: Context-aware simplification.

analysis of the trajectory data might be heavily distorted after the simplification.

In [41, 42], we proposed the polyline bundle simplification problem. Here, given a collection of trajectories, we want to find the smallest simplification for the trajectories governed by a distance measure and an error threshold $\varepsilon > 0$ as before, but with the additional constraint that points that occur in multiple trajectories have either to be kept in all of them or deleted from all of them. This consistency criterion unfortunately makes the problem NP-hard. But we showed that an efficient bicriteria approximation algorithm exists and also devised an efficient heuristic based on partitioning the trajectories into bundles that exhibit a tree structure.

Note that in real-world data, trajectories rarely share exactly the same location measurements, even if the respective humans or animals are in the same place. To capture close-by points we can hence again use an uncertainty radius and then impose the consistency criterion whenever two points from different trajectories are within the respective circle. Furthermore, we can also make the criterion more strict by demanding that such points have to be kept in the simplified output. In that way, we could also easily safe points close to important landmarks (as rivers or waterholes) from being simplified away, see Figure 10, right. This stricter criterion does not only help to ensure that relevant parts of the trajectories are kept but also makes the problem easier again from a computational perspective. Experiments conducted on large data sets demonstrate that even large bundles can be processed within a few seconds. Figure 11 shows the produced solu-



Figure 11: Context-aware simplification of car trajectories. Colors indicate tree structures that were identified for faster consistent simplification.

tion on a set of several thousand car trajectories with over a million data points.

4 Uncertainty-awareness for temporally varying trajectory data

As discussed in the last section, in many cases trajectories are not standing for themselves, but need to be considered in context with other trajectories. If trajectories share common points, we can also interpret these points as nodes of a joint graph.

This is sensible, for example, when considering trajectories that stem from driving from one place to another through a series of in-between cities. The cities then form the nodes of the graph. Individual travel results in temporal uncertainty coming from different driving times along trajectories. One natural goal in the visualization of travel networks then is to enrich the network visualization with a representation of that uncertainty, for which we describe an approach in the following.

The problem can be modeled as a probabilistic graph, in which for every edge (a part of a trajectory between two cities) a probability density function of a certain random variable exists, which in this case describes the distribution of the driving times for different times and days of the week.

In Fig. 12 the modeling is shown. Given the probabilistic graph with probability functions along the edges, we sample the edge functions and compute individual layouts for each of the graphs using an anchored force-directed

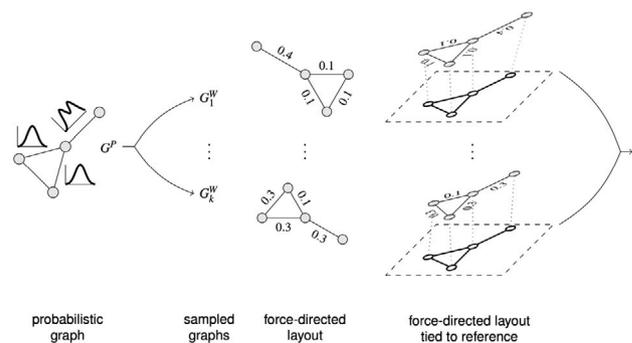


Figure 12: Modeling of a graph with uncertain trajectories. Uncertainty is represented by probability functions along the edges. To render it, the functions are sampled, layouts for the individual graphs are computed, later all graphs are combined to visualize the uncertainty-aware result [40].

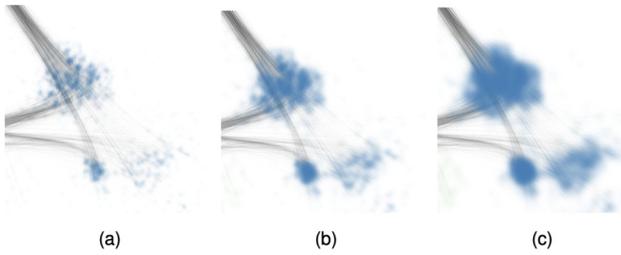


Figure 13: Rendering a combined graph with uncertainty using blending: for each point of the sampled graphs a partly transparent dot is created that is blended. Using different dot sizes, different appearances of the distribution can be achieved [40].

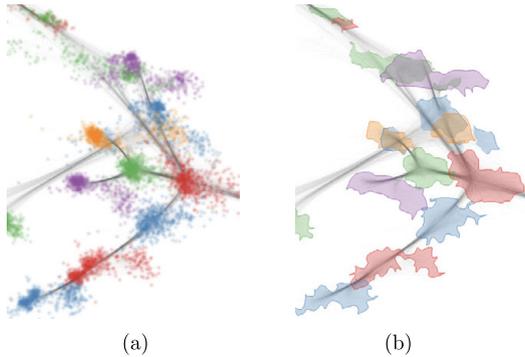


Figure 14: Rendering a combined graph with uncertainty using outlining: in (a) the sampling method is used, while in (b) all points that belong to one node of the probabilistic graph representation are combined and represented by an outline. The parameters of this outline can be used to model the general uncertainty [40].

layout [43]. Having obtained the overlaid graphs, the next challenge is to render them appropriately. In [40] we developed and evaluated different rendering methods in order to visualize the uncertainty in different situations and for different probability functions. In Fig. 13 we show results in which the points of the sampled probabilistic graphs are displayed using partly transparent discs. The result is combined using alpha blending. By using different dot sizes, different appearances of the distribution can be achieved. In Fig. 14(b), a different rendering method is shown. Here all the samples that belong to the same node in the probabilistic graph are combined and represented using an outline. The parameters of the outlines allow us to display different amounts of uncertainty.

Figure 15 shows the overlaid graphs for the uncertain driving times between major south-German cities. The nodes are represented using sampling. The shape of the points that represent one node in the probabilistic graph shows the viewer the variance of the probability function; e. g., when driving to Zurich, times will differ much more than when driving from Wuerzburg to Stuttgart.

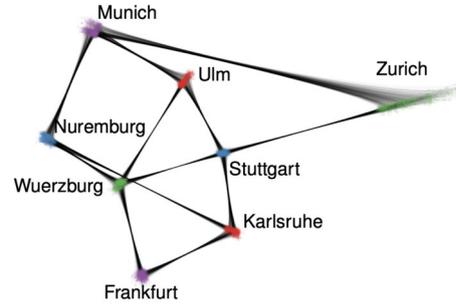


Figure 15: Rendering the full graph with uncertain driving times using a node placement that is geo-referenced.

5 Discussion and future work

We gave an overview on issues of trajectory visualization related to robustness with respect to features in the data relevant for analysis, e. g. by exploiting uncertainty in the trajectory data or subsequent analysis results. Two major aspects are how to visualize the uncertainty itself, and how to take it into account for analysis and the visualization methods. We presented methods for both tasks and showcased their applicability on animal and human movement data. However, as can be seen in Figure 1, robustness might need to be considered at every step of the data analysis pipeline. Therefore, additional methods may need to be integrated to cover all steps and, in particular, to further improve the basis for visual analysis of trajectory data.

Funding: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161 (projects A01, B06, D04).

References

1. Daniel Weiskopf. Uncertainty visualization: Concepts, methods, and applications in biological data visualization. *Frontiers in Bioinformatics*, 2, 2022.
2. Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5(2):110–141, 1986.
3. Matteo Zago, Matteo Luzzago, Tommaso Marangoni, Mariolino De Cecco, Marco Tarabini, and Manuela Galli. 3d tracking of human motion using visual skeletonization and stereoscopic vision. *Frontiers in Bioengineering and Biotechnology*, 8, 2020.
4. Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitkreutz, Philipp Zimmermann, Bastian Goldlücke, Tobias Schreck, Gennady Andrienko, Michael Grossniklaus, and Daniel A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE*

- Transactions on Visualization and Computer Graphics*, 24(1):13–22, 2018.
5. M Schroeck, R Shockley, J Smart, Dolores Romero Morales, and P Tufano. Analytics: the real-world use of big data: How innovative enterprises extract value from uncertain data, executive report. 2012.
 6. Clement Calenge, contributions from Stephane Dray Royer, and Manuela. *adehabitatLT: Analysis of Animal Movements*. 2020.
 7. Georges-Pierre Bonneau, Hans-Christian Hege, Chris R. Johnson, Manuel M. Oliveira, Kristin Potter, Penny Rheingans, and Thomas Schultz. *Overview and State-of-the-Art of Uncertainty Visualization*, pages 3–27. Springer London, London, 2014.
 8. F. Cagnacci, L. Boitani, R. A. Powell, and M. S. Boyce. Animal ecology meets gps-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1550):2157–2162, 2010.
 9. B. Kranstauber, A. Cameron, R. Weinzerl, T. Fountain, S. Tilak, M. Wikelski, and R. Kays. The movebank data model for animal tracking. *Environmental Modelling & Software*, 26(6):834–835, 2011.
 10. Karsten Klein, Michael Aichem, Ying Zhang, Stefan Erk, Björn Sommer, and Falk Schreiber. Teamwise: synchronised immersive environments for exploration and analysis of animal behaviour. *Journal of Visualization*, 24(4):845–859, 2021.
 11. Karsten Klein, Björn Sommer, Hieu Nim, Andrea Flack, Kamran Safi, Mate Nagy, Stefan Feyer, Ying Zhang, Kim Rehberg, Alexander Gluschkow, Michael Quetting, Wolfgang Fiedler, Martin Wikelski, and Falk Schreiber. Fly with the flock: immersive solutions for animal movement visualization and analytics. *J R Soc Interface*, 16(153):20180794, 2019.
 12. Piotr Sapiezynski, Arkadiusz Stopczynski, Radu Gatej, and Sune Lehmann. Tracking human mobility using wifi signals. *PLOS ONE*, 10(7):1–11, 07 2015.
 13. Jonathan R. Potts, Luca Börger, D. Michael Scantlebury, Nigel C. Bennett, Abdulaziz Alagaili, and Rory P. Wilson. Finding turning-points in ultra-high-resolution animal movement data. *Methods in Ecology and Evolution*, 9(10):2091–2101, 2018.
 14. Ashley Bennison, Stuart Bearhop, Thomas W. Bodey, Stephen C. Votier, W. James Grecian, Ewan D. Wakefield, Keith C. Hamer, and Mark Jessopp. Search and foraging behaviors from movement data: A comparison of methods. *Ecology and Evolution*, 8(1):13–24, 2018.
 15. Patrick Laube and Ross S. Purves. How fast is a cow? Cross-Scale Analysis of Movement Data. *Transactions in GIS*, 15(3):401–418, 2011.
 16. Nikos Pelekis, Ioannis Kopanakis, Evangelos E. Kotsifakos, Elias Frentzos, and Yannis Theodoridis. Clustering uncertain trajectories. *Knowledge and Information Systems*, 28(1):117–147, 2011.
 17. Guan Yuan, Penghui Sun, Jie Zhao, Daxing Li, and Canwei Wang. A review of moving object trajectory clustering algorithms. *Artificial Intelligence Review*, 47(1):123–144, 2017.
 18. Yu Zheng. Trajectory Data Mining. *ACM Transactions on Intelligent Systems and Technology*, 6(3):1–41, 2015.
 19. Corey J. A. Bradshaw, David W. Sims, and Graeme C. Hays. Measurement error causes scale-dependent threshold erosion of biological signals in animal movement data. *Ecological applications: a publication of the Ecological Society of America*, 17(2):628–638, 2007.
 20. David Lusseau, Hal Whitehead, and Shane Gero. Incorporating uncertainty into the study of animal social networks. *Animal Behaviour*, 75(5):1809–1815, 2008.
 21. D. P. Croft, R. James, A. J. W. Ward, M. S. Botham, D. Mawdsley, and J. Krause. Assortative interactions and social networks in fish. *Oecologia*, 143(2):211–219, 2005.
 22. Kevin Buchin, Maarten Löffler, Aleksandr Popov, and Marcel Roeloffzen. Uncertain Curve Simplification International Symposium on Mathematical Foundations of Computer Science (MFCS 2021).
 23. W. S. Chan, and F. Chin. Approximation of polygonal curves with minimum number of line segments or minimum error International Journal of Computational Geometry and Applications, 6(1):59-77, 1996.
 24. Sabine Storandt, and Johannes Zink. Polyline Simplification under the Local Fréchet Distance has Subcubic Complexity in 2D arXiv, 2022.
 25. Jana Seep and Jan Vahrenhold. K-means for semantically enriched trajectories. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Animal Movement Ecology and Human Mobility*, HANIMOB’21, page 38-47, New York, NY, USA, 2021. Association for Computing Machinery.
 26. Meggan E Craft. Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669):20140107, 2015.
 27. Margaret C. Crofoot, Roland Kays, and Martin Wikelski. Data from: Study “collective movement in wild baboons”, 2021.
 28. Roi Harel, J. Carter Loftus, and Margaret C. Crofoot. Locomotor compromises maintain group cohesion in baboon troops on the move. *Proceedings of the Royal Society B: Biological Sciences*, 288(1955):20210839, 2021.
 29. A. Kölzsch, GJDM Müskens, S. Moonen, H. Kruckenberg H, P. Glazov, M. Wikelski. Margaret C. Crofoot, Roland Kays, and Martin Wikelski. Data from: Longer days enable higher diurnal activity for migratory birds [greater white-fronted geese]. Movebank Data Repository, 2021. doi: 10.5441/001/1.254rd102.
 30. Ivan Pokrovsky, Andrea Kölzsch, Sherub Sherub, Wolfgang Fiedler, Peter Glazov, Olga Kulikova, Martin Wikelski, and Andrea Flack. Longer days enable higher diurnal activity for migratory birds. *Journal of Animal Ecology*, 90(9):2161–2171, 2021.
 31. Maté Nagy, Zsuzsa Akos, Dora Biro, and Tamás Vicsek. Hierarchical group dynamics in pigeon flocks. *Nature*, 464:890–3, 04 2010.
 32. Pratik Rajan Gupte, Christine E. Beardsworth, Orr Spiegel, Emmanuel Lourie, Sivan Toledo, Ran Nathan, and Allert I. Bijleveld. A guide to pre-processing high-throughput animal tracking data. *Journal of Animal Ecology*, 91(2):287–307, 2022.
 33. David Tedaldi, Alberto Pretto, and Emanuele Menegatti. A robust and easy to implement method for imu calibration without external equipments. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3042–3049, 2014.
 34. Aidan Slingsby and Emiel van Loon. Exploratory visual analysis for animal movement ecology. In *Computer Graphics Forum*, volume 35, pages 471–480. Wiley Online Library, 2016.
 35. Daniel Keim, Gennady Andrienko, Jean-Daniel Fekete, Carsten

Görg, Jörn Kohlhammer, and Guy Melançon. *Visual Analytics: Definition, Process, and Challenges*, pages 154–175. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

36. Gennady Andrienko, Natalia Andrienko, Wei Chen, Ross Maciejewski, and Ye Zhao. Visual analytics of mobility and transportation: State of the art and further research directions. *IEEE Transactions on Intelligent Transportation Systems*, 18(8):2232–2249, 2017.
37. Gennady Andrienko, Natalia Andrienko, Georg Fuchs, and Jo Wood. Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE transactions on visualization and computer graphics*, 23(9):2120–2136, 2016.
38. Urska Demsar, Jed A. Long, Fernando Benitez-Paez, Vanessa Brum Bastos, Solène Marion, Gina Martin, Sebastijan Sekulic, Kamil Smolak, Beate Zein, and Katarzyna Sila-Nowicka. Establishing the integrated science of movement: bringing together concepts and methods from animal and human movement analysis. *International Journal of Geographical Information Science*, 35(7):1273–1308, 2021.
39. Hendrik Edelhoff, Johannes Signer, and Niko Balkenhol. Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement ecology*, 4(1):1–21, 2016.
40. Christoph Schulz and Arlind Nocaj and Jochen Görtler and Oliver Deussen and Ulrik Brandes and Daniel Weiskopf. Probabilistic Graph Layout for Uncertain Network Visualization *IEEE Transactions on Visualization and Computer Graphics*, 531–540, 23, 2017.
41. J. Spoerhase, S. Storandt and J. Zink. Simplification of Polyline Bundles SWAT, 35:1–35:20, 2020.
42. Yannick Bosch, Peter Schäfer, Joachim Spoerhase, Sabine Storandt and Johannes Zink. Consistent Simplification of Polyline Tree Bundles COCOON, 231-243, 2021.
43. Ulrik Brandes and Martin Mader. A Quantitative Comparison of Stress-Minimization Approaches for Offline Dynamic Graph Drawing *Graph Drawing*, 99–110, 2011.

Bionotes



Ying Zhang
University of Konstanz, Department of
Computer and Information Science,
D-78457, Konstanz, Germany
ying.zhang@uni-konstanz.de

Ying Zhang is a PhD student in computer science at the University of Konstanz and the Max-Planck Institute for Animal Behavior. Her research is on immersive analytics for animal behavior research.



Dr. Karsten Klein
University of Konstanz, Department of
Computer and Information Science,
D-78457, Konstanz, Germany
karsten.klein@uni-konstanz.de

Dr. Karsten Klein is a Postdoctoral Researcher with the University of Konstanz, Germany, where he works on the design and implementation of visual and immersive analytics approaches for complex data from application areas, especially the life sciences, as well as on network analysis and visualization methods and the underlying graph theoretic principles.

Prof. Dr. Oliver Deussen
University of Konstanz, Department of Computer and Information
Science, D-78457, Konstanz, Germany
oliver.deussen@uni-konstanz.de

Prof. Dr. Oliver Deussen graduated at Karlsruhe Institute of Technology (KIT) and is professor at University of Konstanz. He served as visiting professor at the Chinese Academy of Science in Shenzhen (SIAT), was President of the Eurographics Association and is speaker of the Excellence Cluster 'Centre for the Advanced Study of Collective Behaviour'. His areas of interest encompass Information Visualization, modeling and rendering of complex systems as well as non-photorealistic rendering.

Theodor Gutschlag
University of Konstanz, Department of Computer and Information
Science, D-78457, Konstanz, Germany
theodor.gutschlag@uni-konstanz.de

Theodor Gutschlag is a Master student in computer science at the University of Konstanz.

Prof. Dr. Sabine Storandt
University of Konstanz, Department of Computer and Information
Science, D-78457, Konstanz, Germany
sabine.storandt@uni-konstanz.de

Prof. Dr. Sabine Storandt is a professor at the University of Konstanz, leading the group on Algorithmics. Her research interests are algorithm engineering, combinatorial optimization and graph algorithms.