Results of the break-out group: Aggregation

Group discussion with Mark de Berg, Jörg Sack, Bettina Speckmann, Anne Driemel, Maike Buchin, Monika Sester, and Marc van Kreveld.

The assumption is that trajectories are acquired from a similar area, possibly at different points in time. Trajectories can e.g. be traces of animals, of hiker or of cars.

1.1 Why aggregation?

Aggregation of trajectories is needed for several reasons: one is to simplify the information and thus ease the communication (similar to cartographic generalization). The other reason is to extract meaning from the possibly noisy input data and extract the underlying common behaviour or the underlying structure which guides the behaviour. The latter is important for the application of digitizing roads by the GPS traces of car drivers.

1.2 Input data in detail

Input data are a set of trajectories which shows the movement behaviour of different objects or phenomena. They can be inaccurate and outliers can be involved.

1.3 Goal

The goal of the aggregation is to determine an aggregated trajectory, which represents the common behaviour. The representative geometry can either be one of the original trajectories (or a part of an original trajectory), or a geometry average of the input geometries.

It can also be a generalized version of the aggregated structure, e.g. in order to enhance the typical structure and make it more clear to understand.

1.4 Prerequisite

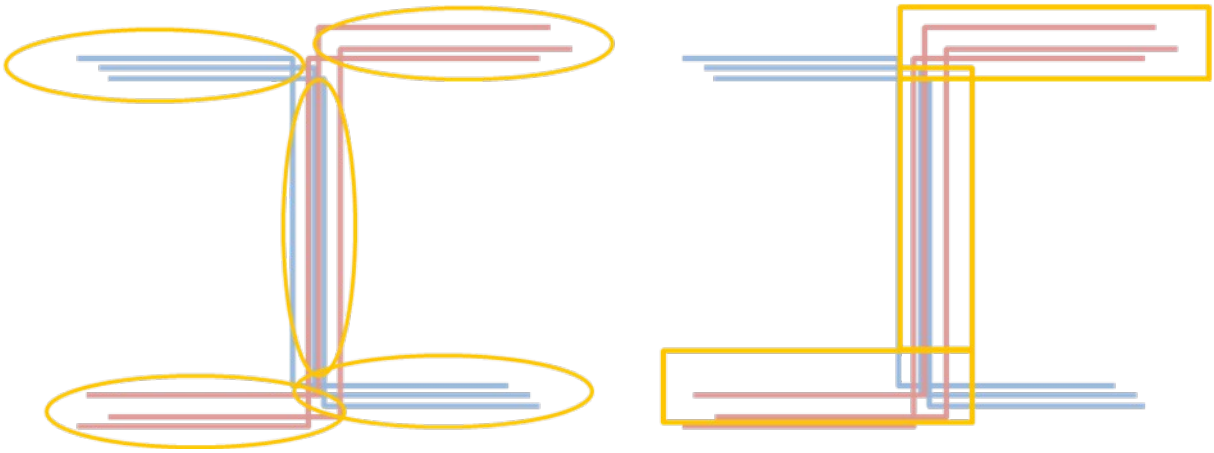
Before trajectories can be aggregated, a similarity measure is needed which defined, which trajectories represent the common behaviour (which is defined as a coincidence in space – and possibly in time). A similarity measure could e.g. be a distance value d. This means that all trajectories (or trajectory parts) are defined a similar, which are less than distance d apart from each other.

Problems and special cases

1.5 Segmentation

When only parts of trajectories are aggregated, segmentation has to be conducted. There arises the problem of what is a correct or adequate segmentation. Consider e.g. the following example of basically two traced (red and blue), which overlap at one section of the path: should the segmentation (and the subsequent aggregation) be based on the similarity of the whole trajectories (as in the case on the right hand side, which leads to two aggregated trajectories (the red and the blue one); or should the segmentation consider the

longest common stretches of all trajectories, in which case the middle section would be the segment with most trajectories. In this case, five aggregates would be generated (see left figure).



1.6 When to select median vs. average geometry?

Depending on the situation, the selection of a median trajectory could be more adequate than taking the average geometry. In the case in the following figure, hiking traces are recorded, which show how hikers walk around a lake. Taking the average geometry would lead to the situation that the path seemingly would go through the lake (left figure); in the case of a median trajectory, one representative would be selected (right figure), which has the benefit that it is a valid trajectory.

Obviously, the decision also depends on the scale: in a smaller scale, the lake may no longer be present and then a more simple (i.e. straight) representation would be required.

1.7 How to simplify aggregated structure?

In cartographic generalization, schematization is often applied (cf. Metro maps). This operation can also be transferred to the visualization of aggregated trajectories. It has to be investigated, which kinds of schematizations are required for spatio-temporal data.