Uncertainty Types in Segmenting and Labeling Time Series Data

M. Bögl[1], C. Bors[1], T. Gschwandtner[1] and S. Miksch[1]

[1]*Institute of Visual Computing and Human-Centered Technology, TU Wien, Vienna, Austria*

The segmenting and labeling of time series can be described as a three stage process. Given a multivariate time series, multiple (pre-)processing steps are applied to make the data useful for subsequent analysis. Secondly, appropriate segmenting and labeling algorithms are utilized to provide meaningful segments and labels. Different algorithms and parameter inputs generate different results and subsequently need to be explored for the most adequate result. However, all steps are affected by uncertainty. Throughout the process they are introduced and altered, thus affecting the reliability and trustworthiness of the results.

In this work we identify four types of uncertainty that influence this process and furthermore the latter decision making. (a) *Value uncertainty* is comprised of uncertainties inherent in the data input and stemming from processing routines that affect the value domain (e.g., noise reduction). (b) *Result uncertainty* stems from the likelihood of a segment actually representing a particular label/result, and the definitive start and end time, being contingent on the implicit transitions between labels. Uncertainty will be (c) *aggregated* by various processing steps but also when being visualized, if multiple segmentation results are shown and screen resolution is insufficient. (d) *Cause and effect uncertainty* is implicit, stemming from consecutively adjusting and comparing different algorithms and/or algorithm parametrizations.

In order for users to judge which algorithm and parameter configurations provide adequate results while trading the influence of uncertainty on the outcome, it is necessary to inspect different stages of analysis with uncertainty information available. To accomplish this we derived different types of visual elements from [1] to effectively convey the aforementioned types of uncertainty throughout the entire analysis pipeline. Depending on the analysis task, the associated uncertainty needs to be visually communicated appropriately within the visual representation of the data. During the (pre-)processing step, inherent *value uncertainty* is visually externalized in an enhanced processing view. When selecting potential algorithms and parameter configurations, the influence of *result uncertainty* is shown for algorithms and parameters individually. In the last step of analysis, the effects of *result* and *aggregation uncertainties* require attention to distinguish if segmenting and labeling results were correctly determined, and how credible they are.

Keywords: Time Series, Uncertainty, Segmenting and Labeling.

References

[1] T. Gschwandtner, M. Bögl, P. Federico, and S. Miksch (2016). Visual Encodings of Temporal Uncertainty: A Comparative User Study. In *IEEE Transactions on Visualization and Computer Graphics*, 539–548.