

A System for Temporally-Sliced Correspondence Analysis Visualization

Artur Šilić,

University of Zagreb, Faculty of Electrical Engineering and Computing, artur.silic@fer.hr

Annie Morin

IRISA, Université de Rennes 1, 35042 Rennes Cedex, France, annie.morin@irisa.fr

Jean-Hugues Chauchat

ERIC-Lyon2, Université de Lyon, France, jean-hugues.chauchat@univ-lyon2.fr

Bojana Dalbelo Bašić*

University of Zagreb, Faculty of Electrical Engineering and Computing, bojana.dalbelo@fer.hr

This work describes the CatViz system that enables a user to visually explore temporal changes in large-scale multivariate data. CatViz is based on Correspondence Analysis - a dimension reduction technique that uses SVD and allows to explore correspondence relations among matrix rows and columns. Using CatViz we depict the relations among moments in time and variable states. Particularly, in two case studies, we show how CatViz is used on the task of large news corpus analysis. We present typical tasks of exploratory text analysis and discuss application procedures that an analyst might perform.

Next, we present the CatViz System, its capabilities and modes of usage. We describe the software architecture of the CatViz System and then we advocate for the design choices that have been made in order to achieve computational efficiency.

Because of intuitiveness, effectiveness and robustness, we expect that the CatViz System will enable exploration and understanding of various data types and especially of texts in huge historical archives and contemporary news streams.

Key Words: dimension reduction, visualization, temporal text mining