



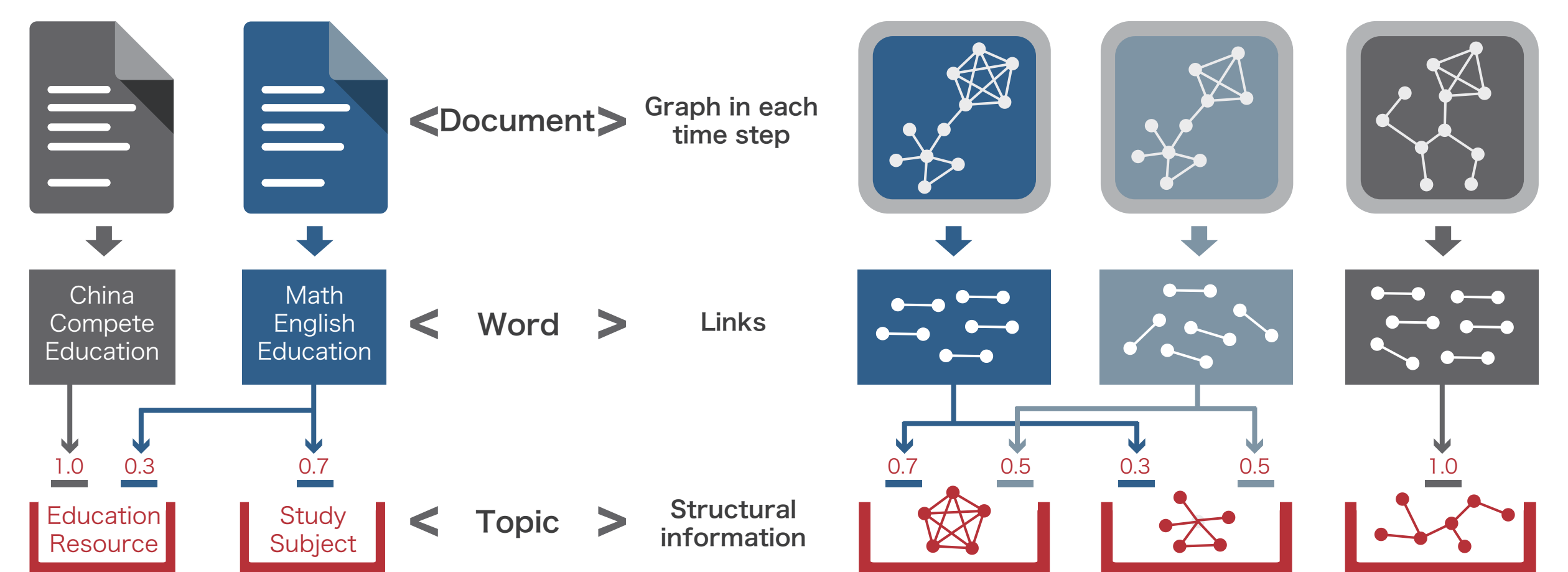
GraphLDA: Latent Dirichlet Allocation-based Visual Exploration of Dynamic Graphs

Lijing Lin Siming Chen Fan Hong Chufan Lai Shuai Chen Xiaoru Yuan
Key Laboratory of Machine Perception (Minister of Education), and School of EECS, Peking University, Beijing, China

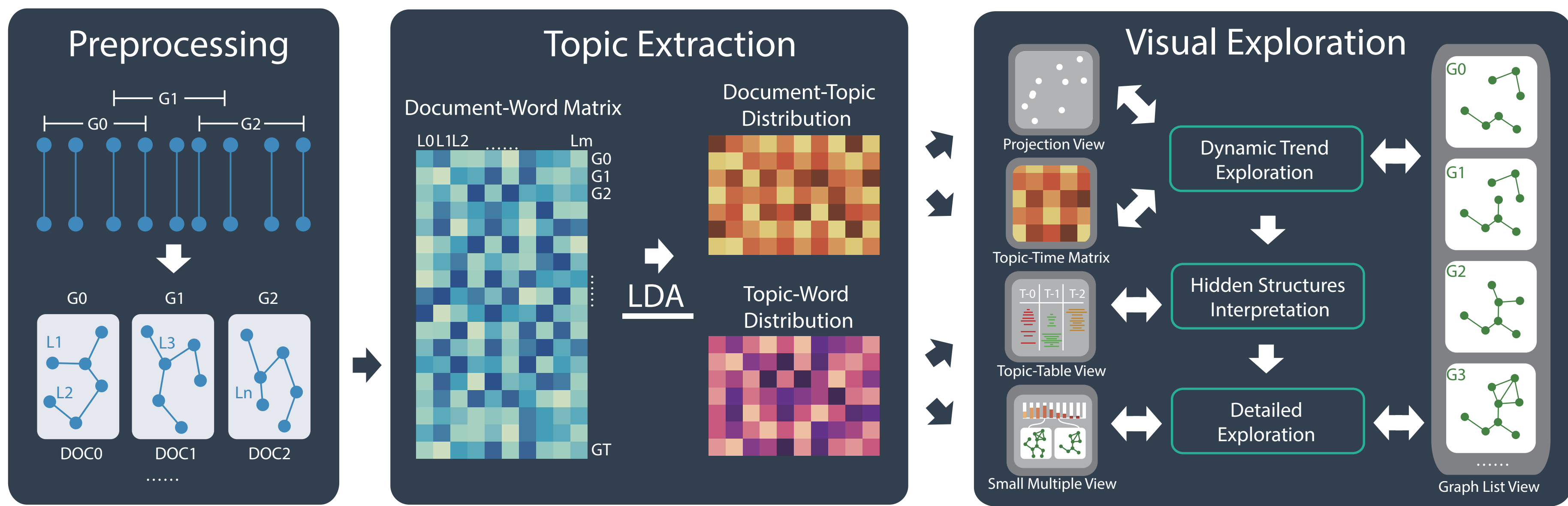
INTRODUCTION

We propose a latent Dirichlet allocation (LDA)-based visual exploration method for dynamic graphs. With the LDA-based analysis, we can reveal hidden structures in the dynamic graph based on the extracted semantic topics. To gain a deeper understanding of the derived structures and their evolution, we design a visual analytics system with an analytical pipeline enabling users to explore the dynamic graph. A new visual interface supports users to iteratively explore and interpret the LDA results based on the salience in the derived structures.

GRAPHLDA



PIPELINE



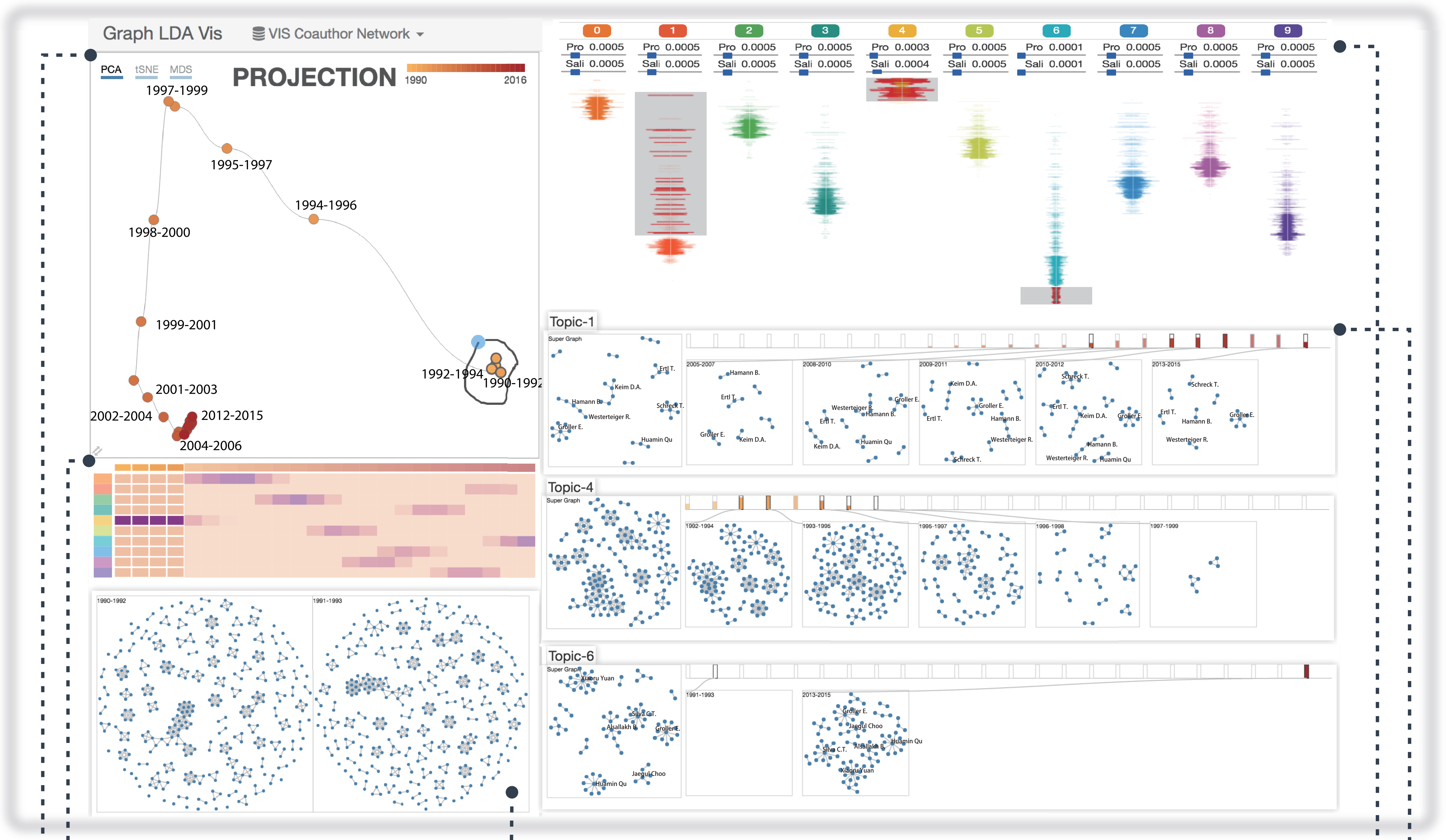
Preprocessing: The raw dynamic graphs data are preprocessed to be the input matrix of LDA.

Topic Extraction: With the input of the document word matrix, the LDA would output two parts of results – the document–topic distribution and topic–word distribution.

Visual Exploration: Visual system supports users to iteratively explore the LDA results from over-view to details within a visual analytics loop.

CO-AUTHOR NETWORK EXPLORATION

Data Papers published in IEEE VIS from 1990 to 2015. Nodes (4813) are authors; links (14033) are cooperations.



Projection View Visualizing the overview of the similarity relation between graphs in different time steps. There are three clusters of projected points (1990–1994, 1996–1999, 2004–2015), which means that there are three stable periods in the co-author graph of VIS.

Topic-Time Matrix View Showing topics (structures) probability distribution in different time steps. Lasso and select cluster during 1990–1994. The view shows that these graphs have a high probability in the same derived topic (Topic-4).

Graph-List View Showing the snapshots of raw graphs. It shows raw graphs of selected points in Projection View.

Topic-Table View Showing words (relations) probability distributions of extracted topics (hidden structures). The salient relations of Topic-4 are positioned at the front, which means that these relations are appeared in the early periods. In some topics, the position of some salient relations is separated from most ones, such as Topic-1, 3, etc.

Small multiples View Showing the evolution of selected salient relations of a specific topic. We select some salient relations in Topic-1. These cooperation relations start in about 2000 and are very active from about 2010. These relations are about Huamin Qu, Ertl T., etc.

ACKNOWLEDGEMENTS

This work is supported by NSFC No. 61672055 and also funded by PKU-Qihoo Joint Data Visual Analytics Research Center.



CONTACT

xiaoru.yuan@pku.edu.cn
http://vis.pku.edu.cn