

Addendum to the Special Issue on Interactive Data Exploration and Analytics (TKDD, Vol. 12, Iss. 1): Introduction by the Guest Editors

Surprisingly, few technologies allow us to freely “wander” around data and make discoveries by following our intuition, or serendipity. While standard data mining aims at finding highly interesting results, it is typically computationally demanding and time consuming, thus may not be well-suited for interactive exploration of large datasets. The focus of this special issue is, therefore, on *interactivity* and the effective *integration* of techniques from *data mining*, *visualization*, and *human–computer interaction*; we explore how the best of these different but related domains can be combined such that the *sum is greater than its parts*. After four successful workshops on *Interactive Data Exploration and Analytics* (IDEA) at ACM SIGKDD 2013, 2014, 2015, and 2016, the goal of this special issue is to further promote and bring attention to this exciting and highly active interdisciplinary research area.

We initially received 43 submissions in response to an open call for papers. All submissions went through the standard, thorough TKDD review process to meet the high quality standard of the journal. In the end, 10 papers were selected for publication in the special issue.

These 10 accepted papers cover many different aspects of interactive data exploration and analytics. Three papers concern the interactive discovery of different types of patterns from data. *GrammarViz 3.0: Interactive Discovery of Variable-length Time Series Patterns*, by Pavel Senin, Jessica Lin, Xing Wang, Tim Oates, Sunil Gandhi, Arnold Boedihardjo, Crystal Chen, and Susan Frankenstein, presents GrammarViz 3.0, a tool that enables the user to interactively discover variable-length patterns in time series based on symbolic discretization and grammar inference. *Interactive Discovery of Coordinated Relationship Chains with Maximum Entropy Models*, by Hao Wu, Maoyuan Sun, Peng Mi, Nikolaj Tatti, Chris North, and Naren Ramakrishnan, introduces MERCER, a visual analytic system that uses novel maximum entropy models to discover surprising chains of relationships. *Mining Redescriptions with Siren*, by Esther Galbrun and Pauli Miettinen, introduces Siren, an interactive tool for mining and visualizing redescriptions, i.e., two semantically different descriptions of roughly the same set of objects.

Three other papers investigate interactive information retrieval for three different types of resources. *VisIRR: An Interactive Visual System for Information Retrieval and Recommendation for Large-scale Document Data*, by Jaegul Choo, Hannah Kim, Edward Clarkson, Zhicheng Liu, Changhyun Lee, Fuxin Li, Hanseung Lee, Ramakrishnan Kannan, Charles Stolper, John Stasko, and Haesun Park, presents VisIRR, an interactive visual information retrieval and recommendation system for large-scale document discovery that organizes retrieved documents into topics. *ATR-Vis: Visual and Interactive Information Retrieval for Parliamentary Discussions in Twitter*, by Raheleh Makki, Eder Carvalho, Axel Soto, Stephen Brooks, Maria Ferreira de Oliveira, Evangelos Milios, and Rosane Minghim, presents ATR-Vis, a user-driven visual analytics tool for the retrieval of Twitter content that enables non-technical users to find relevant tweets with minimal effort. *A Viewable Indexing Structure for the Interactive Exploration of Dynamic and Large Image Collections*, by Frédéric Rayar, Sabine Barrat, Fatma Bouali, and Gilles Venturini, describes a web

platform that jointly indexes and visualizes large image collections, and as such allows for intuitive and interactive exploration of up to millions of images.

Two papers investigate how users can be enabled to better understand models of data by means of visualization and interaction. *Visual Analysis of Brain Networks using Sparse Regression Models*, by Lei Shi, Hanghang Tong, Madelaine Daianu, Feng Tian, and Paul Thompson, proposes an integrated visual analytics framework for the exploration and comparison of human brain network. *CommunityDiff: Visualizing Community Clustering Algorithms*, by Srayan Datta and Eytan Adar, presents CommunityDiff, an interactive system that combines visualization and active learning to support users in the analysis of networks based on multiple community detection methods.

Finally, two papers explore how sampling-based methods can contribute to instant and interactive data exploration and analytics. *A Session-Based Approach to Fast-But-Approximate Interactive Data Cube Exploration*, by Niranjan Kamat and Arnab Nandi, proposes Sesame, a rewrite and caching framework that accelerates interactive query sessions of aggregation queries over sampled data while maintaining accurate results. *Memory-efficient and Accurate Sampling for Counting Local Triangles in Graph Streams: From Simple to Multigraphs*, by Yongsu Lim, Minsoo Jung, and U Kang, presents algorithms for local triangle counting based on edge sampling, so that local triangle counts in a graph stream can be accurately estimated without storing the whole graph.

We sincerely thank the authors of all these papers for their exciting contributions to this special issue. We wish to thank the anonymous reviewers for their invaluable help in selecting a set of high-quality papers and providing constructive comments that have helped the authors to improve their papers. Finally, we hope that you, the reader, will find the papers enticing and that they will inspire you to conduct novel research that will further advance the emerging field of *Interactive Data Exploration and Analytics*.

Matthijs van Leeuwen
Leiden University

Polo Chau
Georgia Tech

Jilles Vreeken
Max Planck Institute for Informatics and Saarland University

Dafna Shahaf
The Hebrew University of Jerusalem

Christos Faloutsos
Carnegie Mellon University

Guest Editors