

# Exploration: Overview

Andreas Nürnberger

Data and Knowledge Engineering Group,  
Faculty of Computer Science, Otto-von-Guericke-University, Germany  
`andreas.nuernberger@ovgu.de`

## 1 Introduction

In the previous chapters of this book quite different approaches to create networks based on existing data collections (Part II) have been discussed and diverse methods for network analysis have been proposed (Part III). All these methods provide powerful means in order to obtain different insights into the properties of huge information networks or graphs. However, one disadvantage of these individual approaches is that each approach provides only specific facets of information to the end user. Integrated exploration tools that enable the interactive exploration of huge graphs and data collections using data visualization, aggregation and mining methods could be much more beneficial for the (interactive) task of finding bisociations. Therefore, we present and discuss in this part of the book methods for interactive exploration that bring these fields together.

## 2 Contributions

This part starts with a more general discussion of methods for interactive data exploration. Therefore, the chapter of Gossen et al. [1] provides first a brief review of methods and tools for the interactive exploration of graphs with a focus on approaches to support bisociative discoveries. Furthermore, a critical discussion of the challenges of evaluating the performance and quality of exploration tools is given.

In the second chapter of Haun et al. [2] the Creative Exploration Toolkit (CET) is presented, which was developed as part of the BISON project. CET is a user interface for graph visualization designed towards explorative tasks. It supports the integration and communication with external data sources and mining tools, especially the data-mining platform KNIME. Besides the tool itself, the chapter also presents the results of a small case study, in which the applicability of the tool for bisociative discoveries is analyzed.

In the chapter “Bisociative Knowledge Discovery by Literature Outlier Detection”, Petrič et al. [4] discuss the role of outliers in literature-based knowledge discovery. They show that outlier documents can be successfully used as means of detecting bridging terms that connect documents of two different literature sources.

The chapter “Exploring the Power of Outliers for Cross-Domain Literature Mining” of Sluban et al. [5] proposes an approach to find outliers that are potential candidates of documents that discuss bisociative discoveries bridging different scientific domains. The discussed approach aims at finding cross-domain

links by mining for bridging concepts or terms (b-terms) and can be used for an exploration task using iterative search and filter methods.

An example of an exploration tool that is based on the idea of Sluban et al. [5] is presented in the final chapter by Juršič et al. [3]. Juršič et al. introduce a system for bisociative literature mining called *CrossBee*. This system is focussed on b-term identification and ranking and supports the search for hidden links connecting two different domains. The chapter contains a detailed description of the proposed methodology that is implemented in *CrossBee*. Furthermore, an experimental evaluation on two datasets, Migraine-Magnesium and Autism-Calcineurin, is reported.

### 3 Conclusions

Interactive exploration methods provide powerful means to support bisociative discoveries in huge data collections and networks. The contributions presented in this part of the book, provide some insights into their capabilities and the challenges these approaches have to face. The tools presented can – together with the critical discussions – serve as a basis for the development of advanced exploration methods that can enable interactive bisociative discoveries.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

### References

1. Gossen, T., Nitsche, M., Haun, S., Nürnberger, A.: Data Exploration for Knowledge Discovery: A brief Overview of Tools and Evaluation Methods. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 287–300. Springer, Heidelberg (2012)
2. Haun, S., Gossen, T., Nürnberger, A., Kötter, T., Thiel, K., Berthold, M.R.: On the Integration of Graph Exploration and Data Analysis: The Creative Exploration Toolkit. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 301–312. Springer, Heidelberg (2012)
3. Juršič, M., Cestnik, B., Urbančič, T., Lavrač, N.: Bisociative Literature Mining by Ensemble Heuristics. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 338–358. Springer, Heidelberg (2012)
4. Petrič, I., Cestnik, B., Lavrač, N., Urbančič, T.: Bisociative Knowledge Discovery by Literature Outlier Detection. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 313–324. Springer, Heidelberg (2012)
5. Sluban, B., Juršič, M., Cestnik, B., Lavrač, N.: Exploring the Power of Outliers for Cross-Domain Literature Mining. In: Berthold, M.R. (ed.) Bisociative Knowledge Discovery. LNCS (LNAI), vol. 7250, pp. 325–337. Springer, Heidelberg (2012)