

# Smart Query Definition for Content-Based Search in Large Sets of Graphs

Tatiana von Landesberger and Sebastian Bremm and Jürgen Bernard and Tobias Schreck

Interactive Graphics Systems Group, Technische Universität Darmstadt, Germany



## Motivation

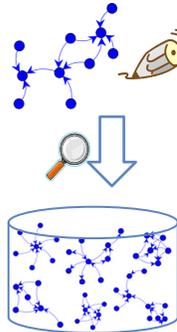
- Finding interesting objects in large databases is a common problem in image retrieval, 3D model retrieval, biochemistry, and other areas.
- The exact definition of the query is the basis for the search process.

### Query definition in graph databases:

- Direct graph editing: precise but time-consuming.
- Pre-defined templates: fast but restricted to set of templates.

Query definition does not reflect the object distribution in the database.

- ➔ Need for smart user-assistance in query definition.



## Approach

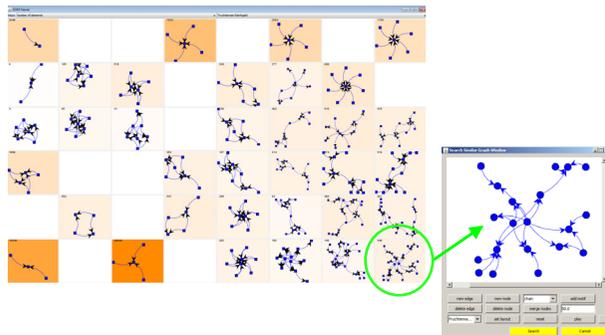
We propose **three ways of defining the query graph**:

- *Smart choice of data samples*: Using an example graph from the database, based on search for graph attributes and the overview of representative graphs in the database.
- *Graph sketching supported by data-dependent graph building blocks*: Using a graph interactively created (drawn, edited, sketched) by the analyst.
- *Combining smart sketching with data samples*: Using a graph, or a set of example graphs from the database as a basis for further editing of the final query graph.

These approaches support **flexible, efficient, and data-oriented query graph definition**.

## Smart Data Samples

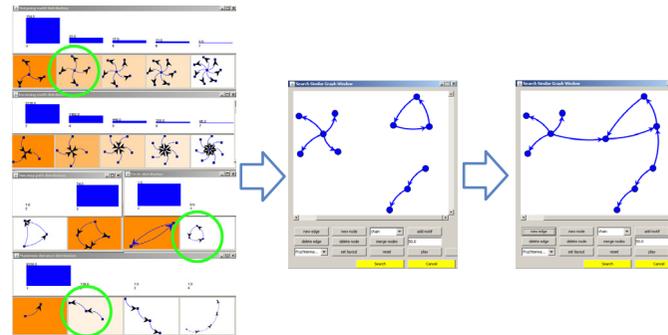
Using SOM clustering results for defining the query graph



Left: The visualization of SOM clustering results – overview of prototype graphs from the database with color-indication of their frequency.  
Right: The selected graph query object from the SOM grid.

## Smart Graph Sketching

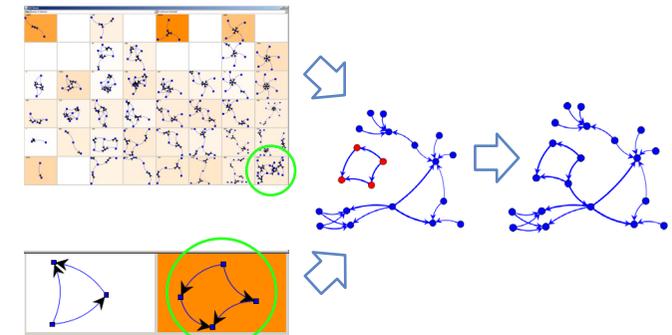
Combining sketching and data-dependent graph building blocks



Left: Distribution of building blocks in the data set showing the composition of the dataset graphs.  
Center: The selection of data-dependent building blocks for creating the query objects.  
Right: The final query graph created by connecting and adapting the chosen baseline building blocks with sketched edges and merged nodes.

## Combination

Synergies of smart sketching and smart data samples



Left: Graph samples as a result of SOM clustering and distribution of graph building blocks.  
Center: The selection of a graph sample and a building block for creating the query graph.  
Right: The final query graph combining a graph sample and building blocks with sketched edges.

## Future Work

- Support for similarity-based graph retrieval with interactive visual result exploration.
- Support for additional graph attributes (node labels and edge weights) and wildcard query definition.
- Experimental evaluation on real-world data sets and incorporating user feedback.

## References

- von Landesberger T., Görner M., Schreck T.: Visual analysis of graphs with multiple connected components. Proceedings of IEEE Symposium on Visual Analytics Science and Technology (2009).
- Wong P. C., Foote H., McKey P., Perrine K.: Generating graphs for visual analytics through interactive sketching. IEEE Transactions on Visualization and Computer Graphics, 12 (2006).
- Adinavicius G., Tuzhilin E.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17 (2005).

## Acknowledgments

This work was partially funded by German Federal Ministry of Economics and Technology within the THESEUS project and by the German Research Foundation DFG within Priority Programme 1335: Scalable Visual Analytics.

