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

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

Smart Graph Sketching
Combining sketching and data-dependent graph building blocks

Combination
Synergies of smart sketching and smart data samples

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

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.