Where we are

Background (15 min)
Graph models, subgraph isomorphism, subgraph mining, graph clustering

Exploratory Graph Analysis (40 min)

Focused Graph Mining (40 min)

Refinement of Query Results (40 min)

Machine Learning and Visualization (40 min)

Challenges and discussion
Approximate Queries

• User query is imprecise

By-Example methods

• User query is an example result

• Only need a partial knowledge on the data
• No need for complicate query languages (use examples, partial descriptions)
• The query adapts to user need
• Enable exploratory search by using small queries on the data
Challenges for Exploratory Graph Analysis

- **Unsupported** in most of the current graph databases
- **No “universal” index** to answer multiple type of queries
- **Partitioning methods** for approximate query answering

- **User interactivity** in the exploration process
- **No solutions for probabilistic graphs**
- **Respond to queries in dynamic graphs**
- **Find examples in streaming settings**

- **Exploiting query logs for personalized query answering**
- **Retrieve results in form of documents converting the query structures**
Summary of Focused Graph Mining

The focus on individual user interest
... as **Query** to the Graph Mining System
... as **Seed Node(s)** for Local Search
... as **Attributes** and **Weights**

- get or infer user interest  ➔ unexpected results
- interactive exploration  ➔ intuitive parametrization
- adaptive graph mining  ➔ individual local search
Challenges for Focused Graph Mining

User interactivity in the graph mining process
• unsupported in most of the current graph mining algorithms
• huge variety of user interactions possible
• feedback loop needs to be unified and become exchangeable

Revolution of formal models and search algorithms
• insufficient extensions of existing models and algorithms
• adaptive steering of algorithms vs. fixed parametrization
• evaluation of algorithms with user studies

Scalability of algorithms for real-time interaction
• NP-hard problems, heuristic algorithms, ..., still not scalable
• exploit the user interest for pruning the search space
Summary of Refinement of Query Results

Refinement

- The user query is too restrictive or too generic

Top-k Results

- Queries typically have inexact matches

Skyline Queries

- Find small set of interesting items with many dimensions and incremental updates

- The user might have a very generic idea of how to describe the structure of interest
- The system guides the user towards the answer with simple steps
- The results are explained with reformulations
- The queries can be inexact
Challenges for Refinement of Query Results

- Profiling of queries for optimized performance
- Provenance and explainability of queries
- Managing uncertainty in data

- Personalized reformulations and interactivity
- Facet search discovery in graphs
- Learning of user preferences while refining

- Real time performance not achieved
- Avoiding traverse the entire space using query workloads and query logs
Towards a graph exploration system

Users

Graph
Towards a graph exploration system

Users

Interactive algorithms

Intuitive queries

Adaptive Methods and Structures

Graph
Missing tiles for Graph Exploration

- Interactivity
- Personalization
- Adaptivity
- Scalability
Graph exploration
Let me Show what is Relevant in your Graph


Not all who wander are lost.

*J R R Tolkien*

More on graphs? Visit our new papers:

- Anton Tsistulin, Davide Mottin, Panagiotis Karras, Alex Bronstein, Emmanuel Müller. **NetLSD: Hearing the Shape of a Graph.** *KDD 2018 – Tuesday 16.00-18.00 - ICC Capital Suite Room 6+13*
- Anton Tsistulin, Davide Mottin, Panagiotis Karras, Alex Bronstein, Emmanuel Müller. **SGR: Self-Supervised Spectral Graph Representation Learning.** *DLDay - Poster session*
- Lukas Faber, Tara Safavi, Davide Mottin, Danai Koutra, Emmanuel Müller. **Adaptive Personalized Knowledge Graph Summarization.** *MLG Workshop @ KDD*
Acknowledgements

We thank the authors of the papers who provided some of the slides included in this tutorial.
References


References


[Staudt14] C. Staudt, Y. Marrakchi, H. Meyerhenke: Detecting Communities Around Seed Nodes in Complex Networks (BigData 2014)


[Iglesias14] Iglesias et al. Local Context Selection for Outlier Ranking in Graphs with Multiple Numeric Node Attributes (SSDBM 2014)

[Iglesias13] Iglesias et al. Statistical Selection of Congruent Subspaces for Mining Attributed Graphs (ICDM 2013)

[Perozzi14] Perozzi et al. Focused Clustering and Outlier Detection in Large Attributed Graphs (KDD 2014)
References


[Ranu14] Ranu, S., Hoang, M. and Singh, A. Answering top-k representative queries on graph databases. SIGMOD, 2014


[Zou10] Zou, L., Chen, L., Özsö, M.T. and Zhao, D. Dynamic skyline queries in large graphs. DASFAA, 2010