Let me show what is relevant in your graph.
Who we are

Davide Mottin
• graph mining, novel query paradigms, interactive methods
• user-centric methods, efficient similarities
http://mott.in

Emmanuel Müller
• data mining, stream mining, clustering and outlier mining
• graphs, streams, and traditional databases
http://data.bit.uni-bonn.de/
Complete slides:

Big data and novice users
Data exploration

Efficiently extracting knowledge from data even if we do not know exactly what we are looking for

Idreos et al., Overview of Data Exploration Methods, SIGMOD 2015
2.07B users
300B friendships
7B$ revenue
350M photos/day
570 M entities
18B relationships

Lost in the graph?
Current: Visualization tools

Several visualization tools:
• General: Gephi, GraphViz, ...
• Biological: Cytoscape, Network Workbench
• Social: EgoNet, NodeXL, ...
• Relational: Tulip

but

• No Scalability to large networks!
• No for novice users
• Limited expressivity
Current: Query languages

**SELECT** ?name ?email

**WHERE**

{  
  ?person a foaf:Person .  
  ?person foaf:name ?name .  
}

**Query languages ARE:**
- Expressive
- Powerful
- Scalable
- Compact

**but**
- **Not** user friendly
- **No** guided search
- **Not** interactive
Efficiently extracting knowledge from graph data even if we do not know exactly what we are looking for.

Graph Exploration: From Users to Large Graphs. CIKM 2016, SIGMOD 2017
This tutorial is about …

- Algorithms for helping the user finding the wanted information
- Approximate search on graphs to assist the user in finding the information
- Interactive methods on graphs based on user feedback
- Automatically discovery of portions of graphs using examples

NOT about

- Visualization methods for graphs
- Query languages and semantics
- Efficient indexing methods
- Pure machine learning on graphs
Our graph exploration taxonomy

- Exploratory Graph Analysis
- Focused Graph Mining
- Refinement of Query Results
Graph exploration taxonomy

Exploratory Graph Analysis

Other politicians like Angela Merkel?

Two exploratory options:
1. An imprecise query
   - Merkel President of ?

2. A by-example query
   - Merkel Chancellor Germany

Query is an example

Kohl chancellor Germany

Schröder chancellor Germany

Gauch
Focused Graph Mining

Targeted analysis on large graphs
1. Focused graph clustering
2. Space restriction methods
3. Graph Reweighting

How can I see only the part of the graph I’m interested in?

They all like the Chelsea
Ego-net analysis
Graph exploration taxonomy

Refinement of Query Results

Where is this molecule contained?

Dealing with generic queries:
1. Reformulation and refinement
2. Top-k results
3. Skyline queries

Dominance relation

Query 1: \(\text{OH} \rightarrow \text{S} \rightarrow \text{O} \)
Query 2: \(\text{OH} \rightarrow \text{S} \rightarrow \text{O} \)

- 270 results
- 220 results

... + 500 results
The graph exploration ... graph

Exploratory graph analytics

- [Mottin14]
- [Jayaram15]
- [Khan13]
- [Yang14]
- [Fan10]
- [Ma14]

Refinement of query results

- [Ranu14]
- [Mottin15]
- [Wu13]
- [Fan13]
- [Vasilyeva16]
- [Gupta14] [Zou10]

Focused graph mining

- [Tong06]
- [Staudt14]
- [Iglesias14]
- [Epasto15]
- [Perozzi14]
- [Iglesias13]
Connection to other tutorials @ KDD

- **Explaining similarities**
  - Graph and Tensor Mining for Fun and Profit
  - Graph Metric Spaces

- **Effective graph mining**
  - Modeling Data With Networks + Network Embedding: Problems, Methodologies and Frontiers

- **Scalability and beyond**
  - Large-Scale Graph Algorithmics: Theory and Practice
  - Graph Sketching, Sampling, Streaming, and Space-Efficient Optimization
Tutorial outline

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