186.868 VU Visual Data Science WS 3.0 ECTS
WS 2018
Visual Graph Analytics

Hsiang-Yun Wu

Institute of Visual Computing & Human-Centered Technology
TU Wien, Austria
A Story of a Start-Up Company

Outline

- What is a Graph? What is a Network?
- Tasks in Visual Graph Analytics
- Factors Affecting Visual Graph Analytics
- A tutorial of Visual Graph Analytics Tool
Outline

■ What is a Graph? What is a Network?
  ■ Graph: A data structure that records a relationship
  ■ Network: A graph associated with attributes

■ Tasks in Visual Graph Analytics

■ Factors Affecting Visual Graph Analytics

■ A tutorial of Visual Graph Analytics Tool
Graph Terminology

- **Vertices** (nodes) connected by **Edges** (links)
  - Graphs can have **cycles**
  - Graph edges can be **directed** or **undirected**
  - The **degree** of a vertex is the number of edges connected to it
    - In-degree and out-degree for directed graphs
  - Graph edges can have values: **edge weights** on them (nominal, ordinal or quantitative)

- Related to
  - Data structure, discrete math, graph theory, computational geometry, optimization, and ... etc.

Graphs with various layout

http://mathworld.wolfram.com/GraphEmbedding.html
Definition

- **Graph**
  \[ G = (V, E) \]

- **Vertices**
  \[ V = \{v_1, v_2, ..., v_n\} \]

- **Edges**
  \[ E = \{\{v_i, v_j\} | v_i, v_j \in V\} \]
Definition

- **Graph** \( G = (V, E) \)
- **Edges** \( E = \{\{u, v\} | u, v \in V, u \neq v\} \)
- **Vertices** \( V = \{v_1, v_2, ..., v_n\} \)
What is a Tree?

- A special form of graph
  - no loops
  - no circuits
  - no self-loops

http://www.treevis.net
Network Visualization

- An intuitive representation for visualizing relationships

**Graph Drawing**
- A pictorial representation of a graph

**Node-Link Diagram**
- Vertices/Nodes: disks or labels
- Edges/Links: lines or curves in the Euclidean space.

**Network Visualization**
- Associated with attributes

[https://flowingdata.com/2016/05/31/air-transportation-network/]
Outline

- What is a Graph? What is a Network?
- Tasks in Visual Graph Analytics
- Factors Affecting Visual Graph Analytics
- A tutorial of Visual Graph Analytics Tool
Amar’s Low-Level Visual Analytic Tasks

- **Retrieve Value**: Given a set of cases, find attributes of those cases.
- **Filter**: Given some conditions on attributes values, find data cases satisfying those conditions.
- **Compute Derived Value**: Given a set of data cases, compute an aggregate numeric representation of those data cases. (e.g. average, median, and count)
- **Find Extremum**: Find data cases possessing an extreme value of an attribute over its range within the data set.
- **Sort**: Given a set of data cases, rank them according to some ordinal metric.
Amar’s Low-Level Visual Analytic Tasks

- **Determine Range:** Given a set of data cases and an attribute of interest, find the span of values within the set.

- **Characterize Distribution:** Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute’s values over the set.

- **Find Anomalies:** Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers.

- **Cluster:** Given a set of data cases, find clusters of similar attribute values.

- **Correlate:** Given a set of data cases and two attributes, determine useful relationships between the values of those attributes.
Tasks in Visual Graph Analytics [Lee et al., 2006]

- Topology-Based Tasks
- Attribute-Based Tasks
- Browsing Tasks
- Overview Task
Topology-Based Tasks

- **Adjacency (direct connection)**
  - Find the set of nodes adjacent to a node.
  - How many nodes are adjacent to a node?
  - Which node has a maximum number of adjacent nodes?

- **Accessibility (direct or indirect connection)**
  - Find the set of nodes accessible from a node.
  - How many nodes are accessible from a node?
  - Find the set of nodes accessible from a node where the distance is less than or equal to $n$.
  - How many nodes are accessible from a node where the distance is less than or equal to $n$?
Common Connection
- Given nodes, find a set of nodes that are connected to all of them.

Connectivity
- Find the shortest path between two nodes.
- Identify clusters.
- Identify connected components.
- Find bridges.
- Find articulation points.
Attribute-Based Tasks

- On the Nodes
  - Find the nodes having a specific attribute value.
  - Review the set of nodes.

- On the Links
  - Given a node, find the nodes connected only by certain types of links.
  - Which node is connected by a link having the largest/smallest value?
Browsing Tasks

- Follow Path
  - Follow a given path.

- Revisit
  - Return to a previously visited node

Pathfinder: Visual Analysis of Paths in Graphs


Pathfinder: Visual Analysis of Paths in Graphs

A compound exploratory task to get estimated values quickly

- Estimate the size of the social network
- Estimate k-core property

Outline

- What is a Graph? What is a Network?
- Tasks in Visual Graph Analytics
- Factors Affecting Visual Graph Analytics
- A tutorial of Visual Graph Analytics Tool
Factors Affecting Visual Graph Analytics

- Data Types & Information Retrieval
- Graph Layout
- Interaction
- Perception and Cognition
Data Types & Information Retrieval

- Graph Aggregation
  - Preferences
  - Knowledge
  - Social networks
  - Clustering
  - Logic
  - Biological network
- Data Mining
- Machine Learning

[Hairball effect]

[Yoghourdjian et al., 2018]
Factors Affecting Visual Graph Analytics

- Data Types & Information Retrieval
- Graph Layout
- Interaction
- Perception and Cognition
Graph Layout

- Force-Based Layout
- Hierarchical Layout
- Schematic Layout
- Minimizing edge crossings
- Minimizing edge bends
- Maximizing symmetry
- Maximizing the minimum angle between edges leaving a node
- Maximizing edge orthogonality
- Maximizing node orthogonality
- Maximizing consistent flow direction (directed graphs only)
Energy-Based Layout Algorithm [Wang et al, 2017]

IEEE InfoVis 2017
Revisiting Stress Majorization as a Unified Framework for Interactive Constrained Graph Visualization

Yunhai Wang, Yanyan Wang, Yinqi Sun, Lifeng Zhu, Kecheng Lu,
Chi-Wing Fu, Michael Sedimair, Oliver Deussen, and Baoquan Chen
Factors Affecting Visual Graph Analytics

- Data Types & Information Retrieval
- Graph Layout
- Interaction
- Perception and Cognition
Shneiderman's mantra:
Overview first, zoom and filter, then details-on-demand

[Yoghourdjian et al., 2018]
Interaction

- Select: Mark something as interesting

- Examples:
  - The marking feature of several graph layout tools
  - Google Maps
    Pin an anchor on a location of interest
Interaction

- Explore: Show me something else
- Examine a different subset of data cases

Examples:
- Panning function: the movement of a camera across a scene or scene movement while the camera stays still
Reconfigure: Show me a different arrangement

Examples:
  - Alignment

LineUp
Visual Analysis of Multi-Attribute Rankings

Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister and Marc Streit

http://caleydo.org/tools/lineup/
Interaction

- Encode: Show a different representation
- Examples:
  - Change visual variables: Color, Size, Orientation, Font, Form

Node-link diagram
https://bl.ocks.org/mbostock/4062045

Adjacency matrix
https://bost.ocks.org/mike/miserables/
Interaction

- Abstract/Elaborate: Show more or less detail
- Examples:
  - Table Lens

https://neo4j.com/

https://stats.oecd.org/OECDregionalstatistics/#story=0
Interaction

- Filter: Show something with some constraints/conditions
- Enable users to change the set of data items being presented based on some specific conditions.
- Examples:
  - Dynamic query controls

http://www.babynamewizard.com/voyager
Connect: Show me related/connected elements

Examples:

- Brushing and linking
Interaction

- Create: Generate the drawings
- Example:
  - Build a relationship graph

http://bl.ocks.org/benzguo/4370043
Factors Affecting Visual Graph Analytics

- Data Types & Information Retrieval
- Graph Layout
- Interaction
- Perception and Cognition
Perception and Cognition

- Mental Map Model
  - Circular Ordering
  - Proximity Relations
  - Topology Preservation
Preserving Mental Maps

- little influence in terms of error rate and response time
- But more confident

Effectiveness of Mental Maps

Lin et al., 2011

[Archambault and Purchase, 2013]
Outline

- What is a Graph? What is a Network?
- Tasks in Visual Graph Analytics
- Factors Affecting Visual Graph Analytics
- A tutorial of Visual Graph Analytics Tool
Graph Databases Advantages

- Natural modeling of highly connected data.
- Special graph storage structure
- Efficient schema less graph algorithms.
- Support for query languages.
- Operators to query the graph structure.
Neo4j [Robinson et al. 2013]

- **Neo4j** [https://neo4j.com/](https://neo4j.com/)

- Labeled attribute **multigraph**.
- Nodes and edges can have **properties**.
- There are **no restrictions** on the number of edges between two nodes.
- **Loops** are allowed.
- Attributes are **single valued**.
- Different types of **indexes**: Nodes & relationships.
- Different types of **traversal strategies**.
- API for Java, Python

Cypher using relationship 'likes'

(a) -[:LIKES]-> (b)
Textbooks

- **Graph Theory**
  - Introduction to Graph Theory

- **Graph Drawing**
  - Graph Drawing – Algorithms for the Visualization of Graphs
  - Handbook of Graph Drawing and Visualization
  - Tutorial (talk slides)
Some Useful Tools & Websites

- GraphViz: http://graphviz.org
- Tulip: http://tulip.labri.fr/TulipDrupal/
- yFiles: https://www.yworks.com/products/yfiles-for-java-2.x
- GraphDrawing: http://graphdrawing.org
- Neo4j: https://neo4j.com/
- HypergraphDB: http://www.hypergraphdb.org
- Sparksee: http://www.sparsity-technologies.com/
Are you interested in network visualization?

Graph drawing is the algorithmic core of network visualization. In this course you learn about

- graph layout aesthetics and their optimization
- fundamental layout algorithms for different types of graphs (trees, planar, non-planar, ...)
- practical aspects as well as theoretical properties and computational limitations

Topics in this course combine the areas of algorithms, visualization, and graph theory.
References


Q & A

Personal website:

https://www.cg.tuwien.ac.at/staff/HsiangYunWu.html