Graph Visualization

Rajmonda Caceres

CSC/SDS235: Visual Analytics, Smith College

March 18, 2021

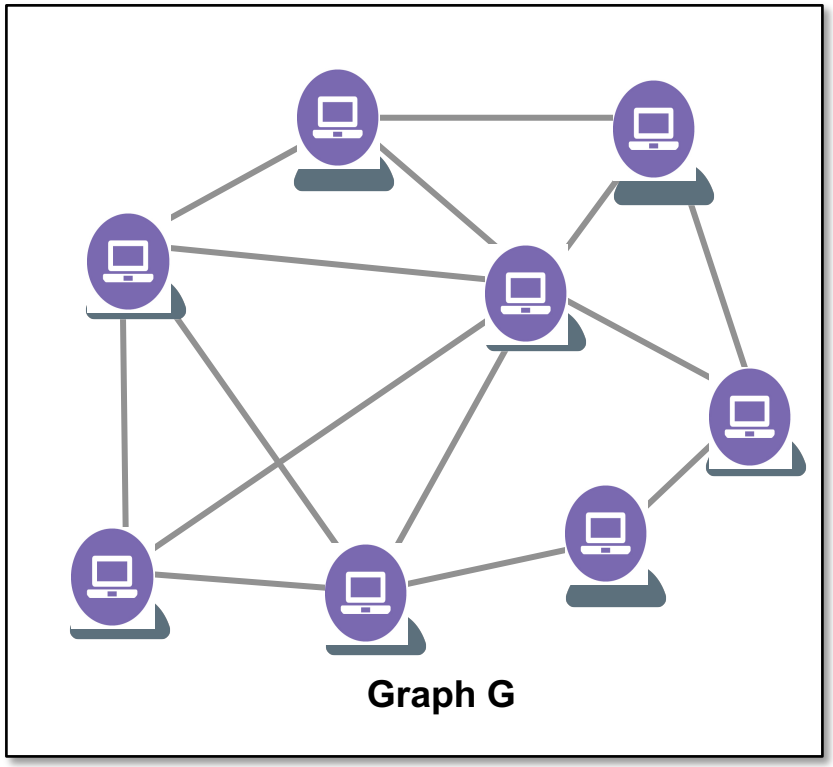
LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Outline

- What is a graph? Why do we visualize graphs?
- Overview of basic graph visualization components
- Challenges with visualizing large graphs
- Few solutions to address scale and complexity of large graphs

Graphs 101

A graph (at times also referred to as a network) is a mathematical abstraction representing a collection of entities and their interactions



G = (V, E)
- **V = vertices (entities)**
- **E = edges (relationships)**

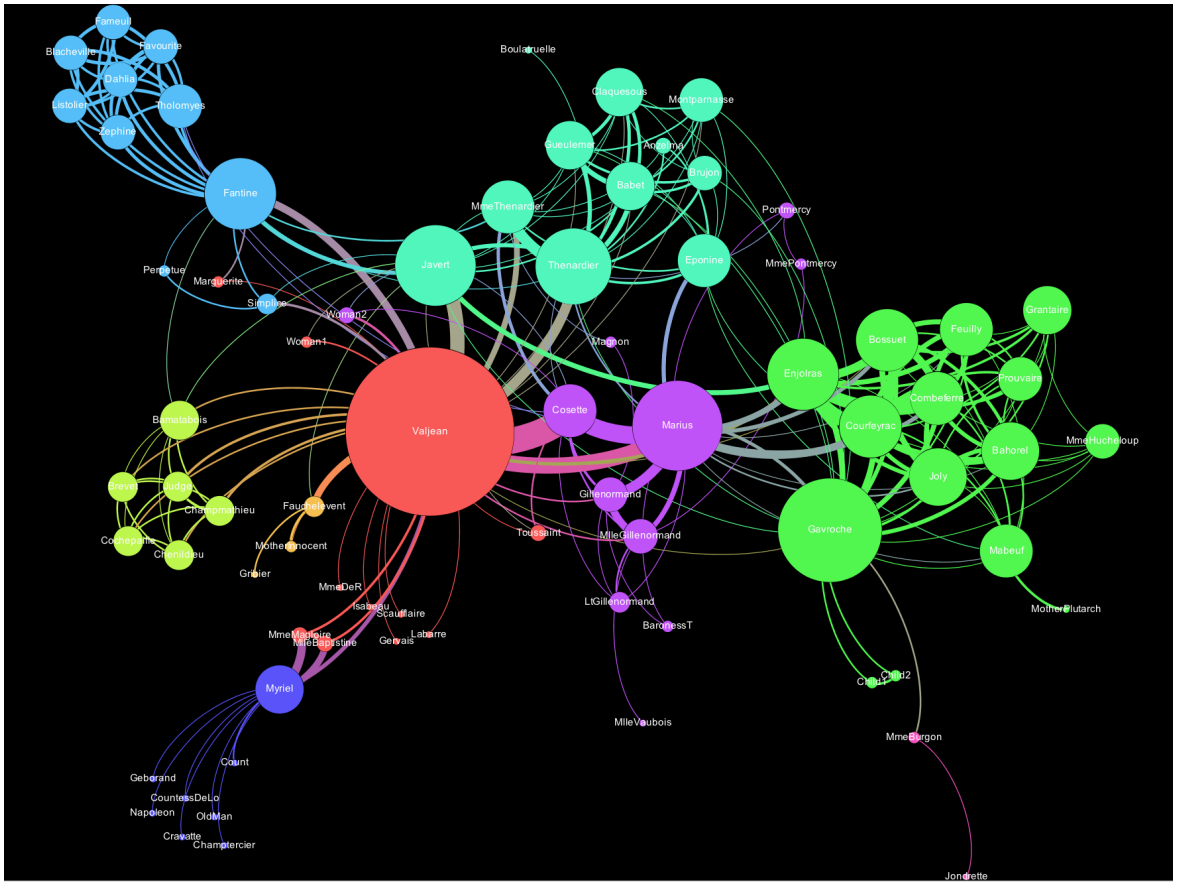
- **V = people**
- **E = social relationships**

- **V = machines**
- **E = network packet flow**

Introduction to Visualization

Visualization (def) :

Using visual representations of abstract data to reinforce human cognition



Ultimate Goal: 1. Lead to insights 2. Aid analysis and interpretation

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

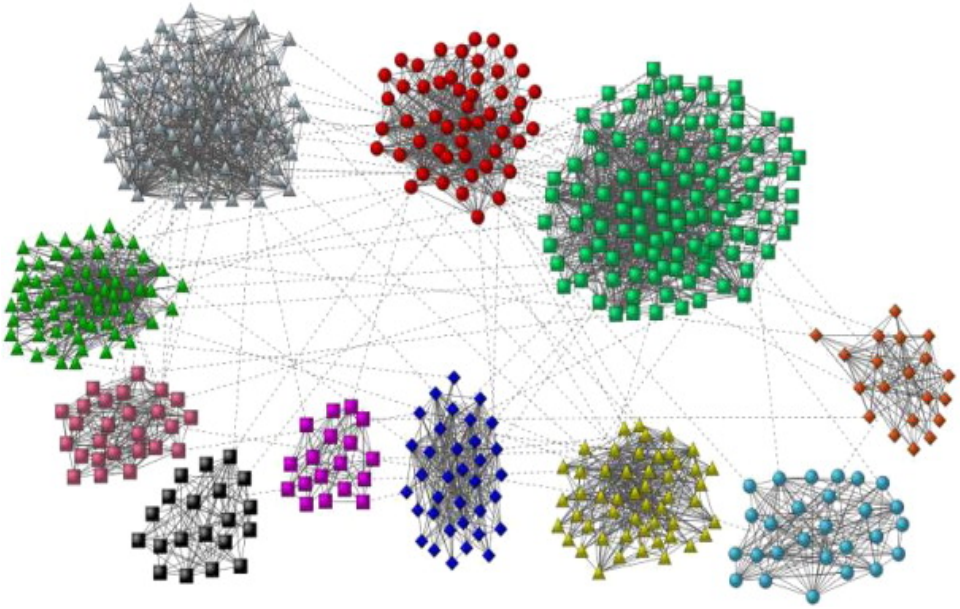


Figure 1

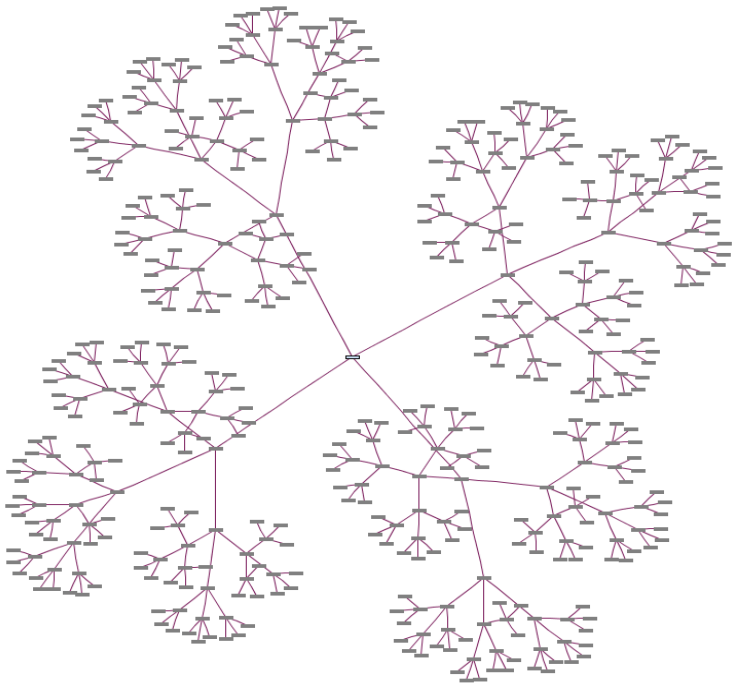


Figure 2

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

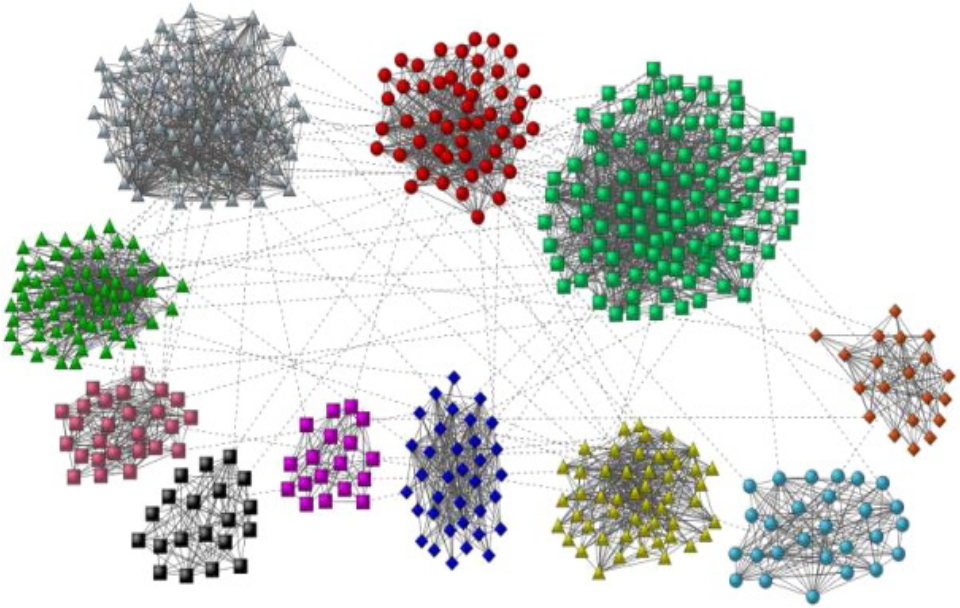


Figure 1: Communities

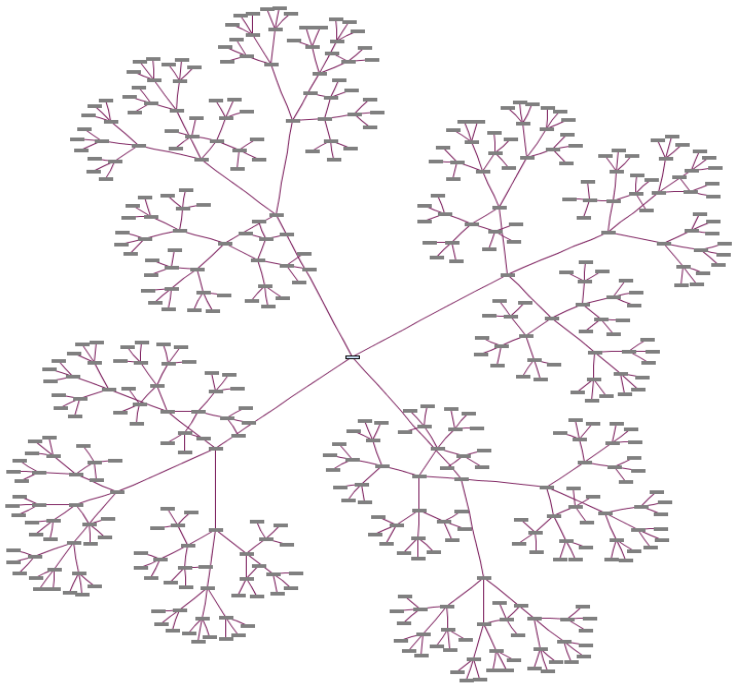


Figure 2: Hierarchy

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

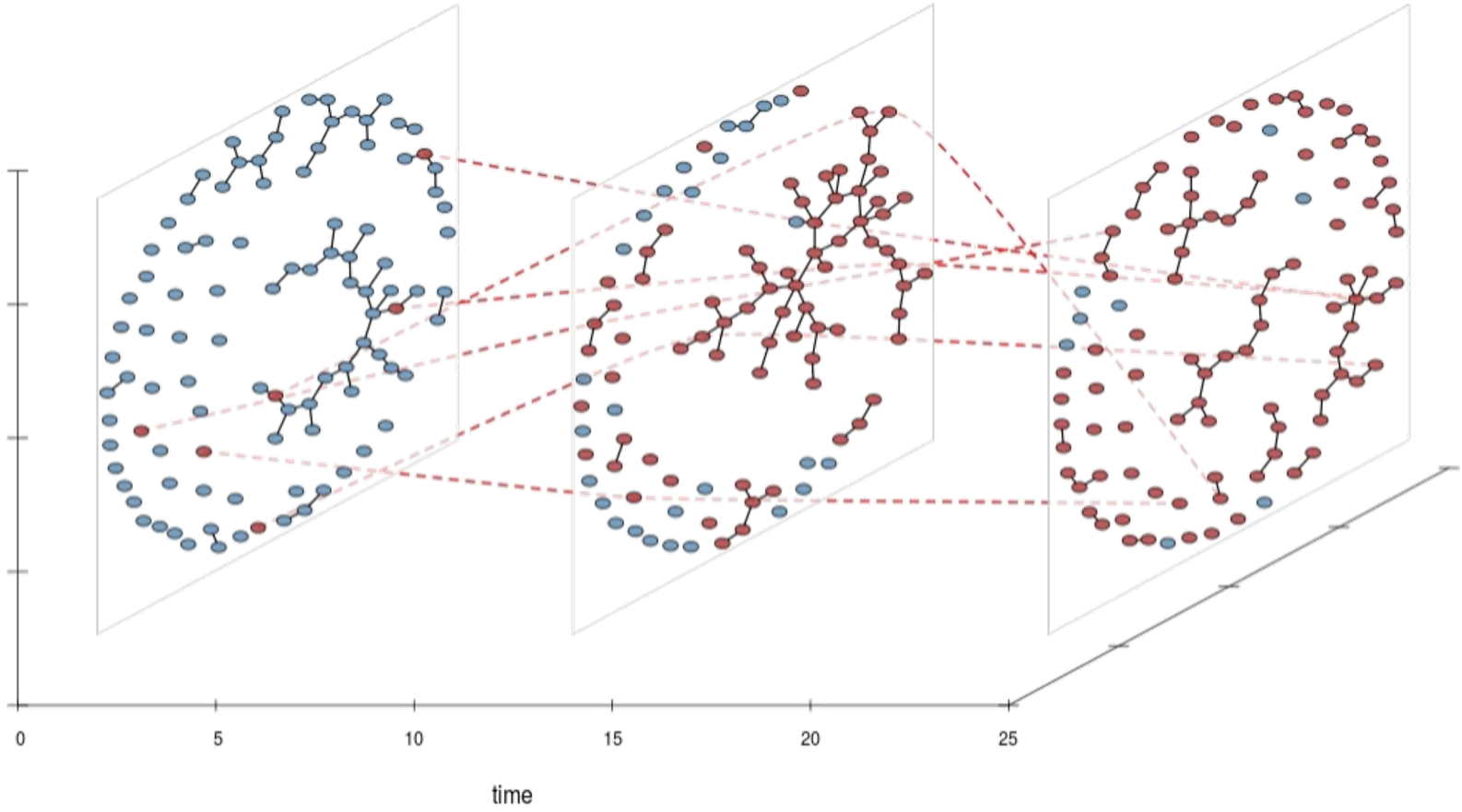


Figure 3

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

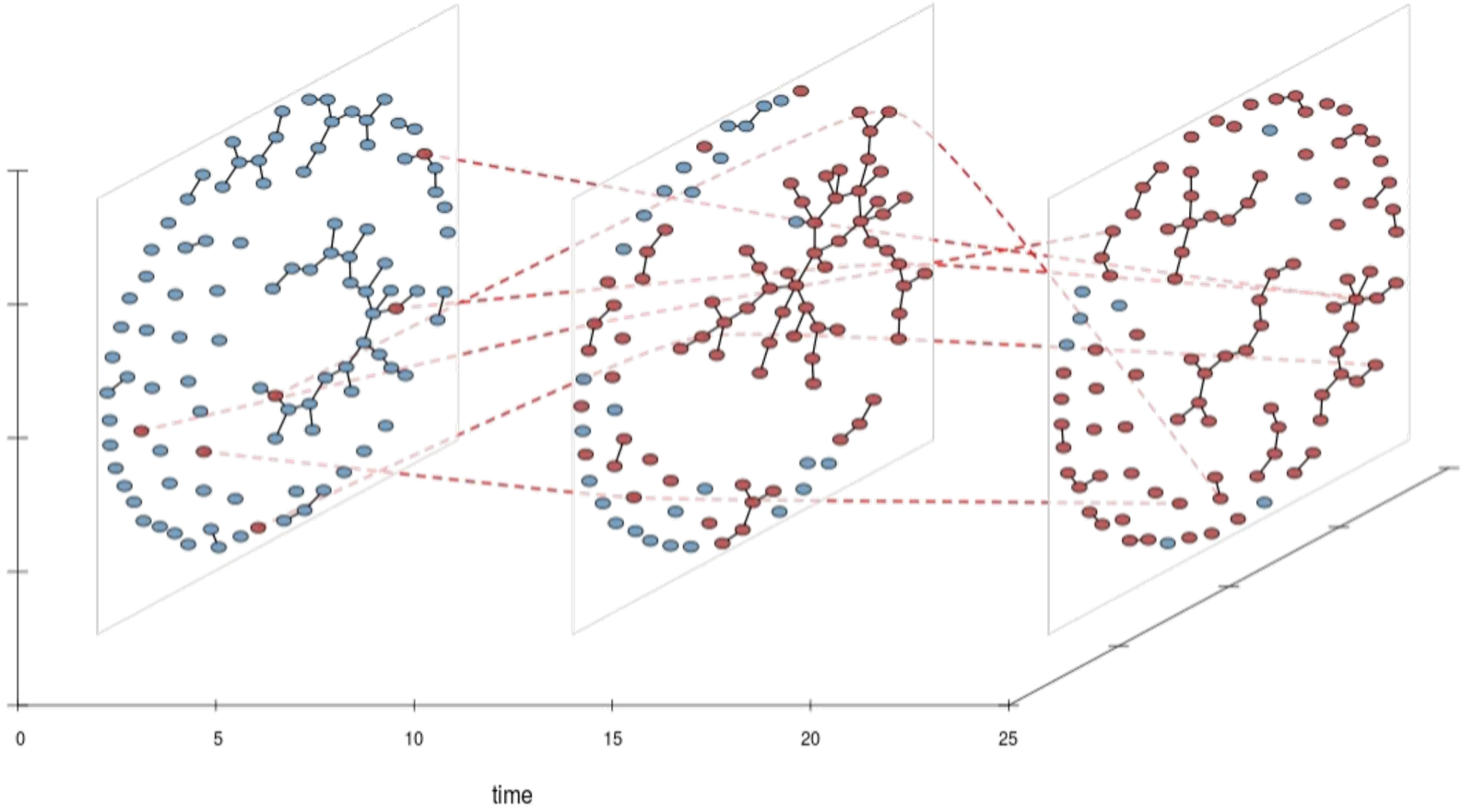


Figure 3: Dynamics, changes over time

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

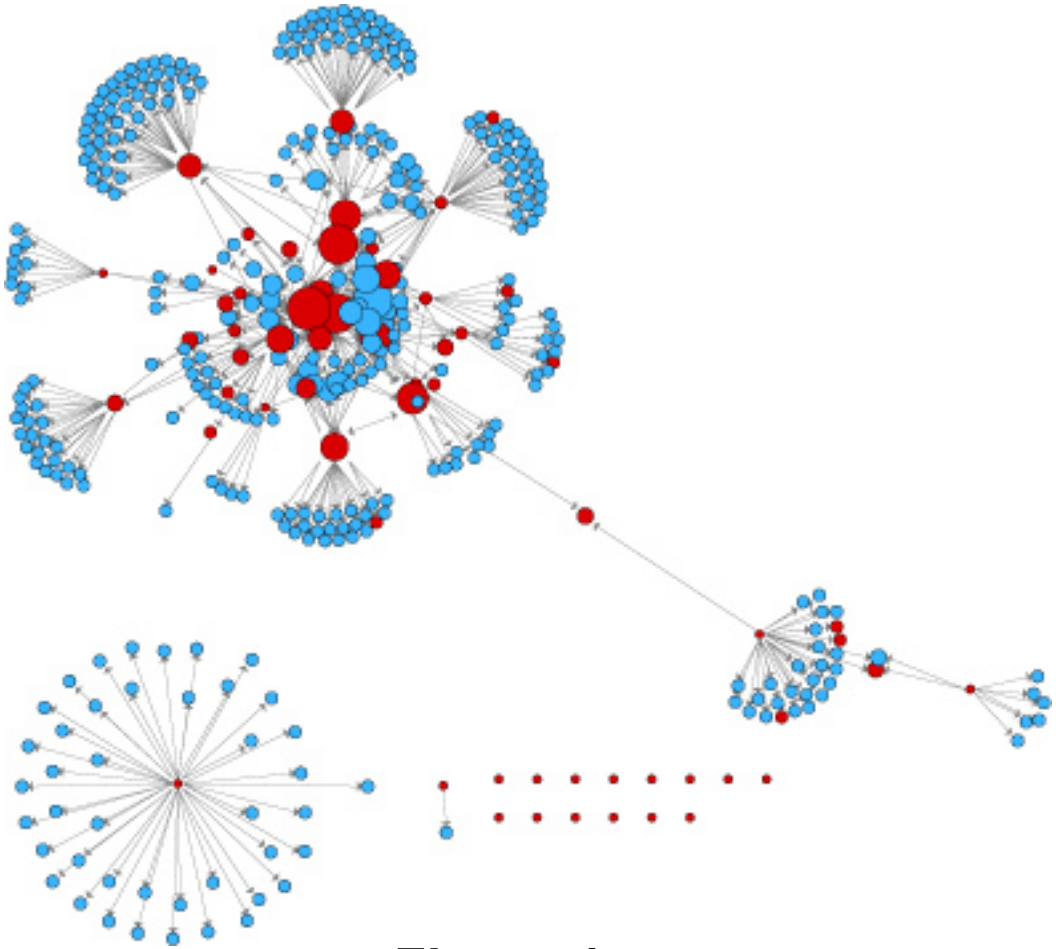


Figure 4

The Power of Visual Pattern Recognition

- **A good visual representation can highlight important features**

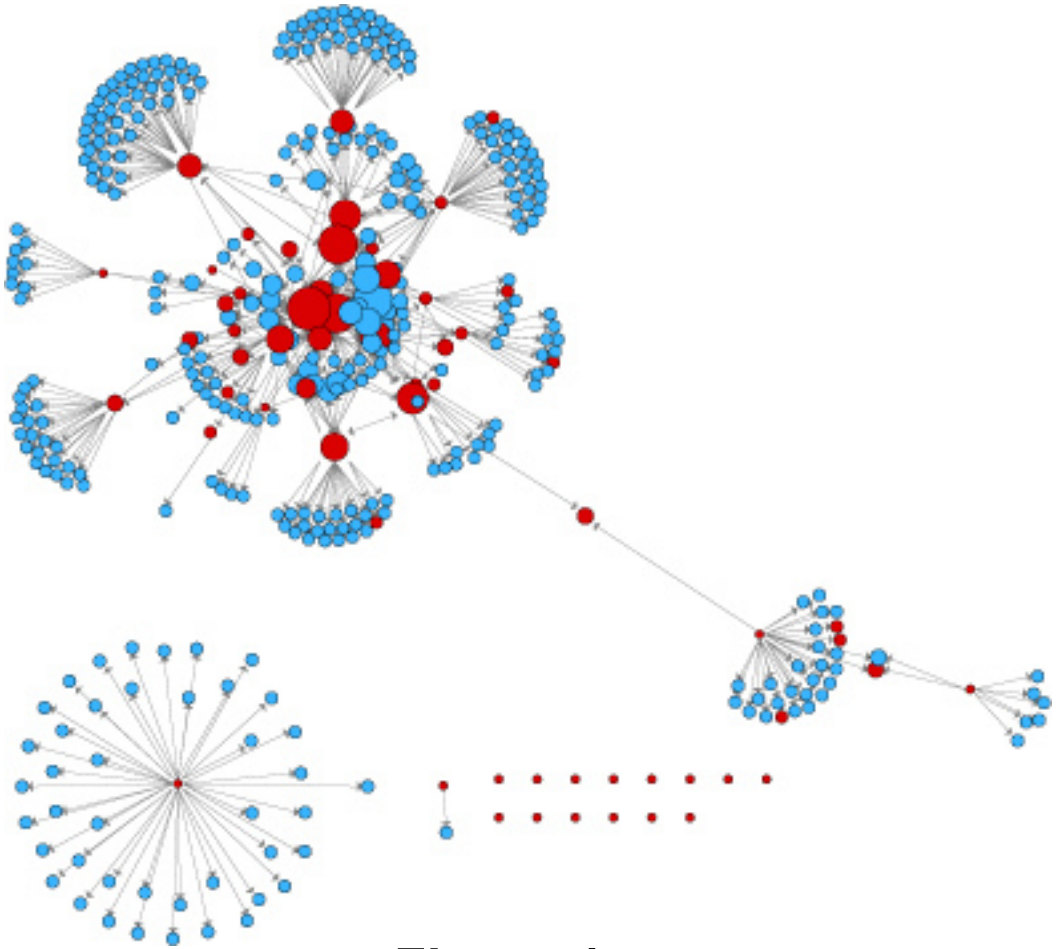
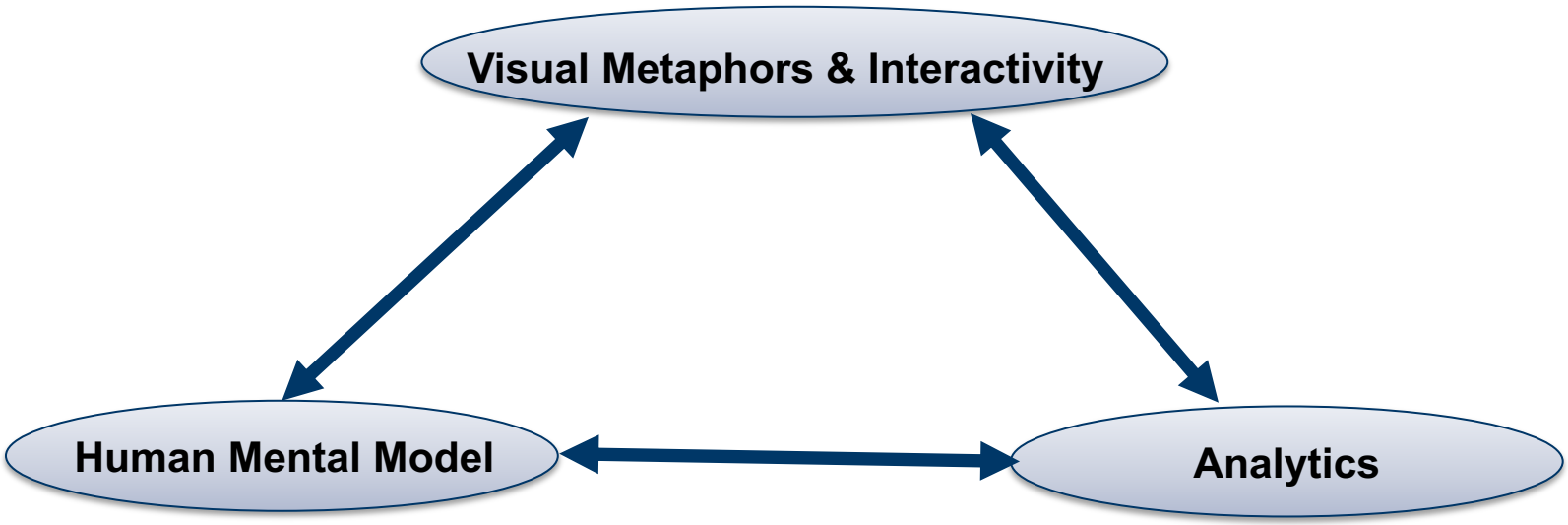


Figure 4: Outliers, anomalies

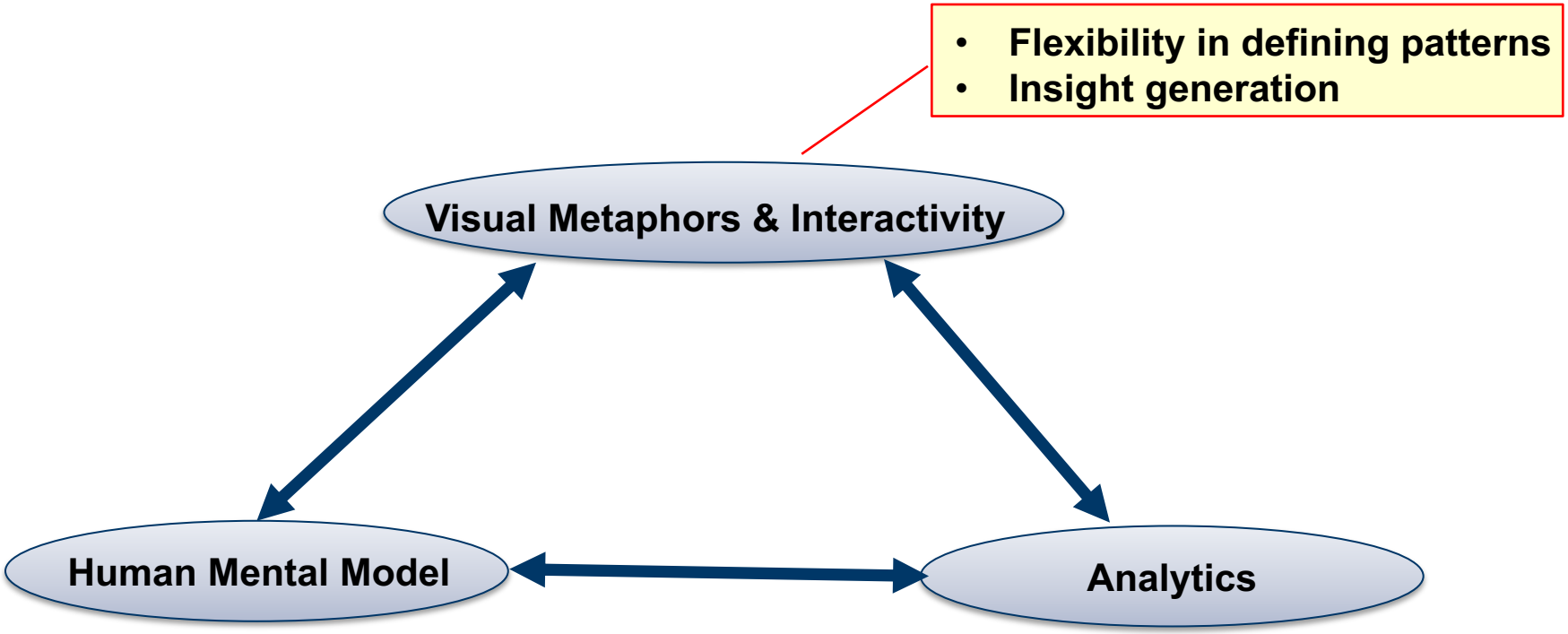
Big Picture: Effective Visualization

- **Three critical components that need to be seamlessly integrated**



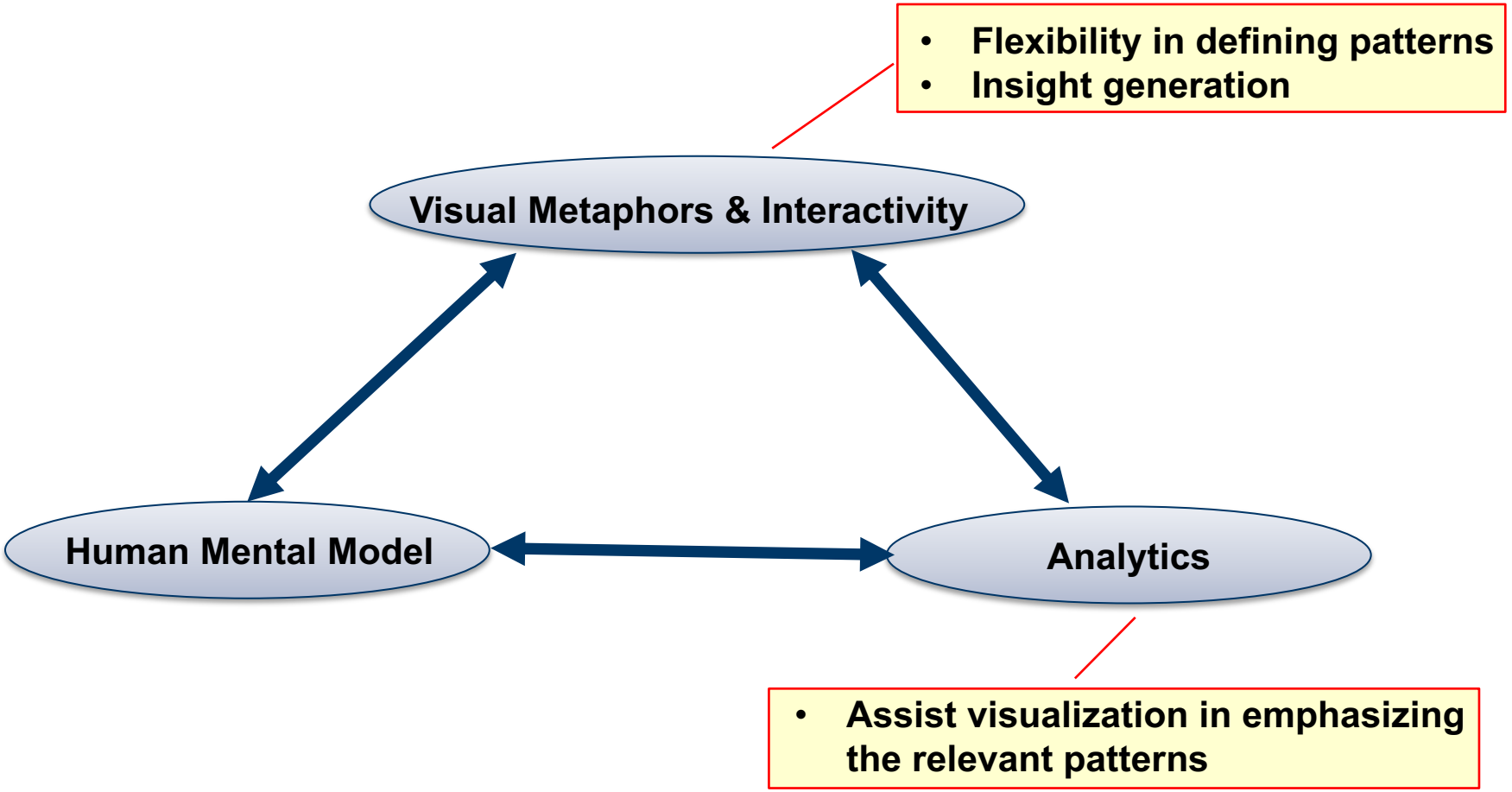
Big Picture: Effective Visualization

- **Three critical components that need to be seamlessly integrated**



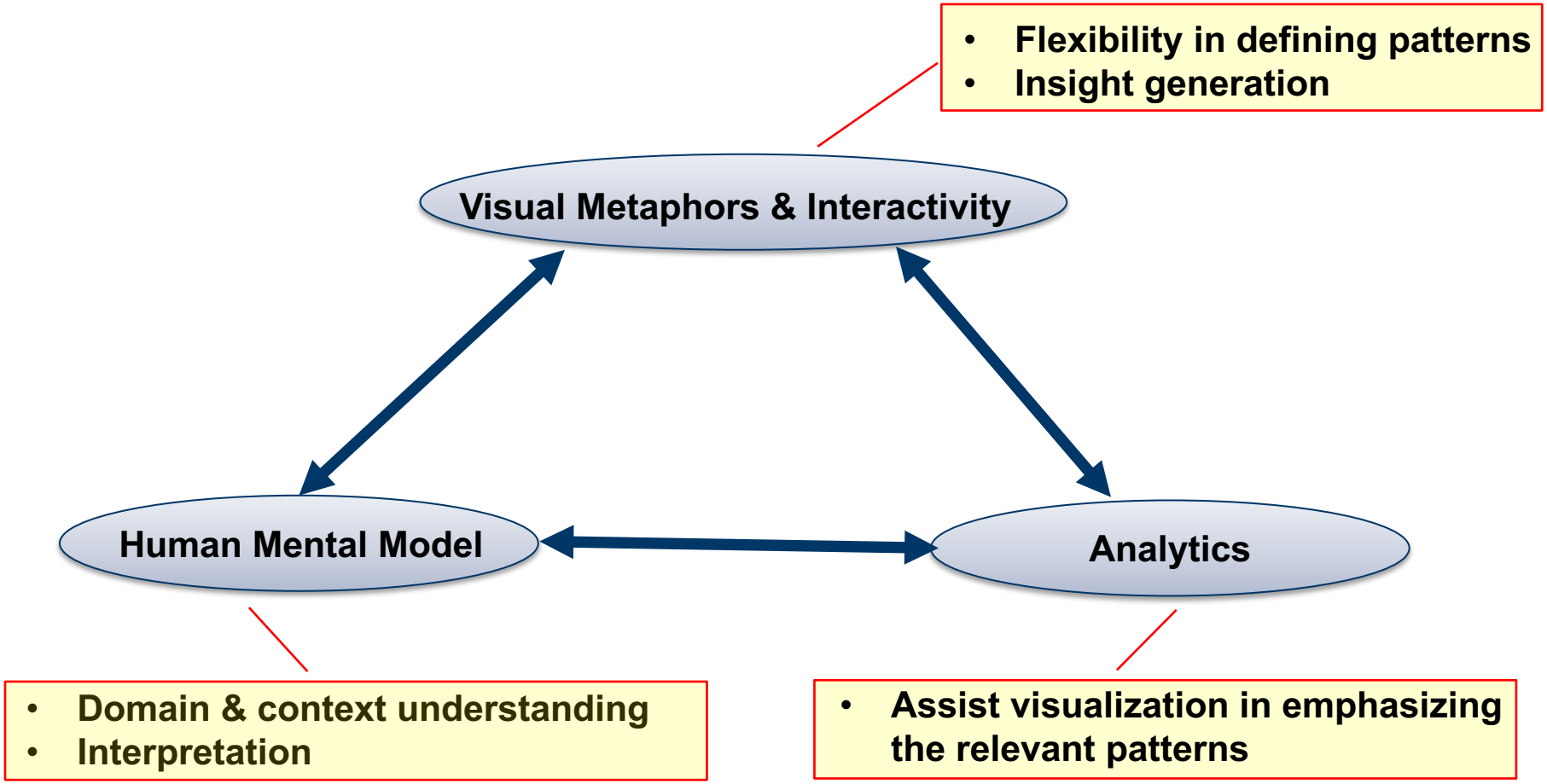
Big Picture: Effective Visualization

- **Three critical components that need to be seamlessly integrated**



Big Picture: Effective Visualization

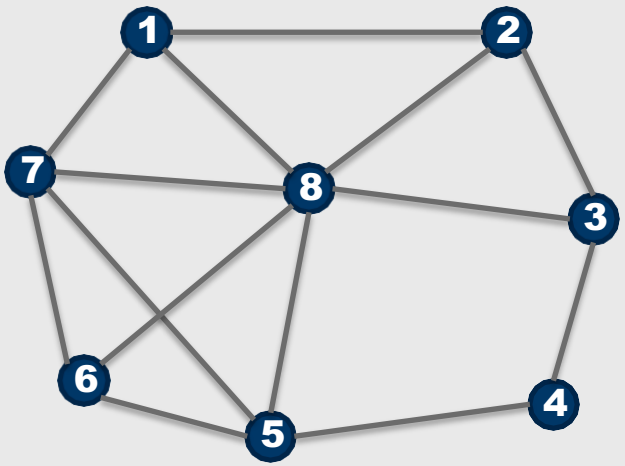
- **Three critical components that need to be seamlessly integrated**



Outline

- What is a graph? Why do we visualize graphs?
- Overview of basic graph visualization components
- Challenges with visualizing large graphs
- Few solutions to address scale and complexity of large graphs

Two Traditional Ways for Visualizing Graphs



Node-link Representation G

	1	2	3	4	5	6	7	8
1	0	1	0	0	0	0	1	1
2	1	0	1	0	0	0	0	1
3	0	1	0	1	0	0	0	1
4	0	0	1	0	1	0	0	0
5	0	0	0	1	0	1	1	1
6	0	0	0	0	1	0	1	1
7	1	0	0	0	1	1	0	1
8	1	1	1	0	1	1	1	0

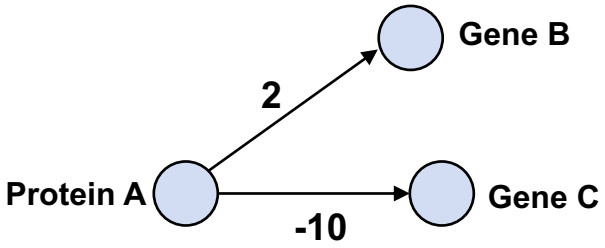
Matrix Representation A

- $G = (V, E)$
- V : **Vertices (entities)**
- E : **Edges (relationships)**

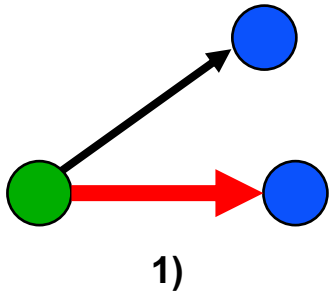
$A(i, j) \neq 0$ **if edge exists between vertex** i **and vertex** j

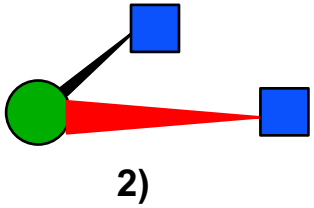
Network Information and Visual Elements

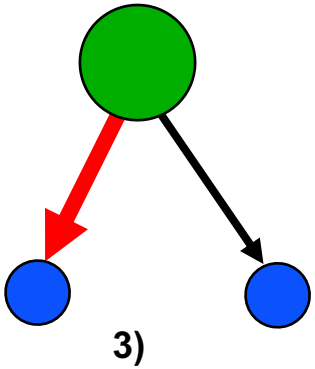
- **Shape, size, color, texture, position**



	A	B	C
A	0	2	-10
B	0	0	0
C	0	0	0

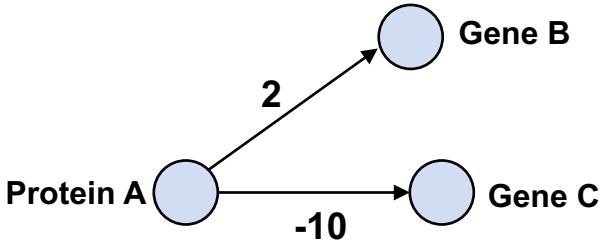




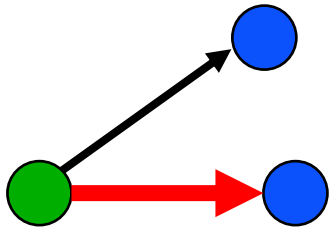


Network Information and Visual Elements

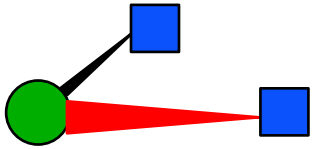
- **Shape, size, color, texture, position**



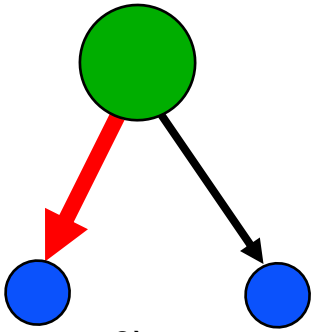
	A	B	C
A	0	2	-10
B	0	0	0
C	0	0	0



1)



2)



3)

	A	B	C
A	0	2	-10
B	0	0	0
C	0	0	0

4)

	B	A	C
B	0	0	0
A	2	0	-10
C	0	0	0

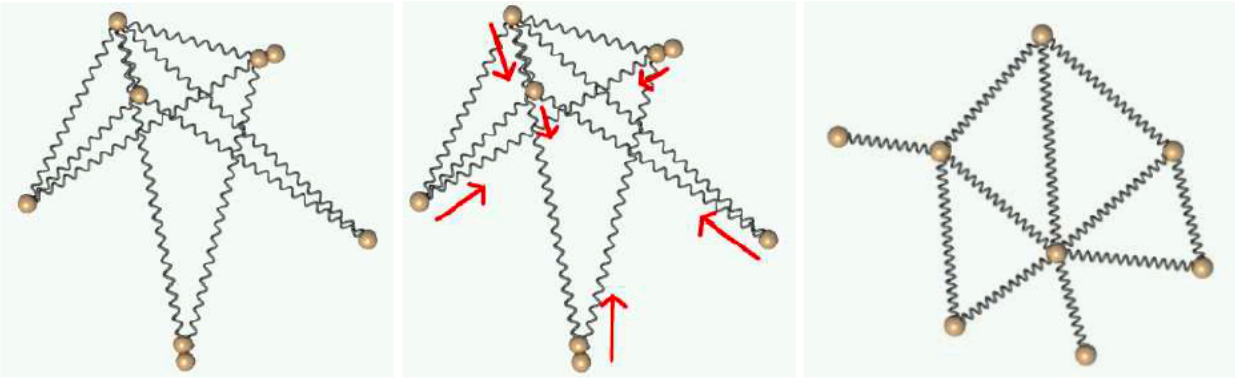
5)

	B	A	C
B	0	0	0
A	2	0	-10
C	0	0	0

6)

Force-directed Layout Algorithms

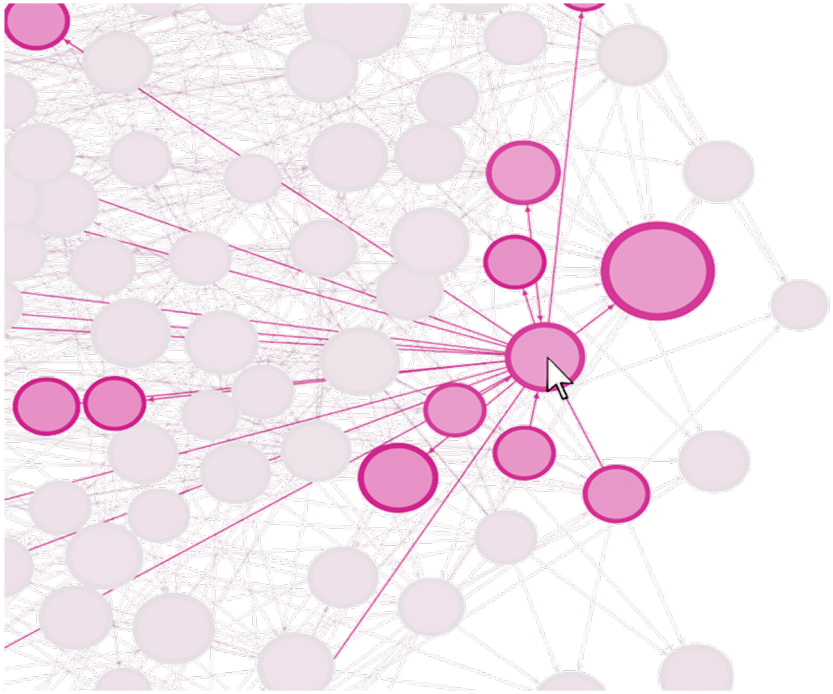
- **Fundamental optimization goal: minimize number of edge crossings**
- **Assign a spring-like force values to nodes on the graph**
  - **Edges are assigned attractive forces, Nodes are assigned repulsive forces**
- **Simulate a dynamic system of this system and let it reach equilibrium**



- **Different variants consider different ideal spring(edge) lengths and how much to move nodes around:**
  - **Fruchterman-Reingold: bound parameter on movement distance**
  - **Kamada-Kawai: proportional to the shortest path distance**

Network Information and Visual Elements

- **Interaction types**
  - **Zoom**
  - **Search and locate**
  - **Filter/ highlight**
  - **Details on demand**
  - **Linked views**
  - **Distortion**



Large Variety of Graph Visualization Tools

WebWeb

Library

Gephi

Tool

Cytoscape

Tool

sigma.js

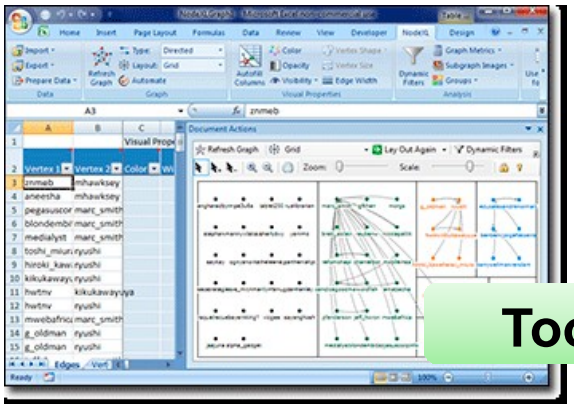
Web network visualization made easy

Library

Data-Driven Documents

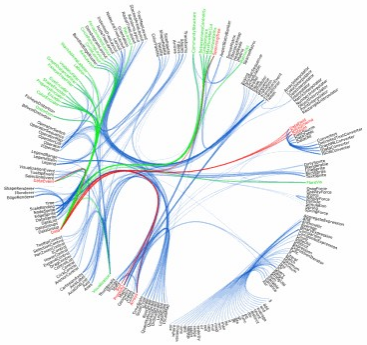
Library

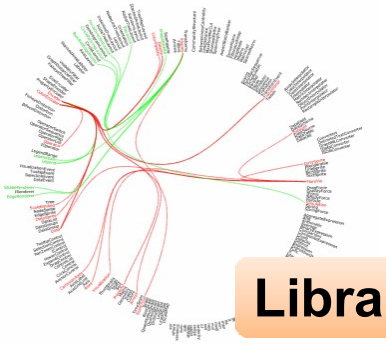
NodeXL



Tool

flare DATA VISUALIZATION FOR THE WEB





Library

Large Variety of Graph Visualization Tools

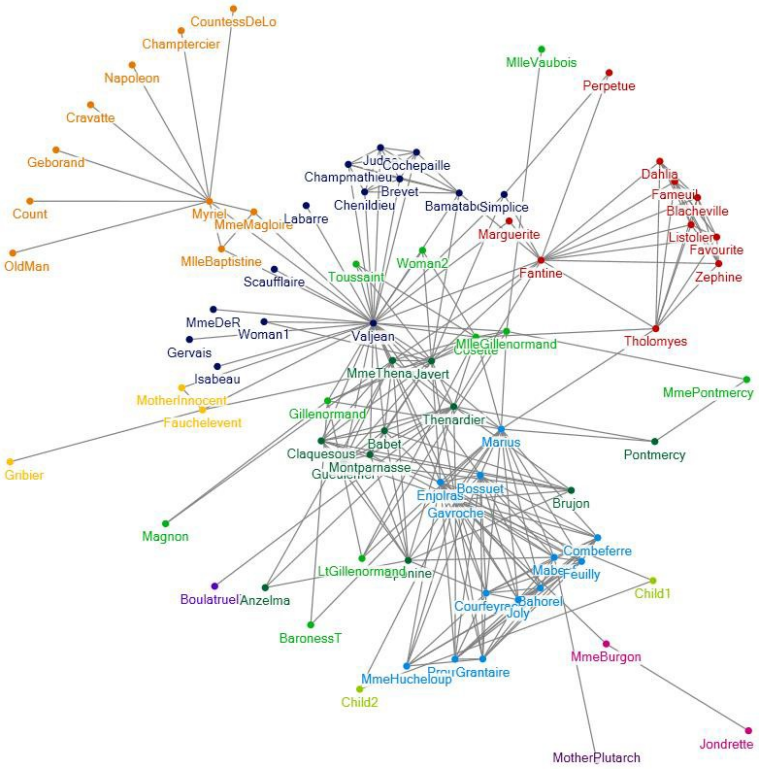
	Functionality	Scalability	Plug-in development	Application-specific Features
Gephi	Full software	Better at large scales	Easy	General
Cytoscape	Full software	Better at large scales	Difficult	Biology
NodeXL	Full software	Slow at large scales	Difficult	General
Flare	Interactive	Slow at large scales	Easy	General
Sigma.js	Interactive, web	Slow at large scales	Easy	General
D3.js	Interactive, web	Slow at large scale	Easy	General
Webweb	Interactive, web	Slow at large scale	Easy	General

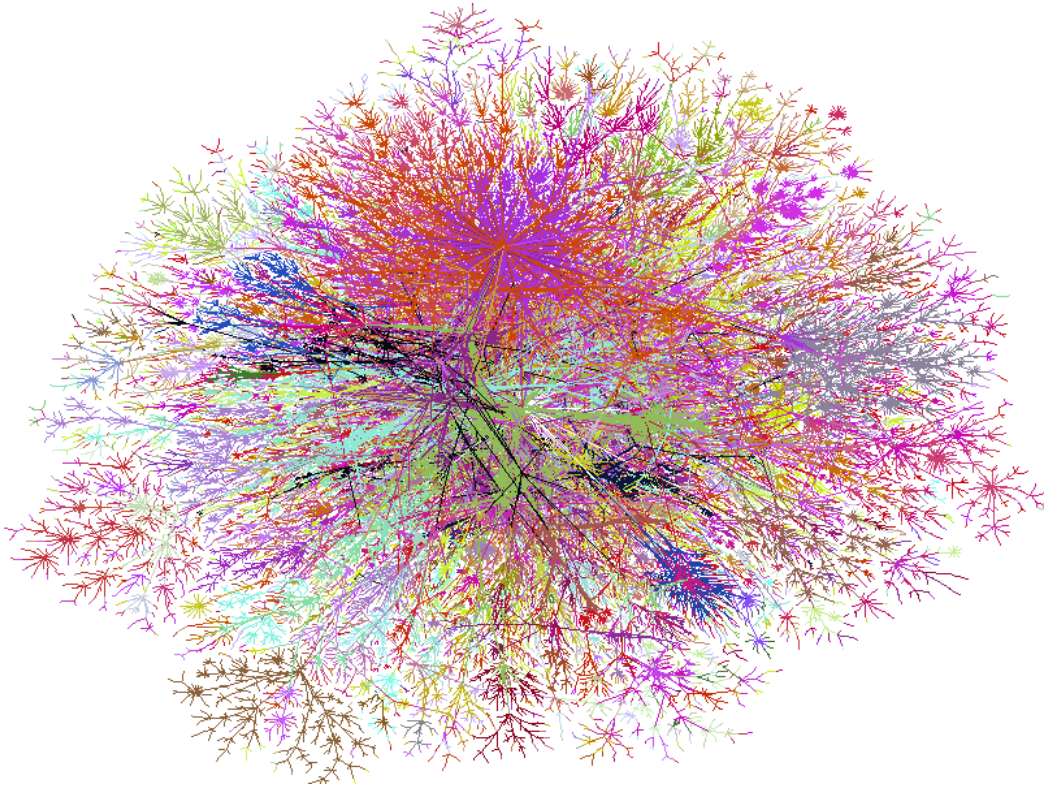
Outline

- What is a graph? Why do we visualize graphs?
- Overview of basic graph visualization components
- Challenges with visualizing large graphs
- Few solutions to address scale and complexity of large graphs

Challenge 1: Resolution

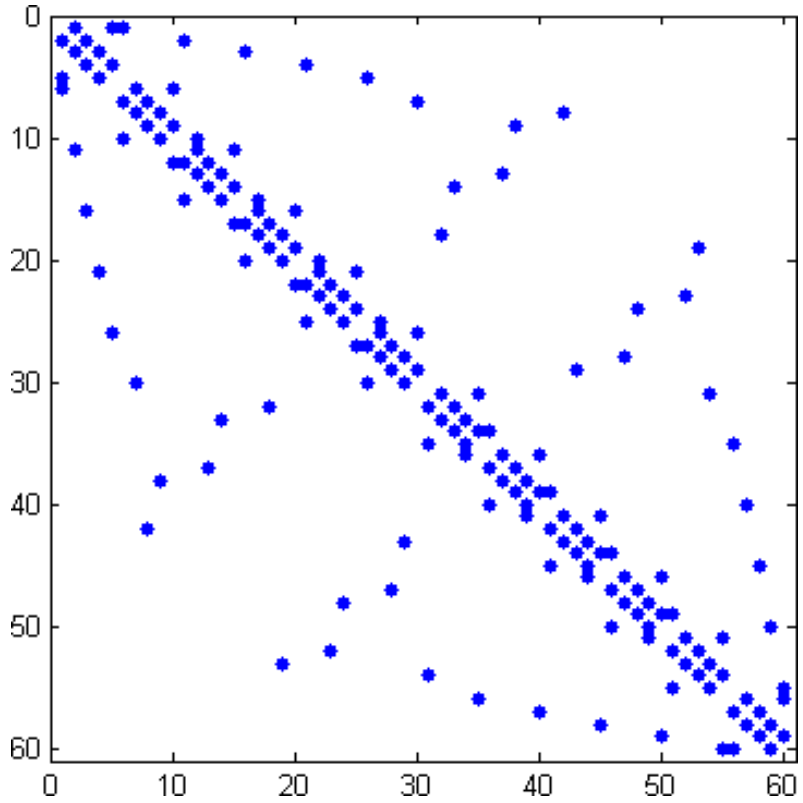
- **How many nodes can we draw with fidelity?**

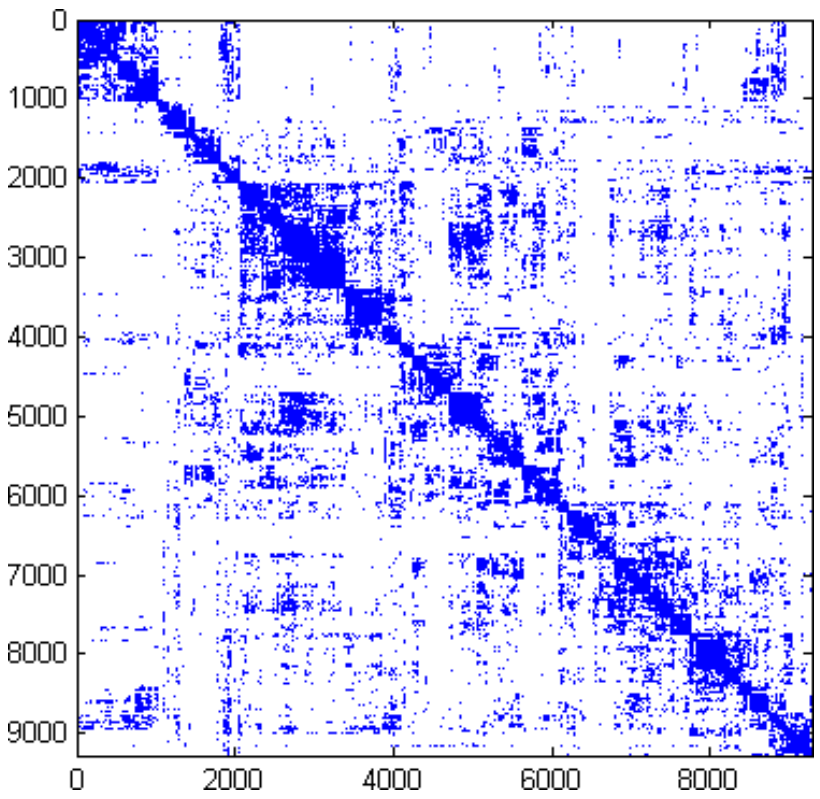




Challenge 1: Resolution

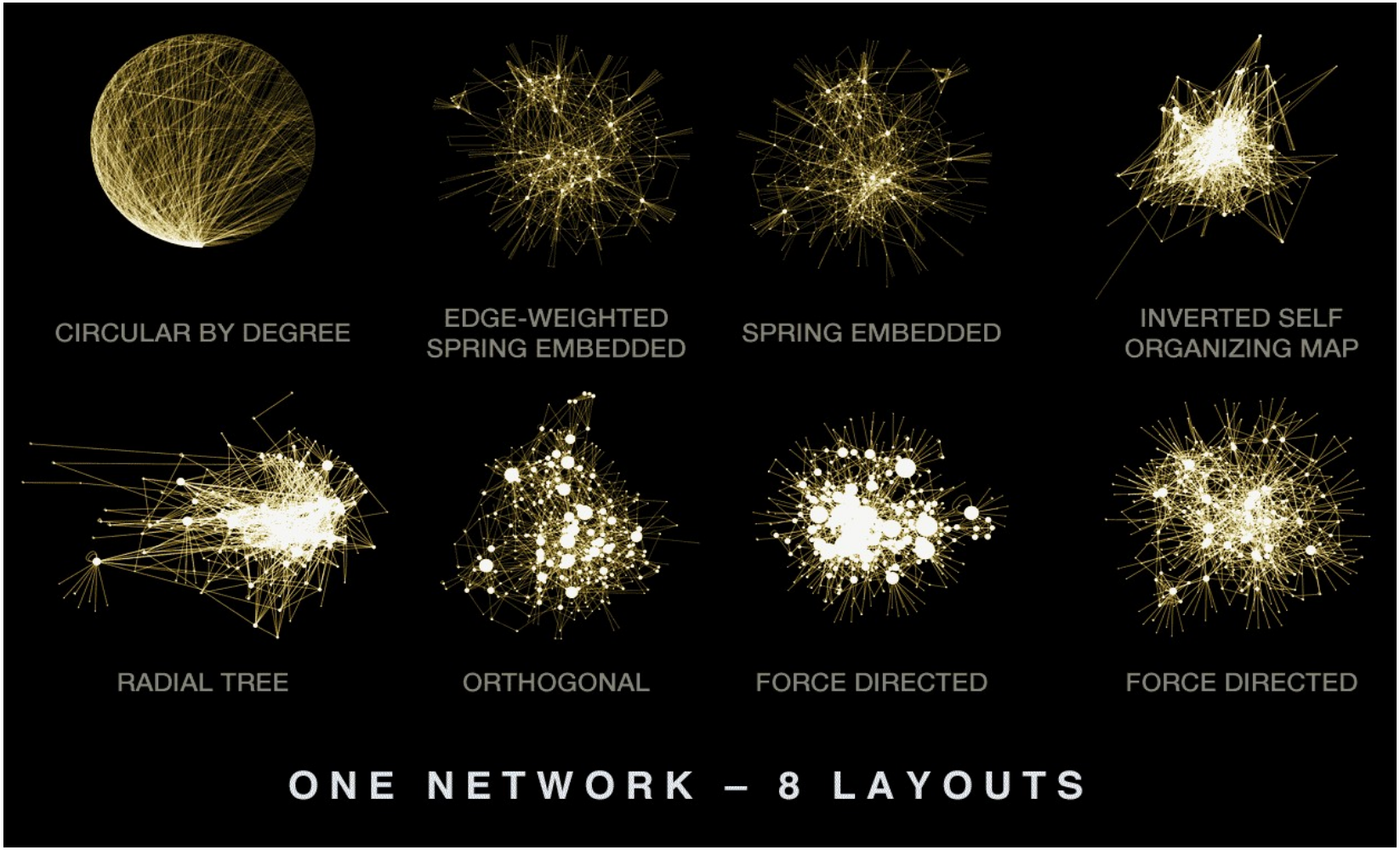
- **How many nodes can we draw with fidelity?**





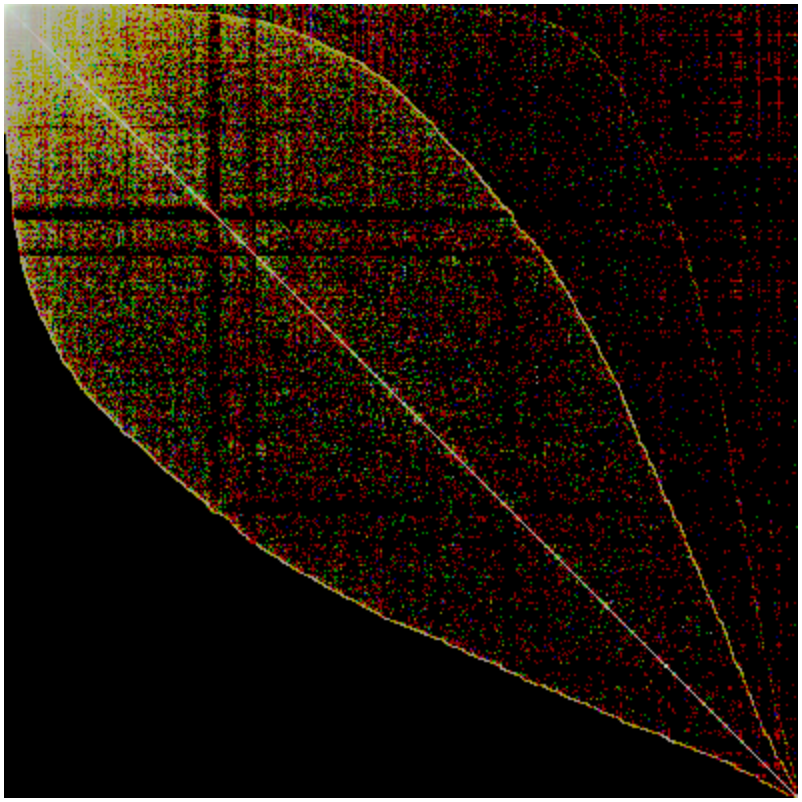
Challenge 2: Layout

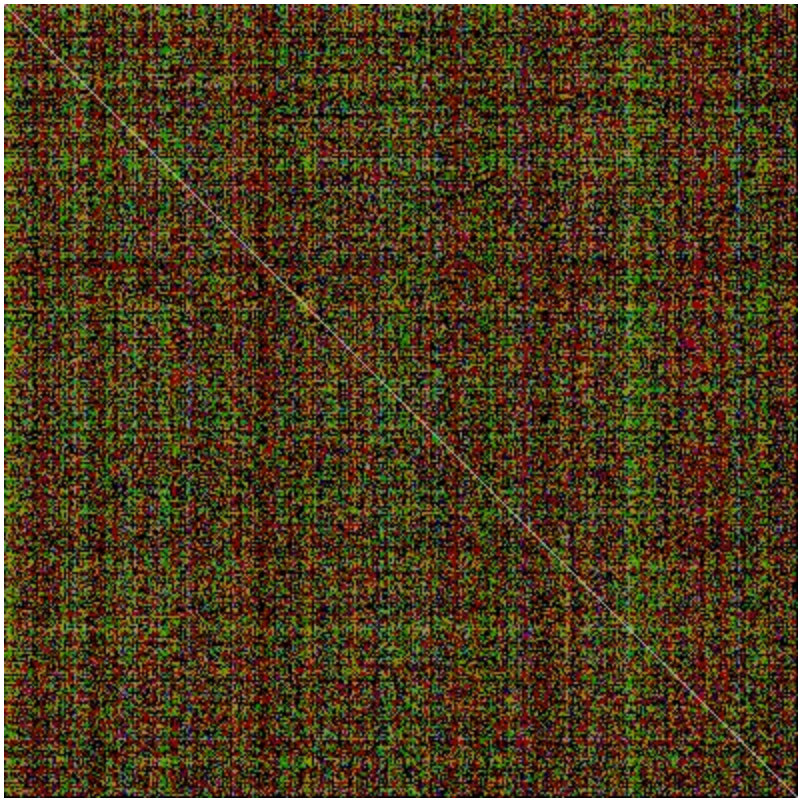
- **How do we spatially organize the nodes?**



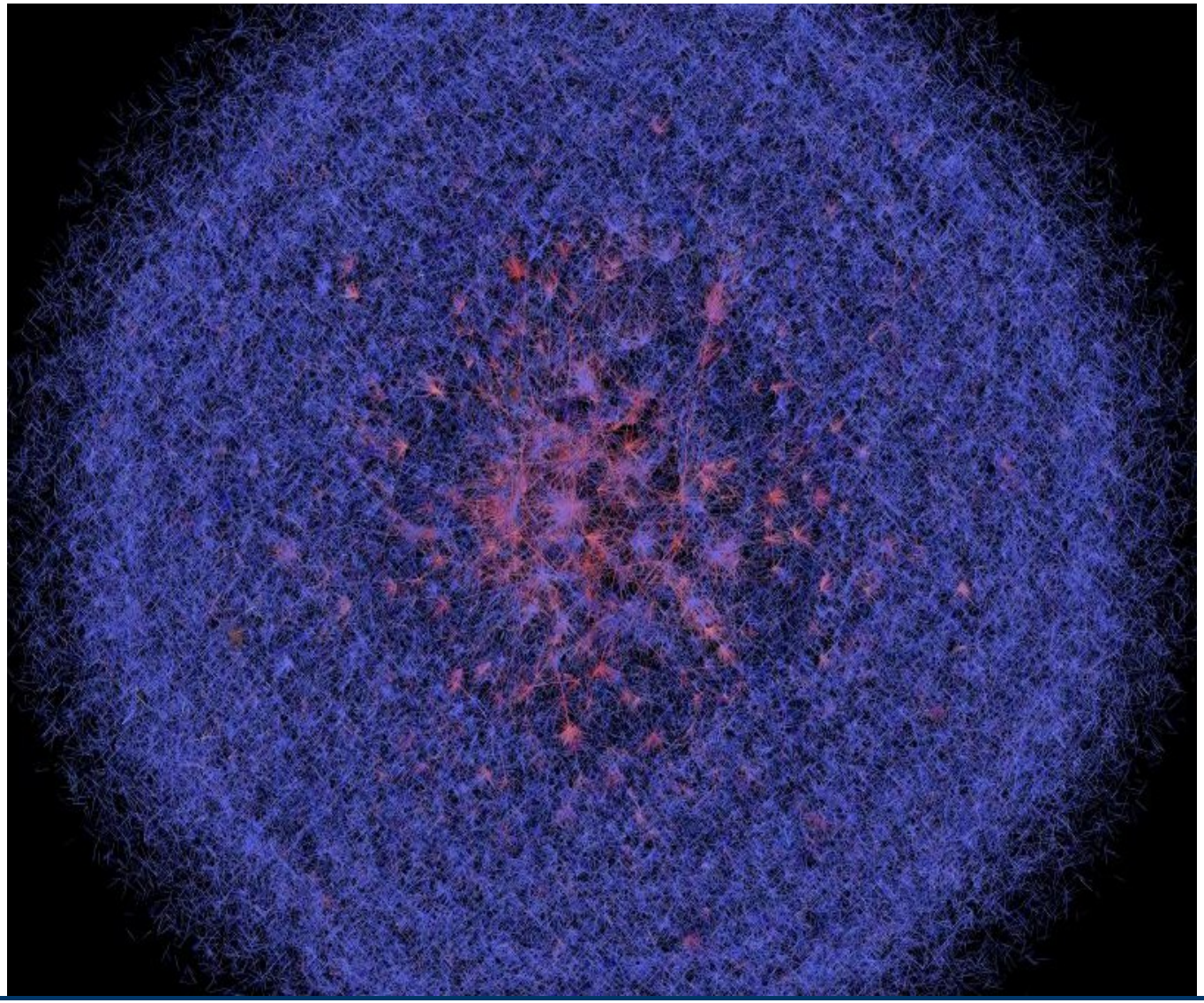
Challenge 2: Layout

- **How do we spatially organize the nodes?**

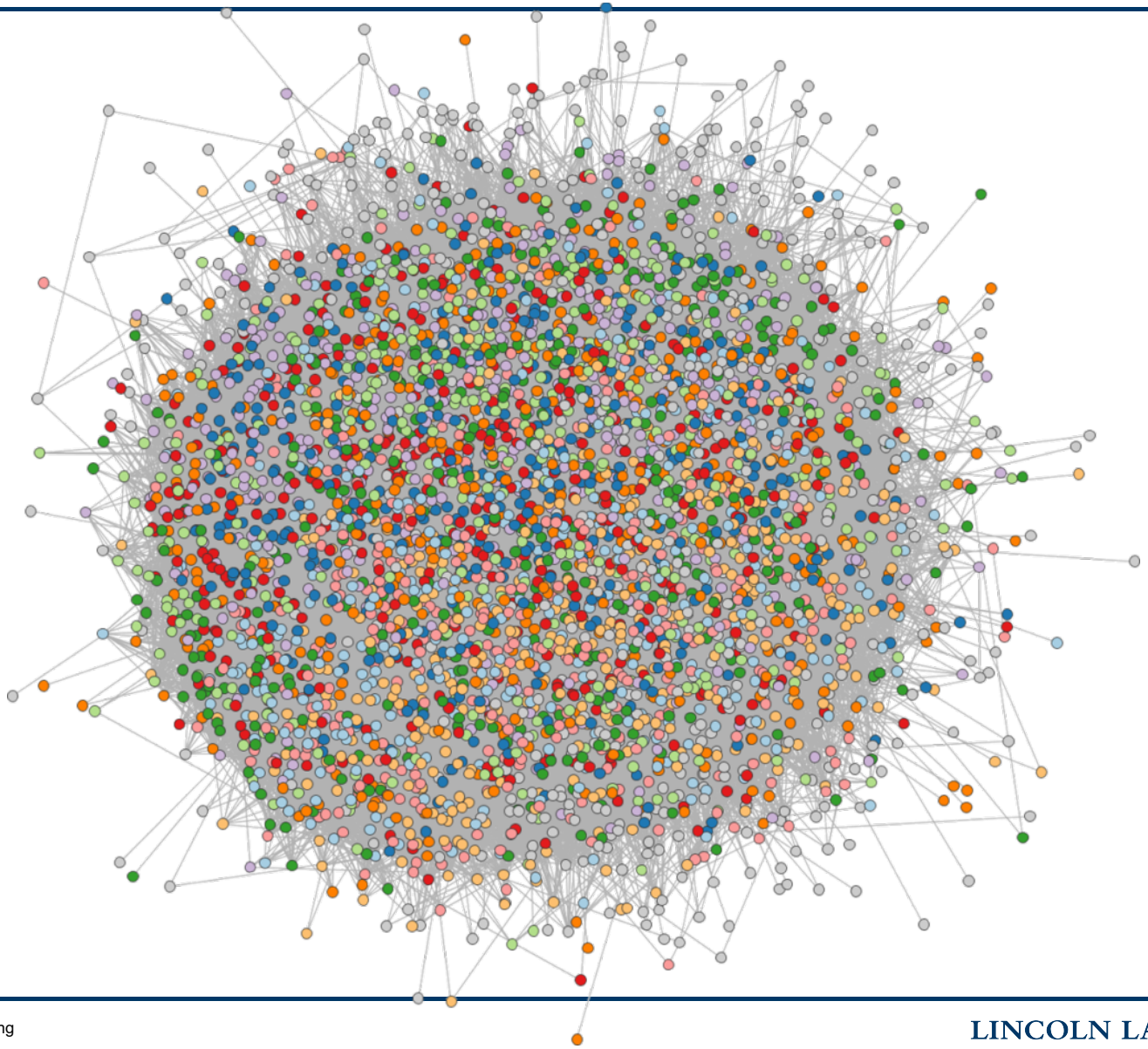




The Hairball Graph



The Hairball Graph



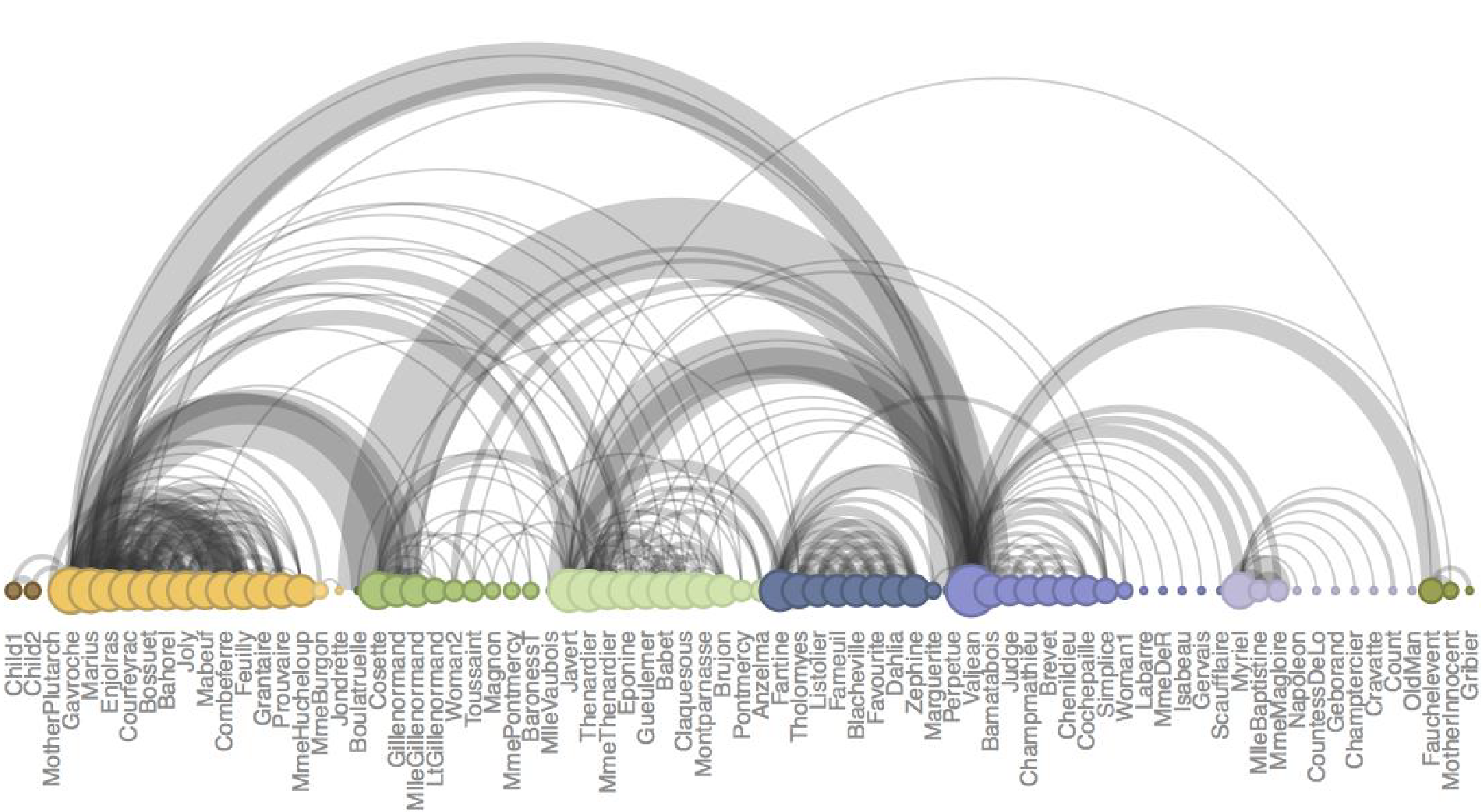
Outline

- What is a graph? Why do we visualize graphs?
- Overview of basic graph visualization components
- Challenges with visualizing large graphs
- Few solutions to address scale and complexity of large graphs

Untangling Hairball Graphs: Some Solutions

- Impose structure on layout
- Edge bundling
- Multi-view representation
- Sampling
- …

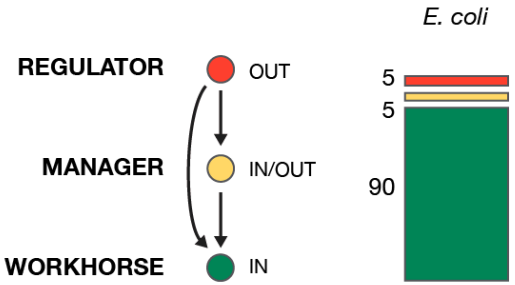
Arc Diagrams

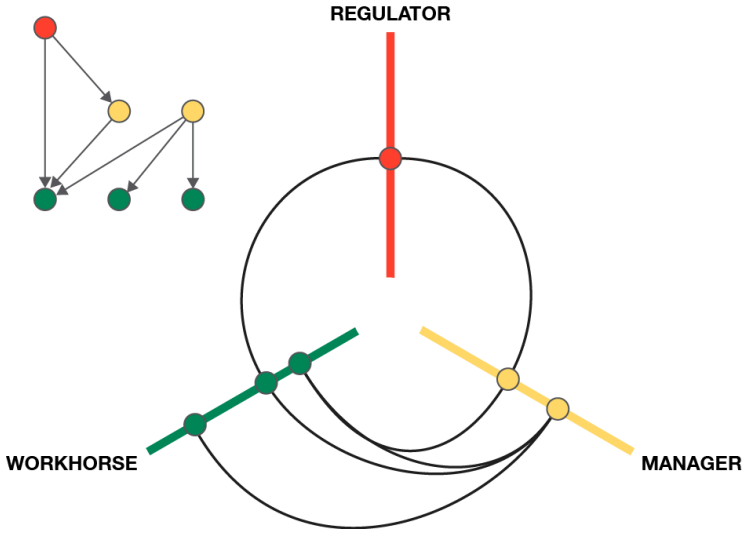


R package: arcdiagram : https://github.com/gastonstat/arcdiagram

Hive Plot
E-coli Gene Network Illustration

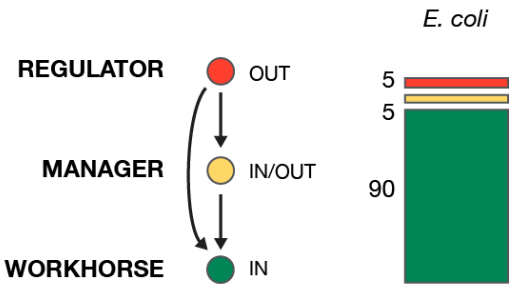
Node Classification

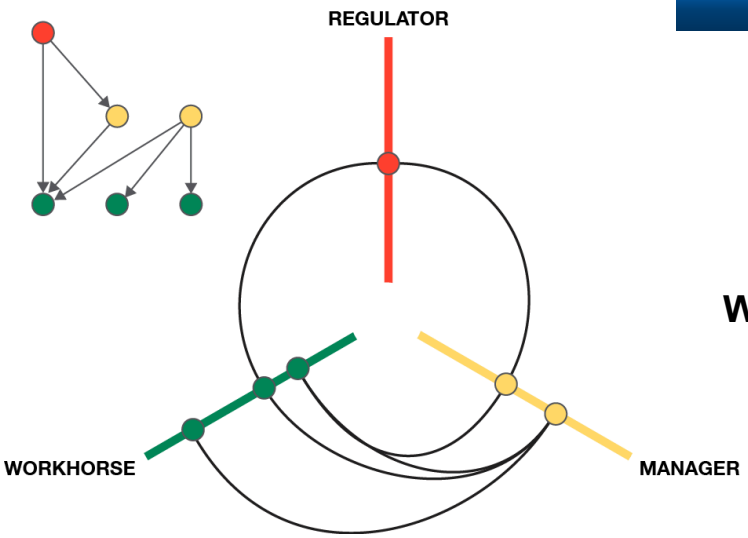


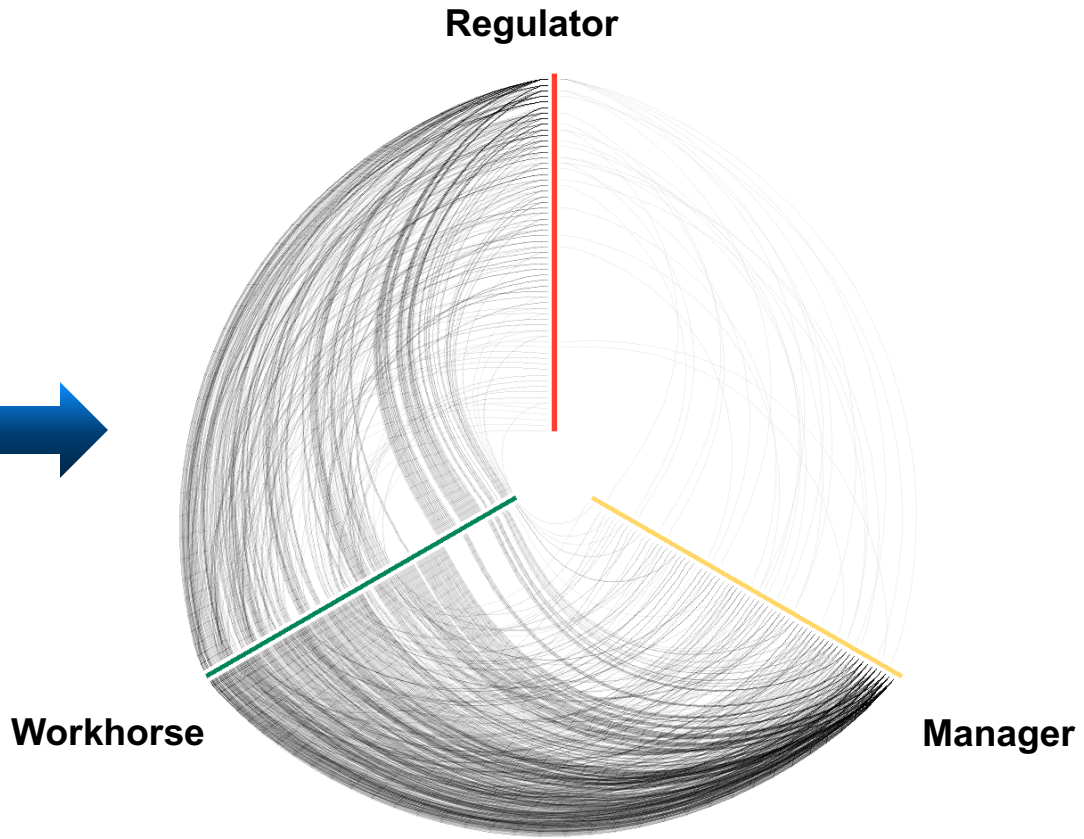


Hive Plot
E-coli Gene Network Illustration

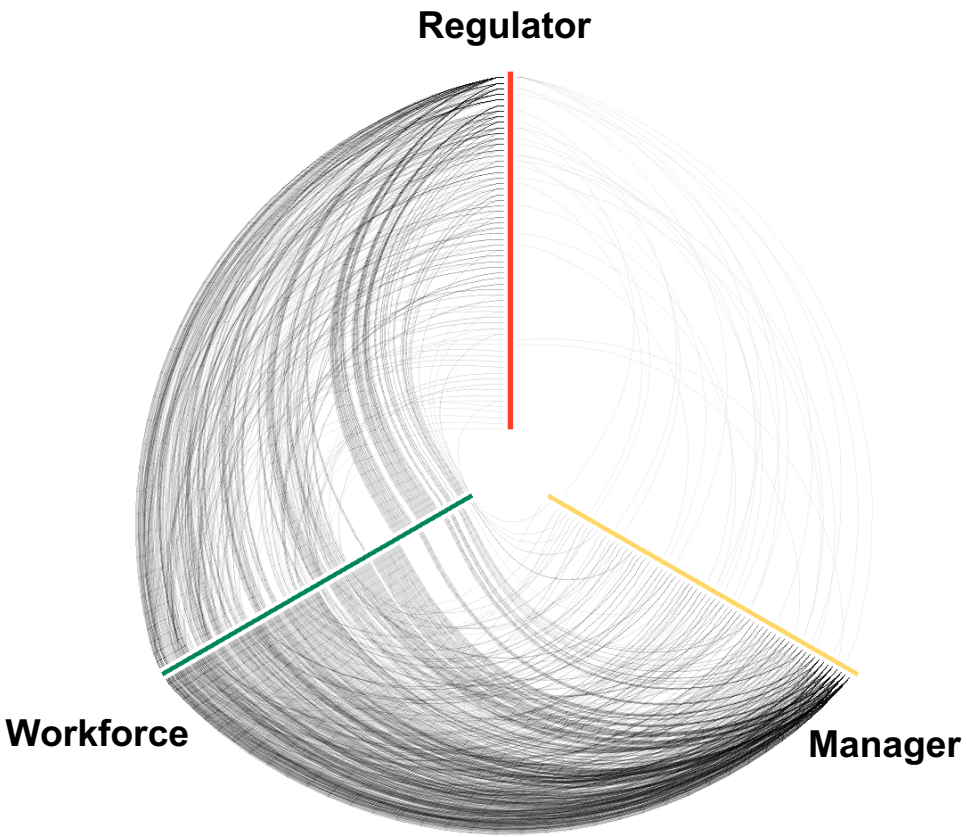
Node Classification



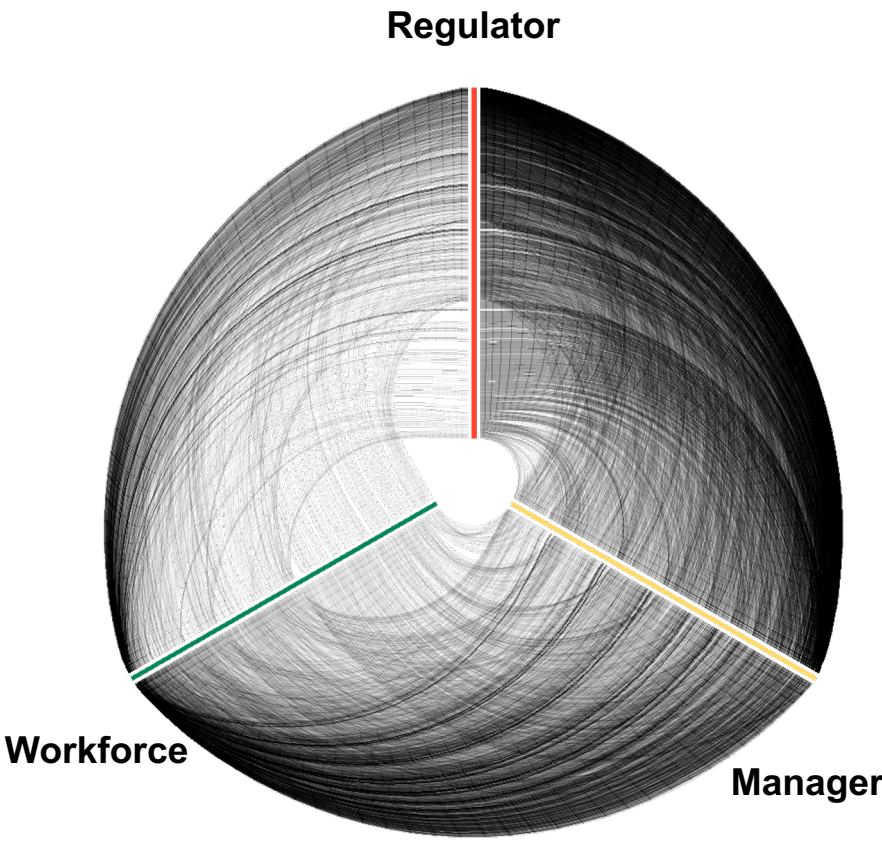




Visual Comparison of Hive Plots



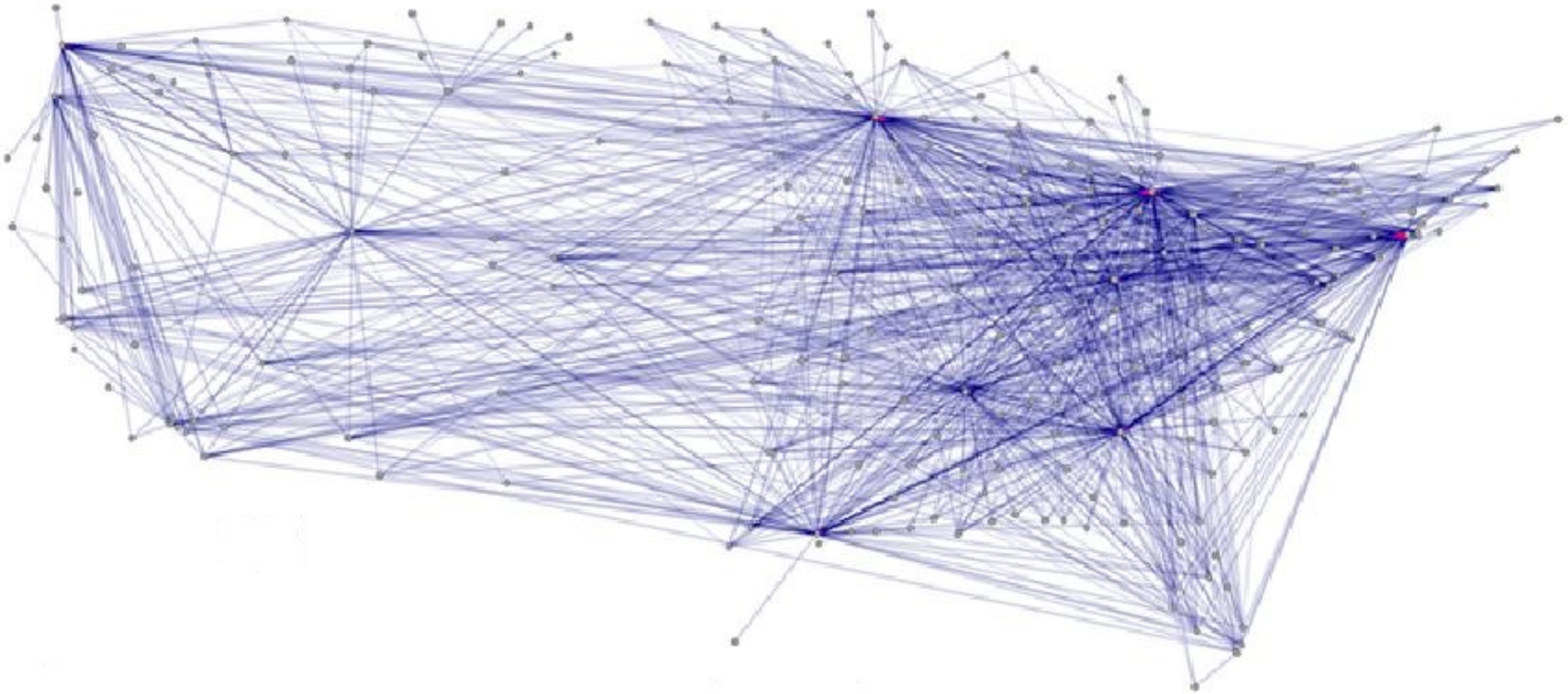
E-coli Network



Linux Network

Yan KK et. al., Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks. Proc Natl Acad Sci, 2010.

Edge Bundling



Holten, Danny, and Jarke J. Van Wijk. "Force-Directed Edge Bundling for Graph Visualization." Computer Graphics Forum. Vol. 28. No. 3. Blackwell Publishing Ltd, 2009.

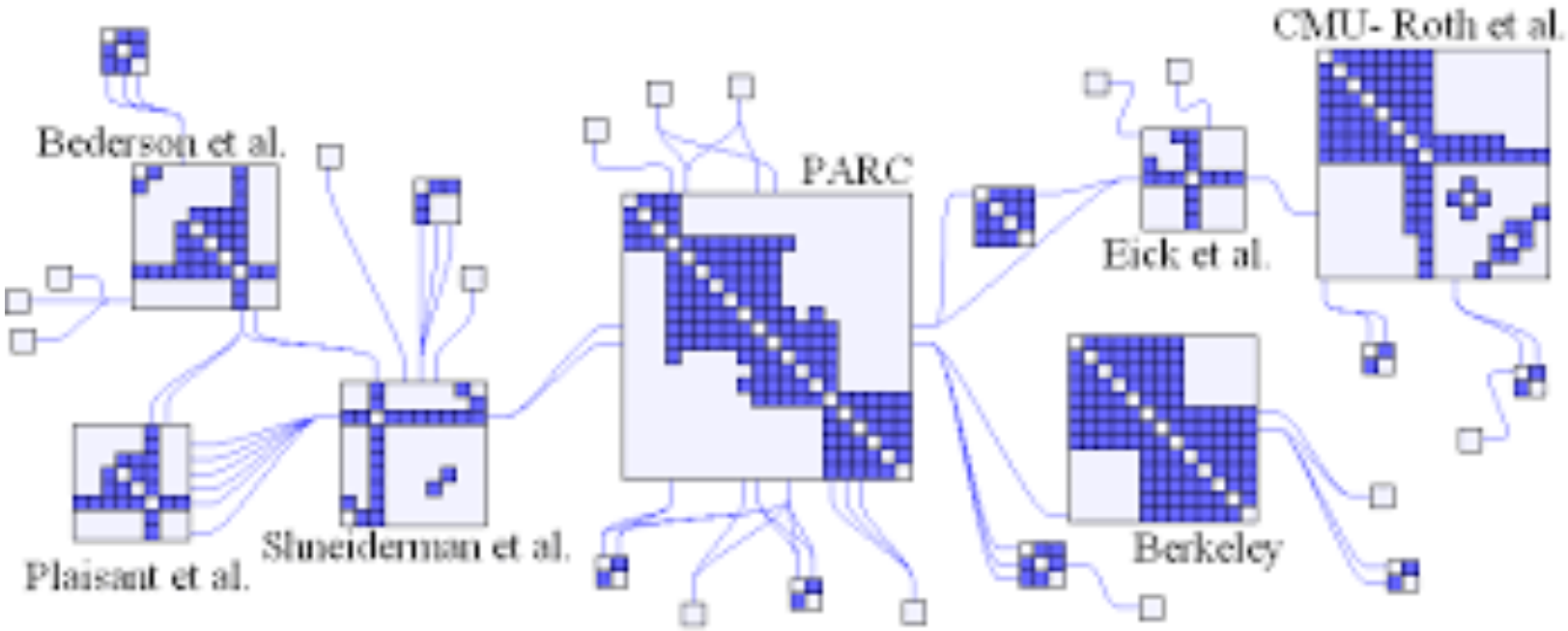
Edge Bundling

Holten, Danny, and Jarke J. Van Wijk. "Force-Directed Edge Bundling for Graph Visualization." Computer Graphics Forum. Vol. 28. No. 3. Blackwell Publishing Ltd, 2009.

Multi-View, Multi-layer Visualizations

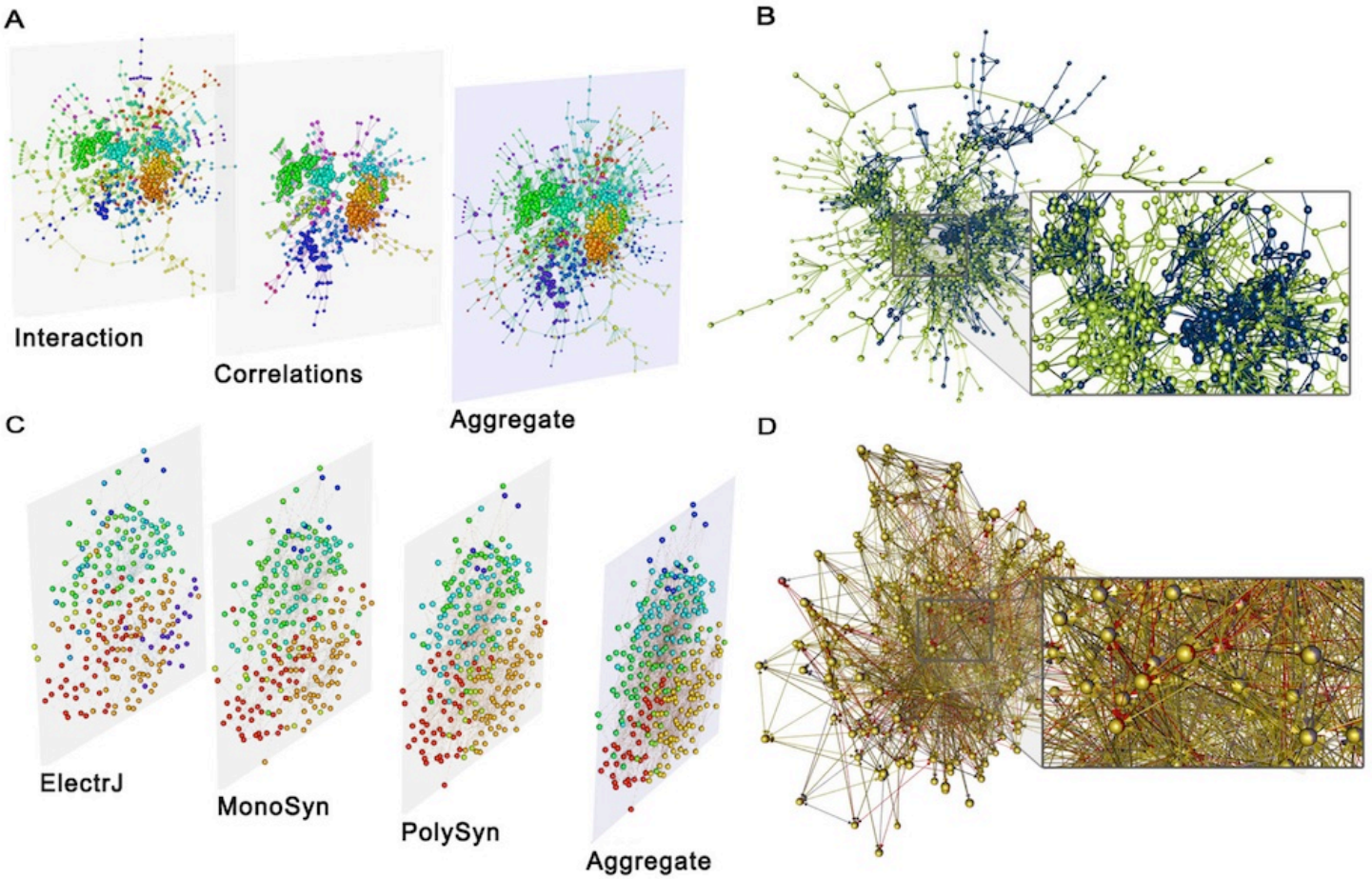
Hybrid Representations

- **Dense subgraphs: use matrix representation**
- **Sparse connections: use node-link representation**



Inria Matrix Explorer : https://hal.inria.fr/hal-00876592

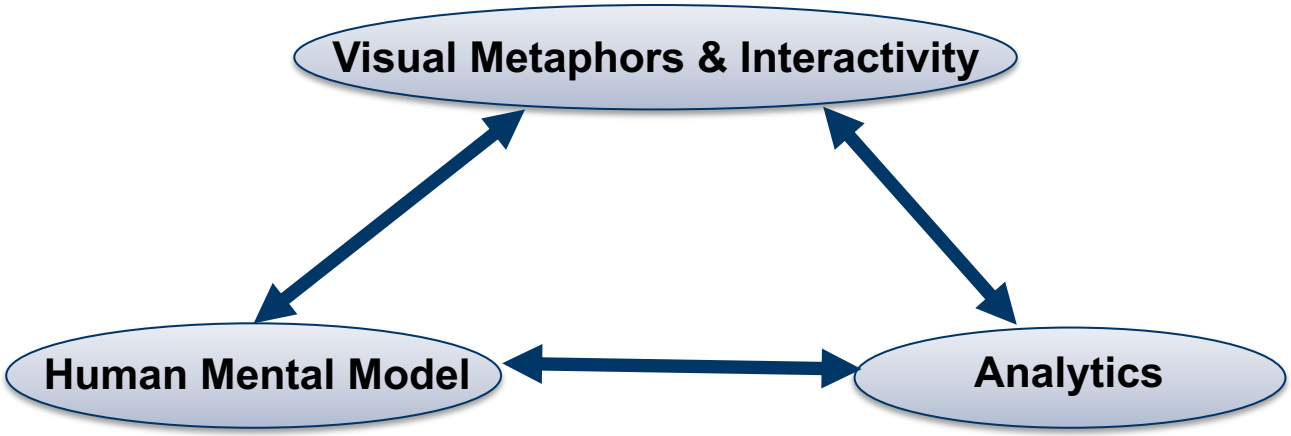
Multi-View, Multi-Layer Visualizations



Muxviz: http://muxviz.net/

Quick Recap

- Visualization can help highlight important features
- Scalability of visualization techniques is still a challenge
- Good news: there are ways to be clever about how to visualize graphs and overcome scale and complexity
- Takeaway 1: start by figuring out what questions you want visualization to help you answer
- Takeaway 2: think carefully about the application/domain specifics and use them to guide visualization design and supporting analytics



Additional Graph Drawing Resources

- **Int. Symposium on Graph Drawing:**

  graphdrawing.org

- **Handbook of Graph Drawing & Visualization:**

  cs.brown.edu/~rt/gdhandbook/

- **Graph Drawing: Algorithms for the Visualization of Graphs**

  dl.acm.org/citation.cfm?id=551884