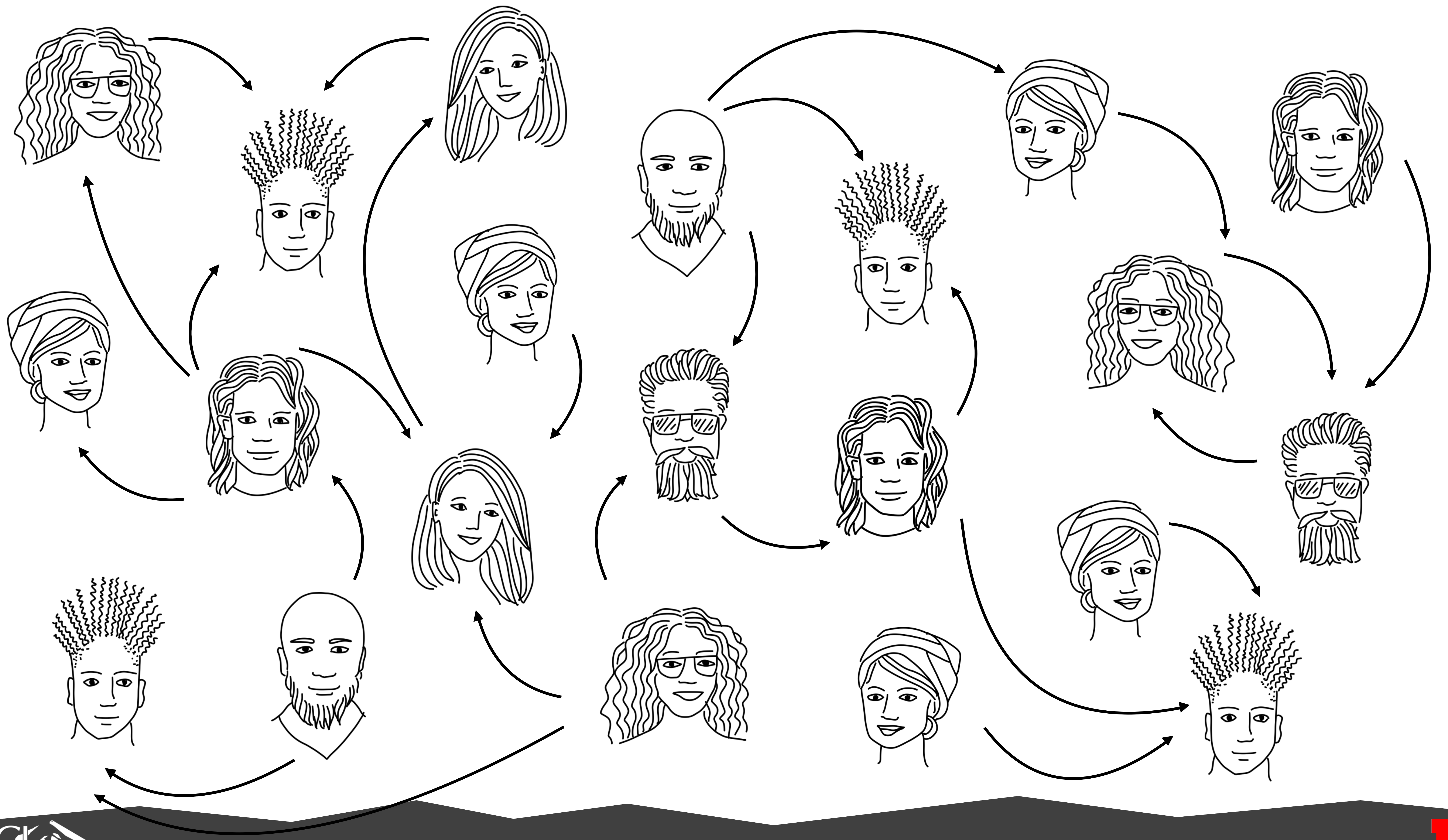
Paul Rosen
paul.rosen@utah.edu
@paulrosenphd
https://cspaul.com

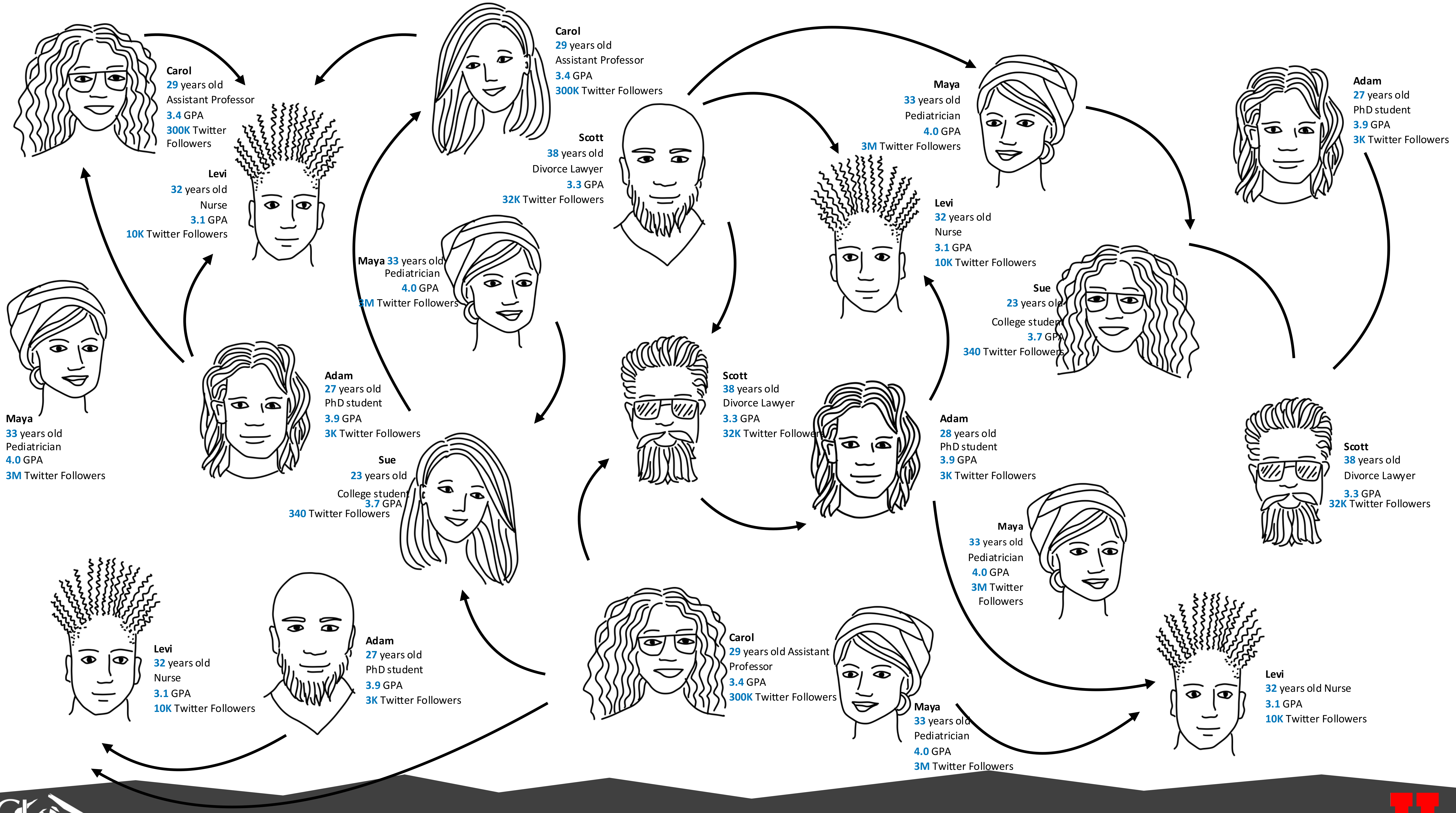
SCI
www.sci.utah.edu
THE UNIVERSITY OF UTAH

Visualization for Data Science
DS-4630 / CS-5630 / CS-6630

Visualizing Multivariate Networks

Based on an IEEE VIS Tutorial held by Carolina Notre, Marc Streit, and Alexander Lex




Carol
29 years old
Assistant Professor
3.4 GPA
300K Twitter Followers
Carol
29 years old
Assistant Professor
3.4 GPA
300K Twitter Followers
Levi
32 years old
Nurse
3.1 GPA
10K Twitter Followers
Scott
38 years old
Divorce Lawyer
3.3 GPA
32K Twitter Followers
Maya 33 years old
Pediatrician
4.0 GPA
3M Twitter Followers
Maya
33 years old
Pediatrician
4.0 GPA
3M Twitter Followers
Adam
27 years old
PhD student
3.9 GPA
3K Twitter Followers
Sue
23 years old
College student
3.7 GPA
340 Twitter Followers
Levi
32 years old
Nurse
3.1 GPA
10K Twitter Followers
Adam
27 years old
PhD student
3.9 GPA
3K Twitter Followers
Maya
33 years old
Pediatrician
4.0 GPA
3M Twitter Followers
Adam
27 years old
PhD student
3.9 GPA
3K Twitter Followers
Levi
32 years old
Nurse
3.1 GPA
10K Twitter Followers
Sue
23 years old
College student
3.7 GPA
340 Twitter Followers
Scott
38 years old
Divorce Lawyer
3.3 GPA
32K Twitter Followers
Adam
28 years old
PhD student
3.9 GPA
3K Twitter Followers
Scott
38 years old
Divorce Lawyer
3.3 GPA
32K Twitter Followers
Maya
33 years old
Pediatrician
4.0 GPA
3M Twitter Followers
Carol
29 years old Assistant Professor
3.4 GPA
300K Twitter Followers
Maya
33 years old
Pediatrician
4.0 GPA
3M Twitter Followers
Levi
32 years old Nurse
3.1 GPA
10K Twitter Followers

Multivariate Network

- Network Topology + Node and Edge Attributes
- Visualization is a tradeoff between Topology and Attributes
  - Choosing efficient encodings for one aspect often interferes with the ability to effectively visualize the other.

EUROVIS 2019
R. S. Laramee, S. Oeltze, and M. Sedlmair
(Guest Editors)

Volume 38 (2019), Number 3
STAR – State of The Art Report

The State of the Art in Visualizing Multivariate Networks

C. Nobre[1], M. Meyer[1], M. Streit[2], and A. Lex[1]

[1]University of Utah, Utah, USA
[2]Johannes Kepler University Linz, Austria

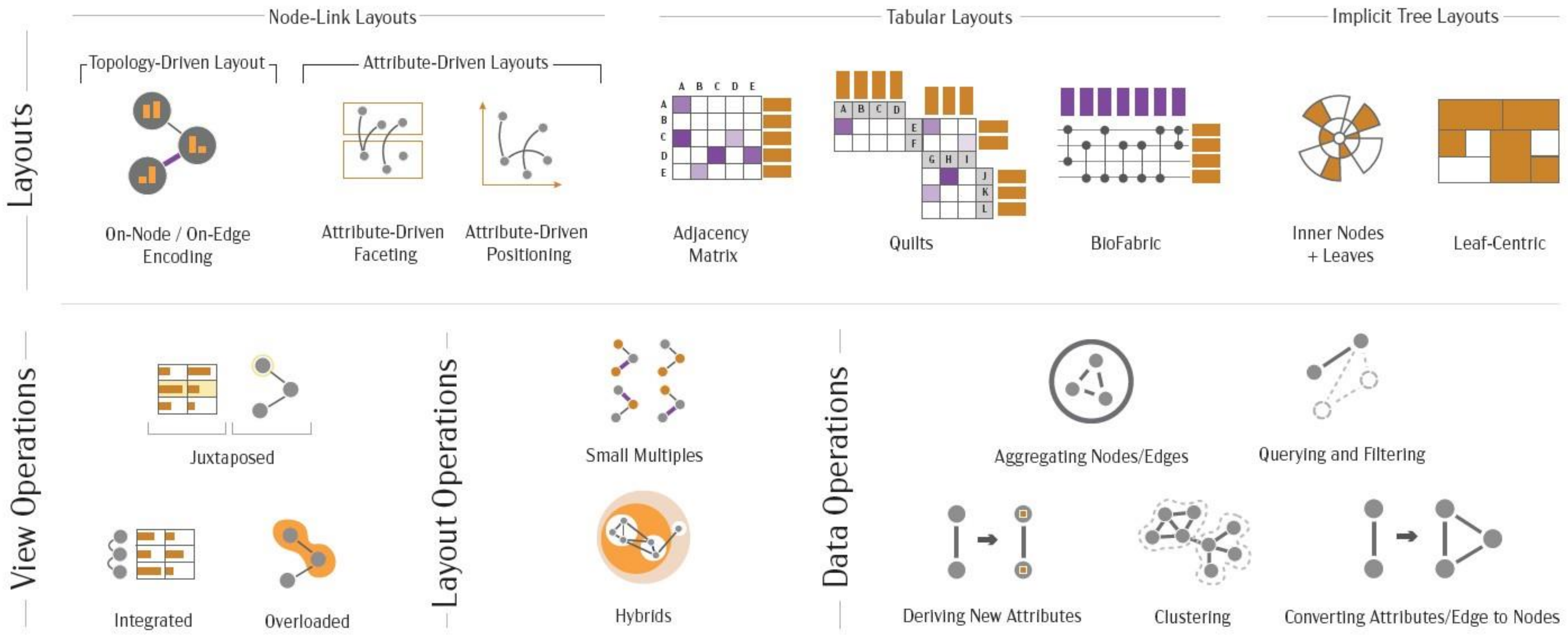
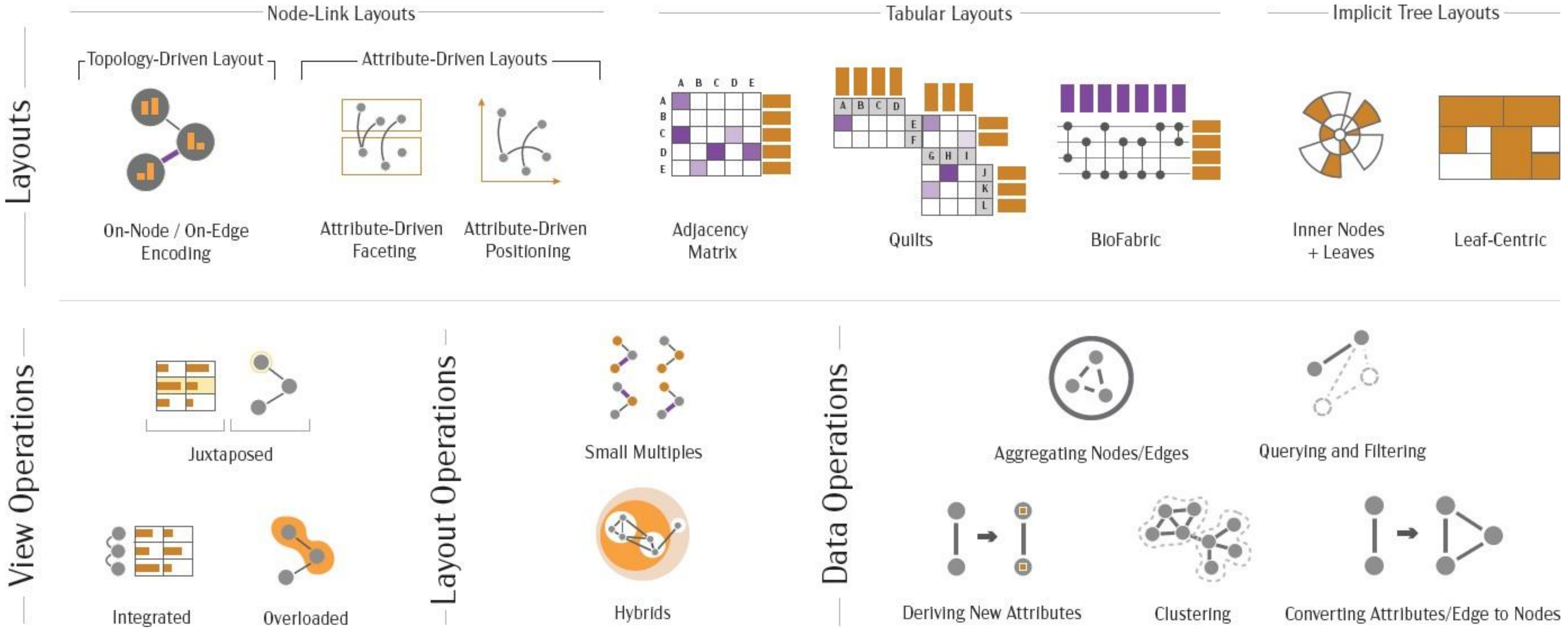


Figure 1: *A typology of operations and layouts used in multivariate network visualization.* ***Layouts*** *describe the fundamental choices for encoding multivariate networks.* ***View Operations*** *capture how topology and attribute focused visualizations can be combined.* ***Layout Operations*** *are applied to basic layouts to create specific visualization techniques.* ***Data Operations*** *are used to transform a network or derive attributes before visualizations. The colors reflect node attributes (orange), edge attributes (purple), and topology (grey).*

Abstract

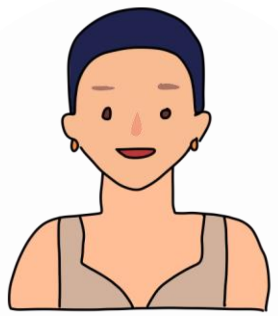
Multivariate networks are made up of nodes and their relationships (links), but also data about those nodes and links as attributes. Most real-world networks are associated with several attributes, and many analysis tasks depend on analyzing both, relationships and attributes. Visualization of multivariate networks, however, is challenging, especially when both the topology of the network and the attributes need to be considered concurrently. In this state-of-the-art report, we analyze current practices and classify techniques along four axes: layouts, view operations, layout operations, and data operations. We also provide an analysis of tasks specific to multivariate networks and give recommendations for which technique to use in which scenario. Finally, we survey application areas and evaluation methodologies.


Layouts
Node-Link Layouts
Topology-Driven Layout
Attribute-Driven Layouts
On-Node / On-Edge Encoding
Attribute-Driven Faceting
Attribute-Driven Positioning
Tabular Layouts
Adjacency Matrix
Quilts
BioFabric
Implicit Tree Layouts
Inner Nodes + Leaves
Leaf-Centric
View Operations
Juxtaposed
Integrated
Overloaded
Layout Operations
Small Multiples
Hybrids
Data Operations
Aggregating Nodes/Edges
Querying and Filtering
Deriving New Attributes
Clustering
Converting Attributes/Edge to Nodes

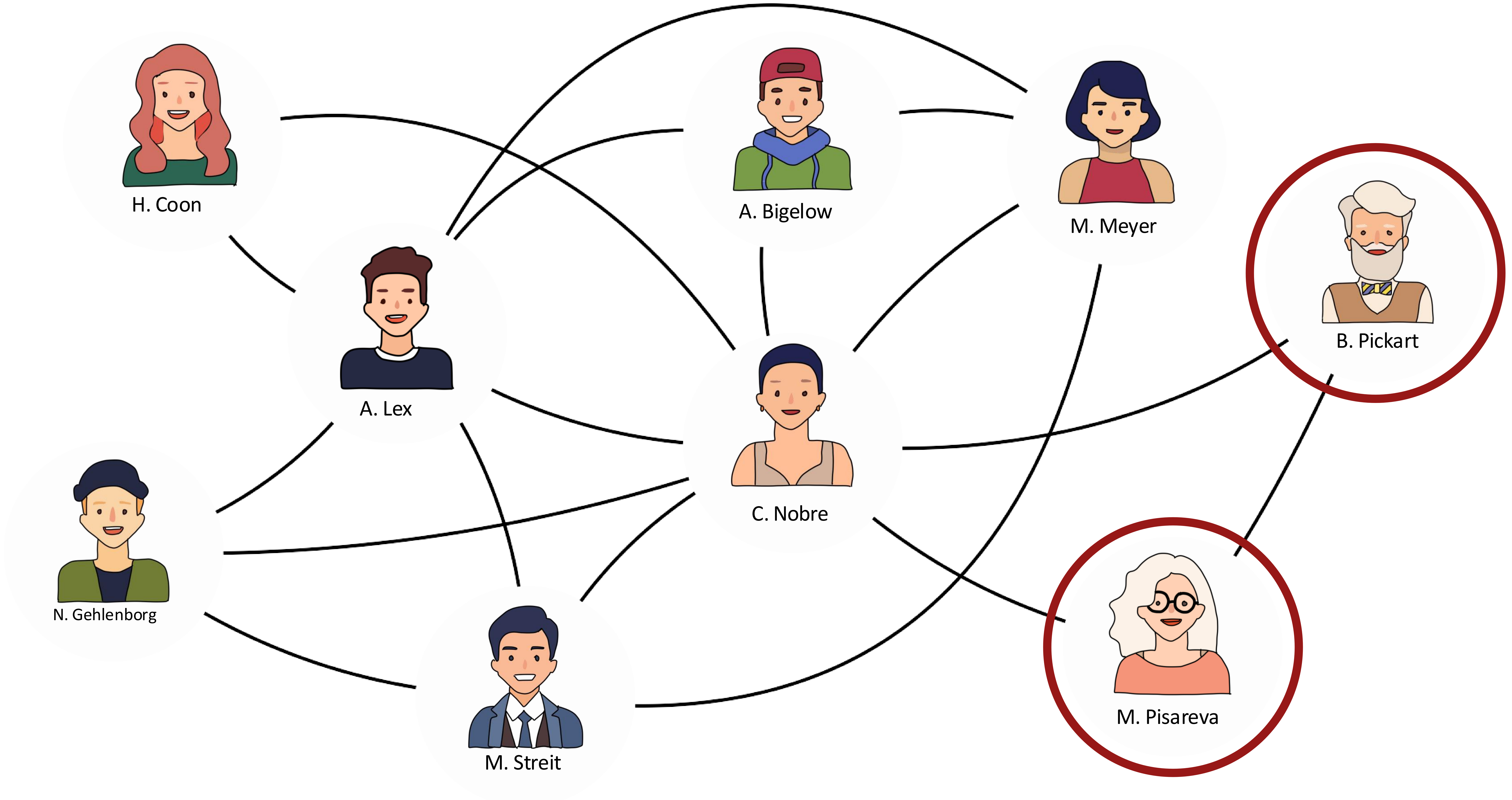
Multivariate Network Tasks

How is a multivariate network task different than a regular graph task?

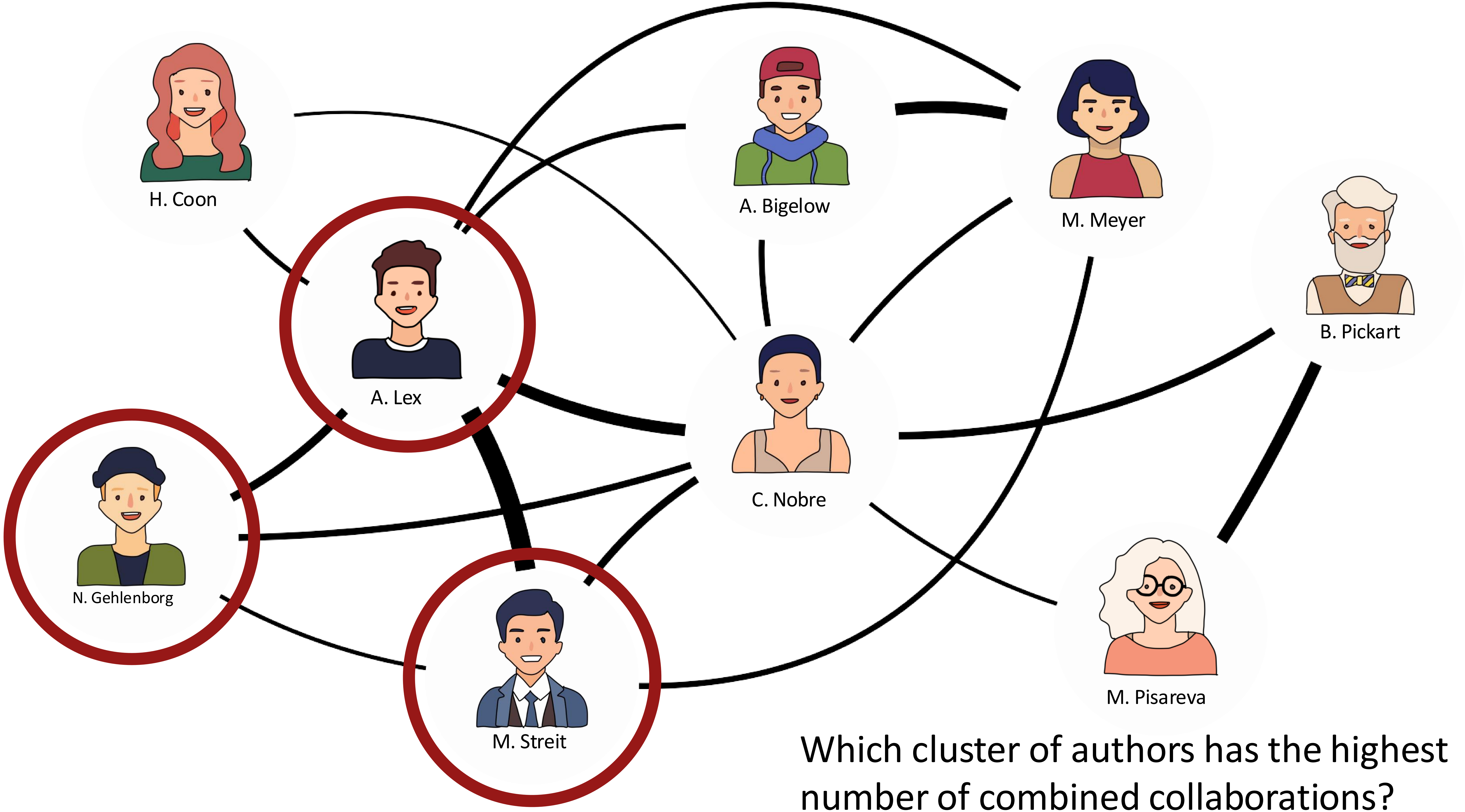
- Rely on both the topology of the network and the attributes of the nodes and edges



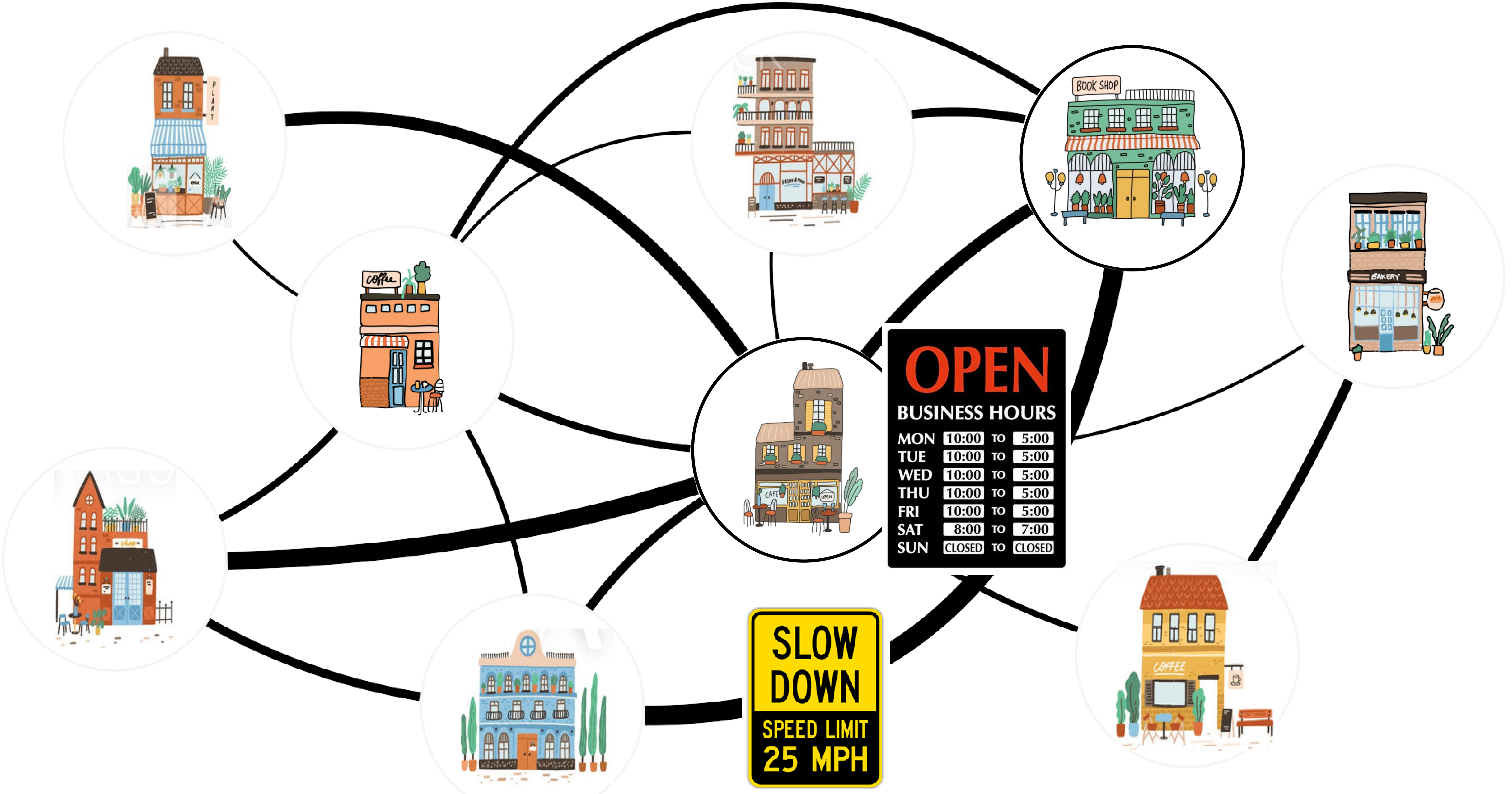
C. Nobre



How many of my collaborators are from the oceanography field?



Which cluster of authors has the highest number of combined collaborations?



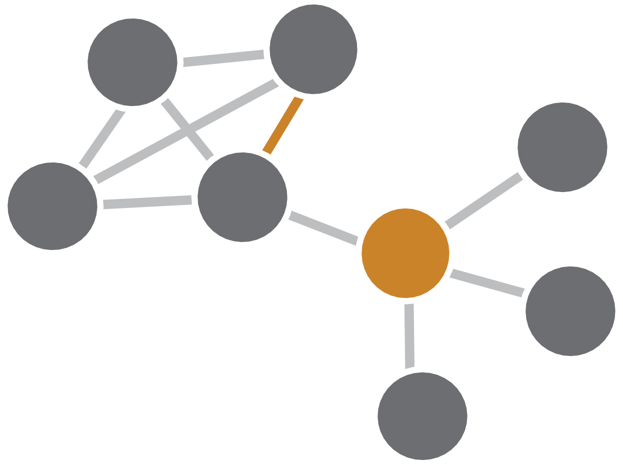
What is an efficient way I can complete all my errands?

- How many of my collaborators are in the oceanography field?
- Which cluster has the highest number of collaborations?
- What is the fastest route to get all my errands done?

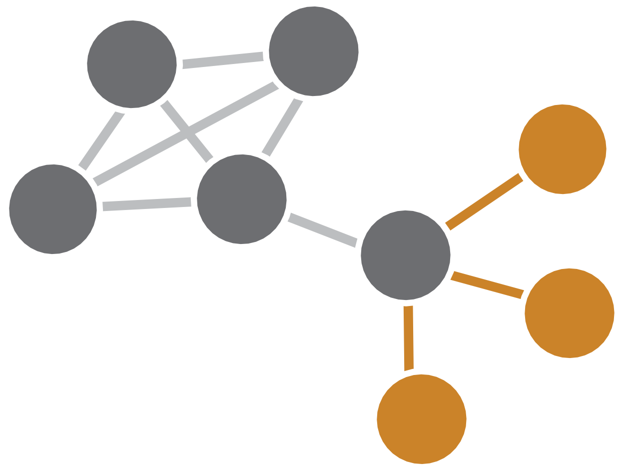
Tasks that rely on the **topology** of the network and the **attributes** of the nodes and edges

MVNV tasks are applied to topological structures

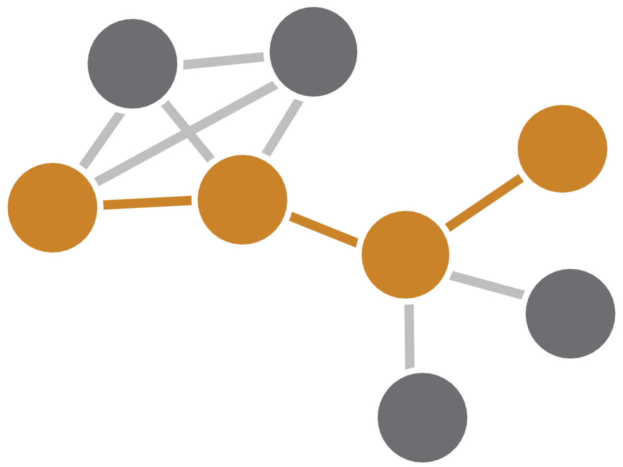
Single Node/Edge



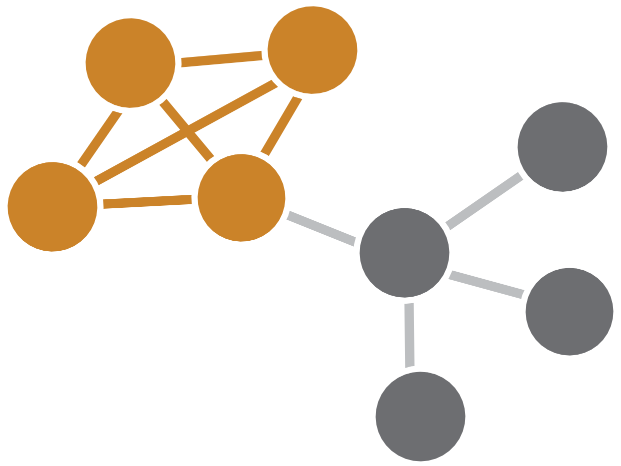
Node Neighbors



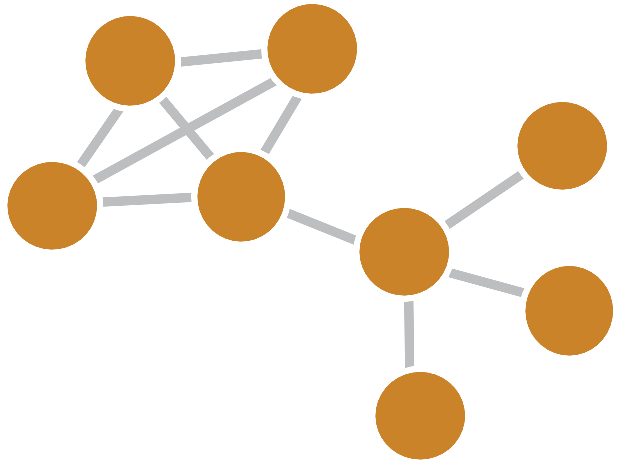
Path



Cluster

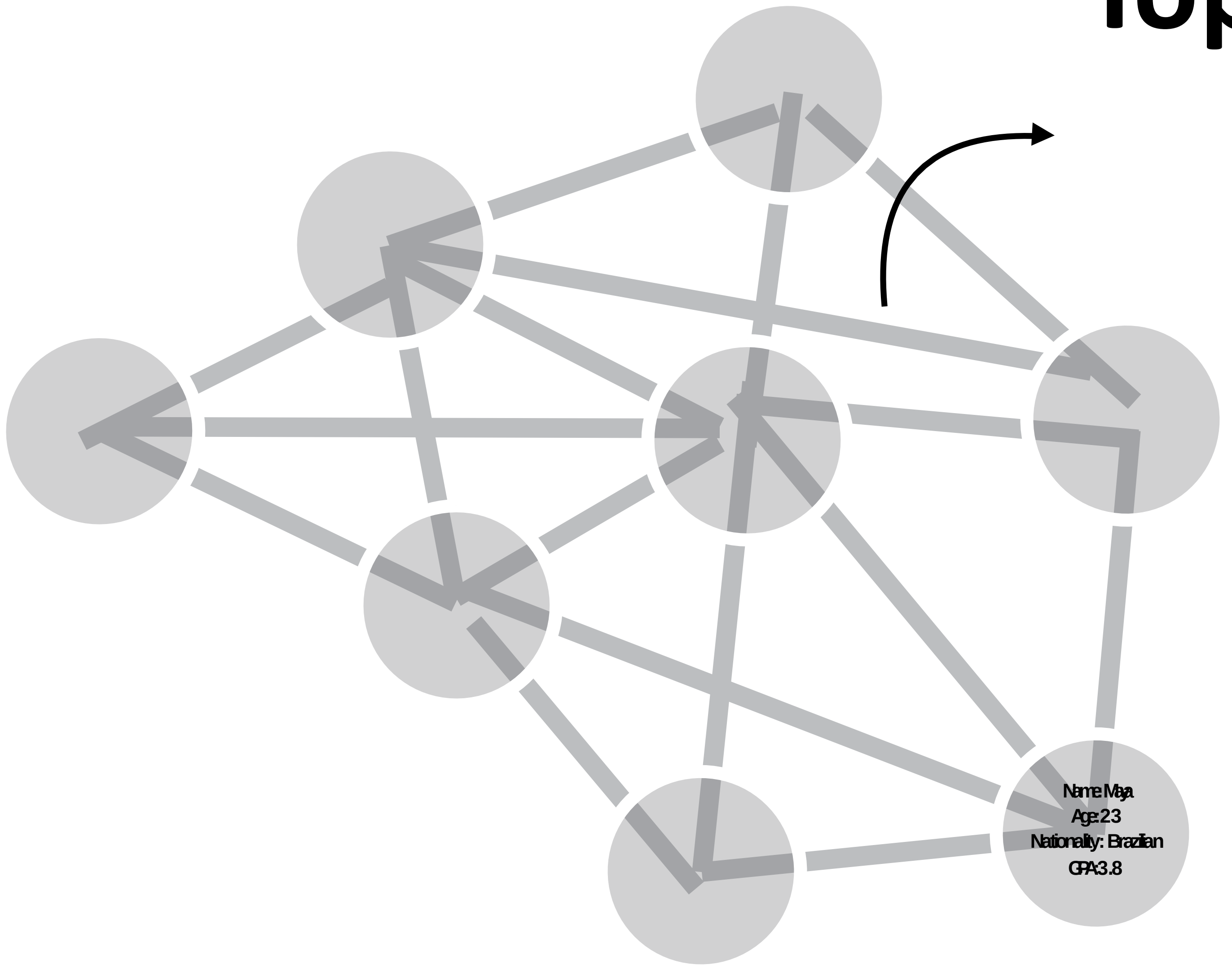


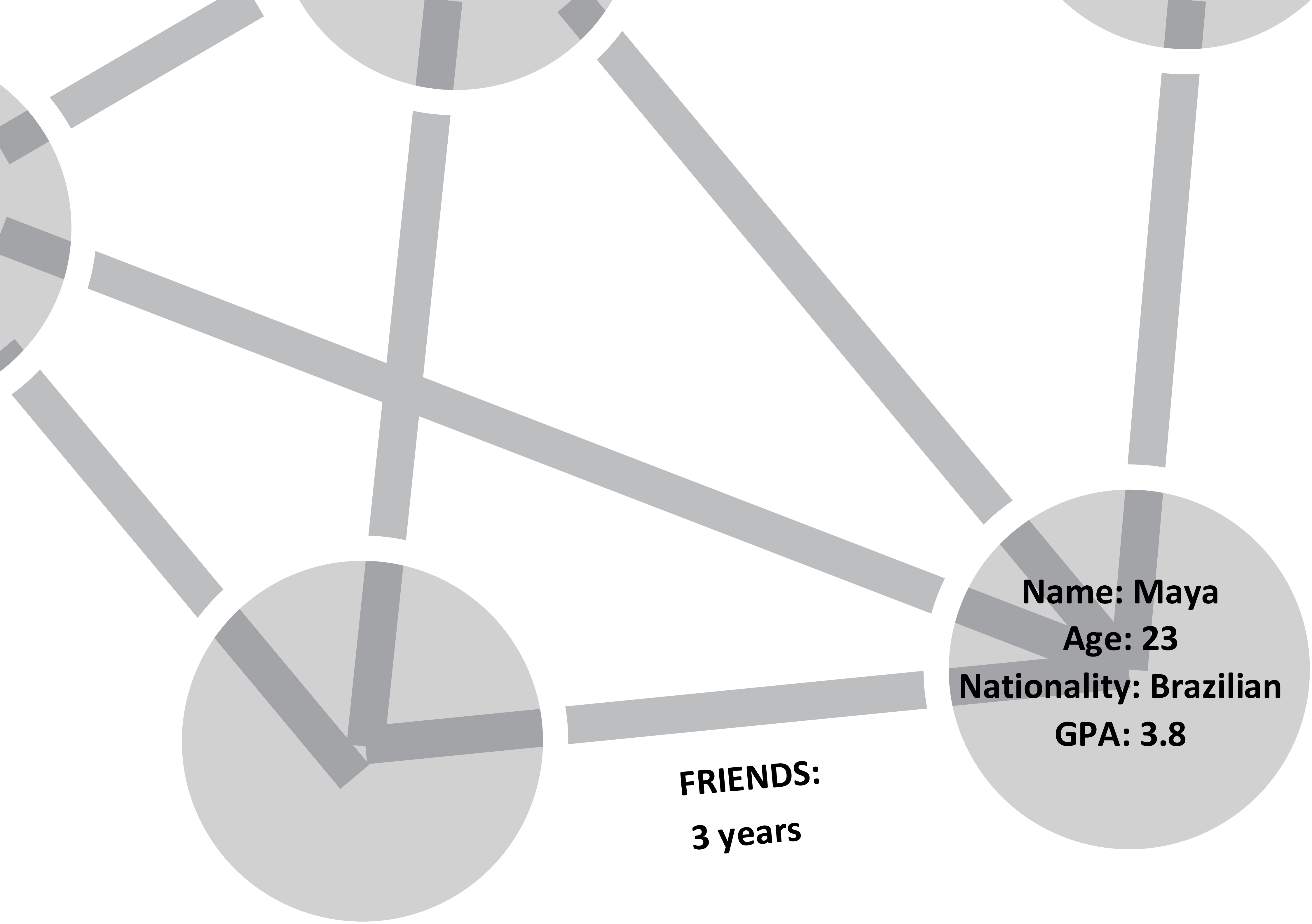
Network/Subnetwork

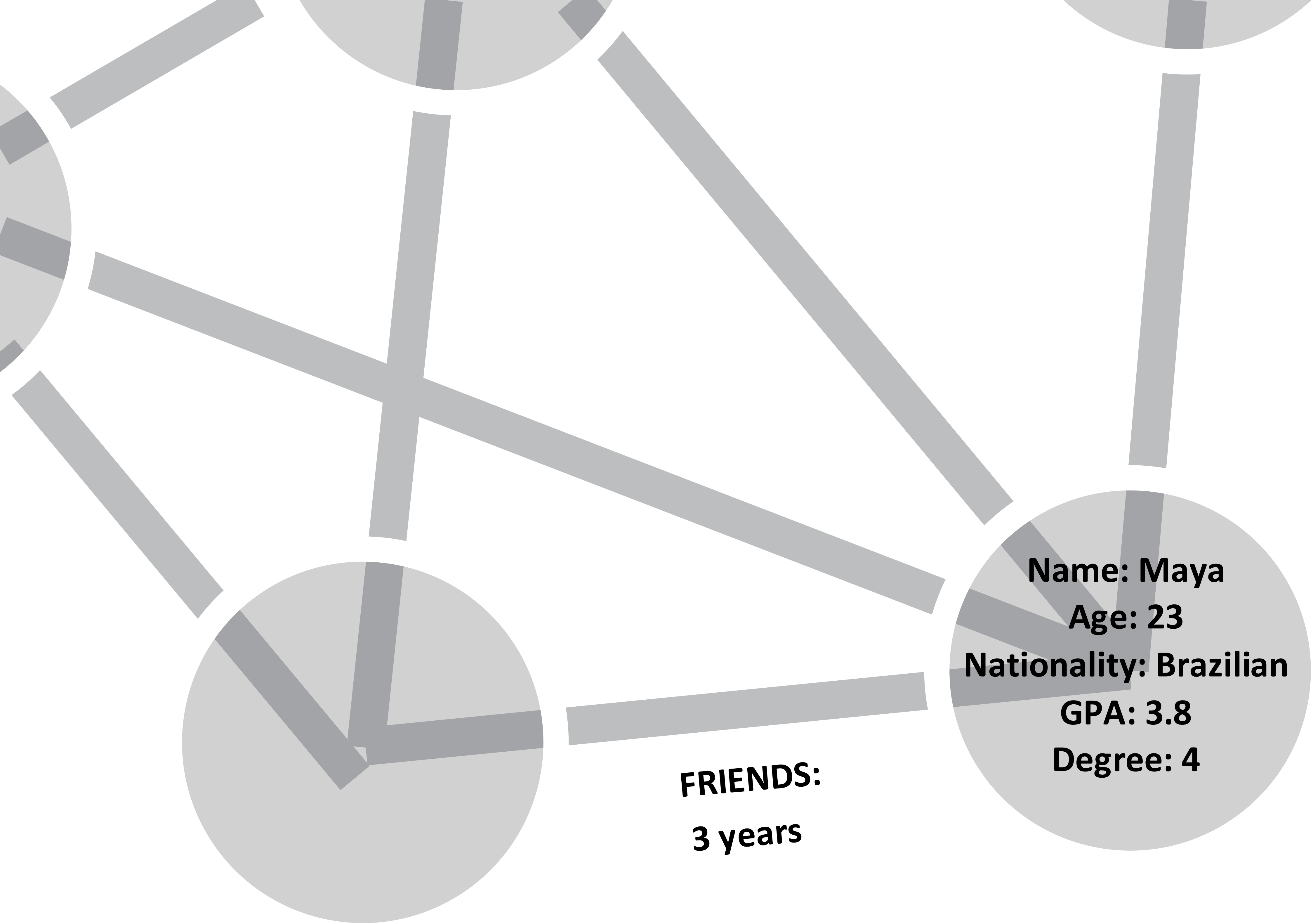


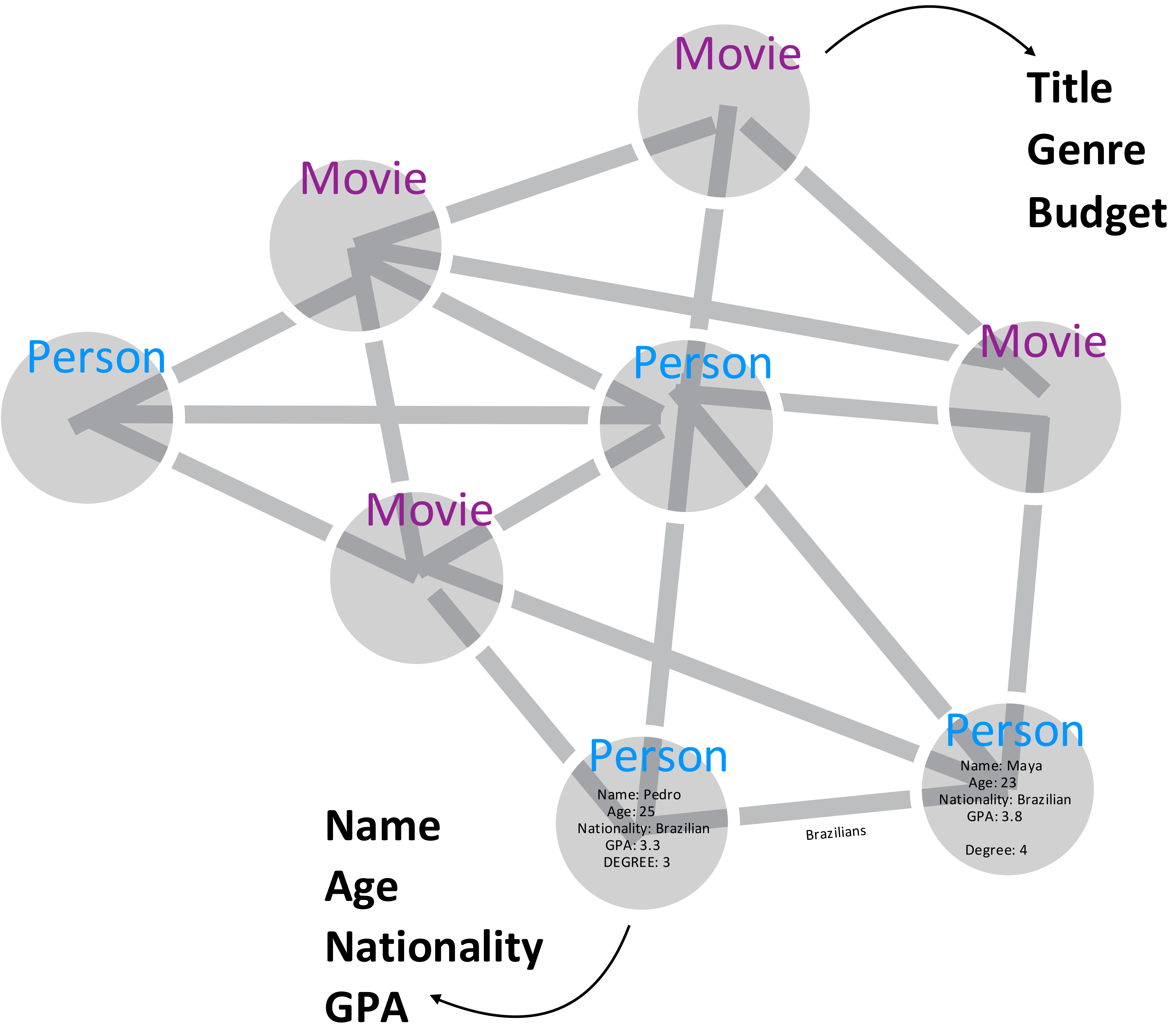
Network and Attribute Characteristics

Topology



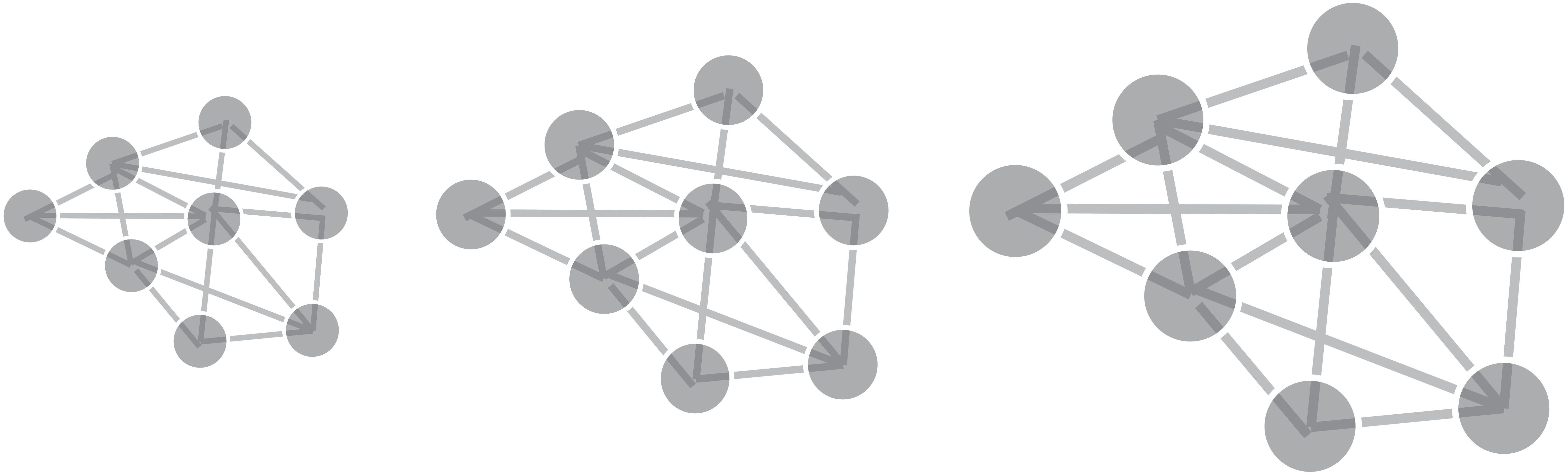

Name: Maya
Age: 23
Nationality: Brazilian
GPA: 3.8
FRIENDS:
3 years


Name: Maya
Age: 23
Nationality: Brazilian
GPA: 3.8
Degree: 4
FRIENDS:
3 years



Movie
Title
Genre
Budget
Movie
Person
Person
Movie
Movie
Person
Person
Name: Pedro
Age: 25
Nationality: Brazilian
GPA: 3.3
DEGREE: 3
Brazilians
Name: Maya
Age: 23
Nationality: Brazilian
GPA: 3.8
Degree: 4
Name
Age
Nationality
GPA

Network Size

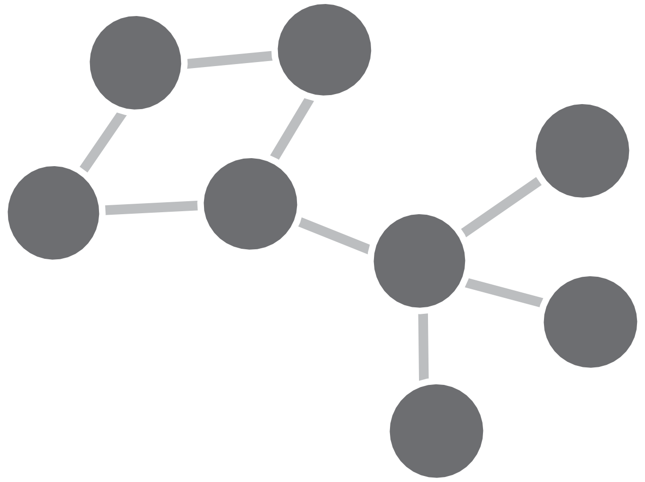


Small
<100

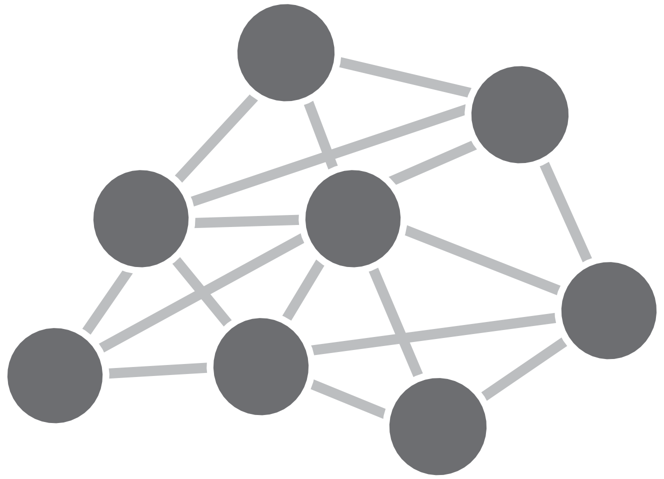
Medium
100-1000

Large
>1000

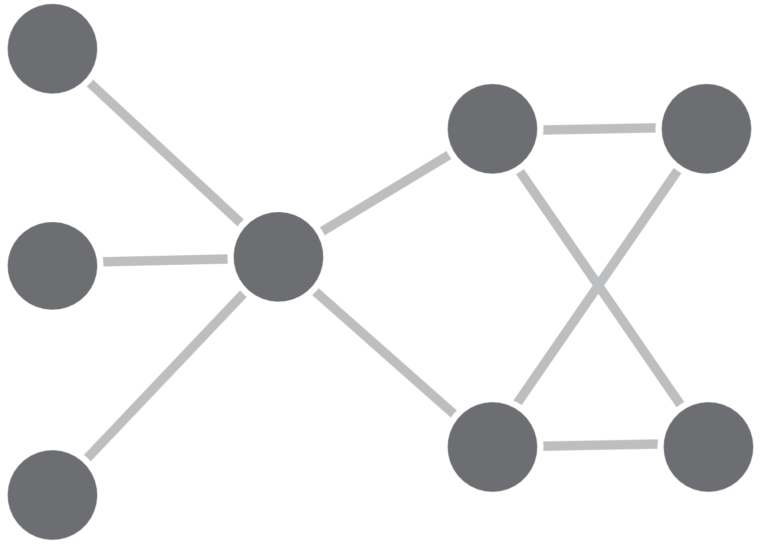
Network Types



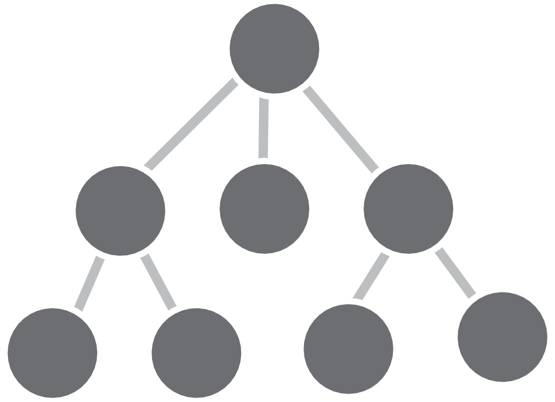
Sparse



Dense

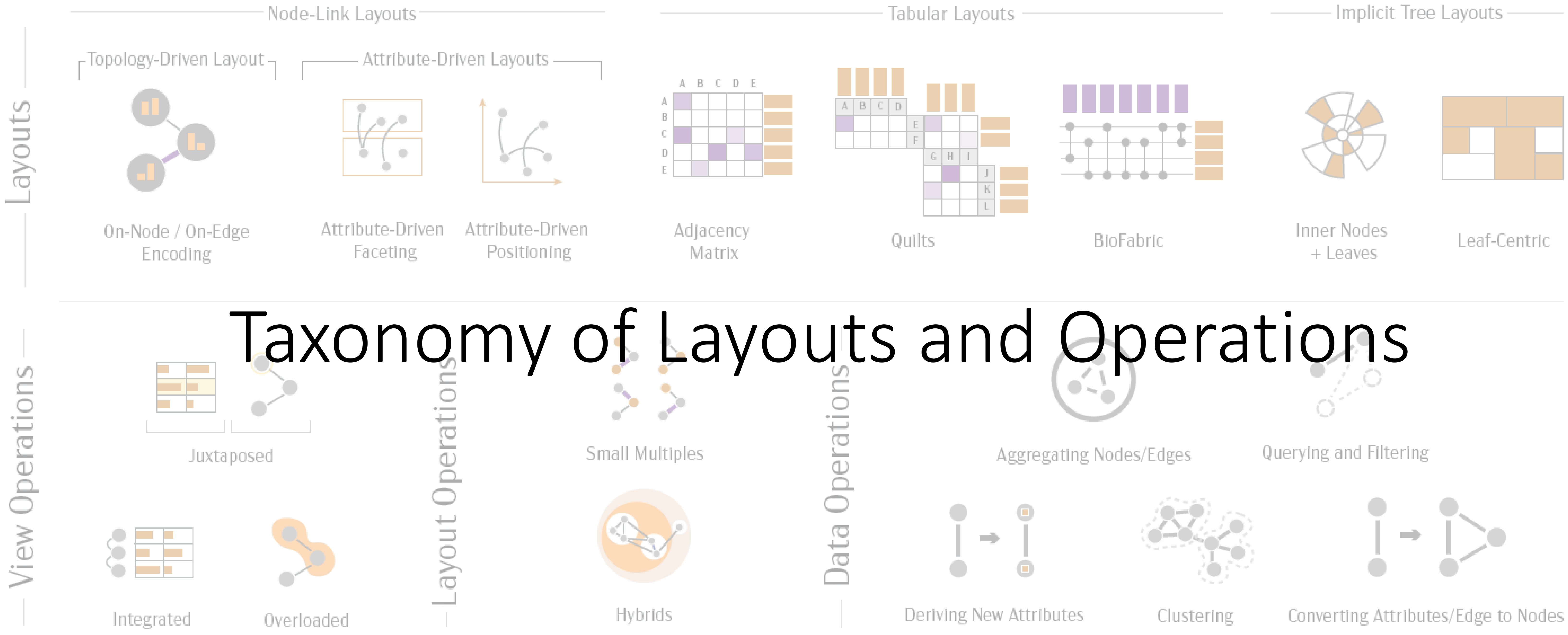


Layered



Trees

Taxonomy of Layouts and Operations

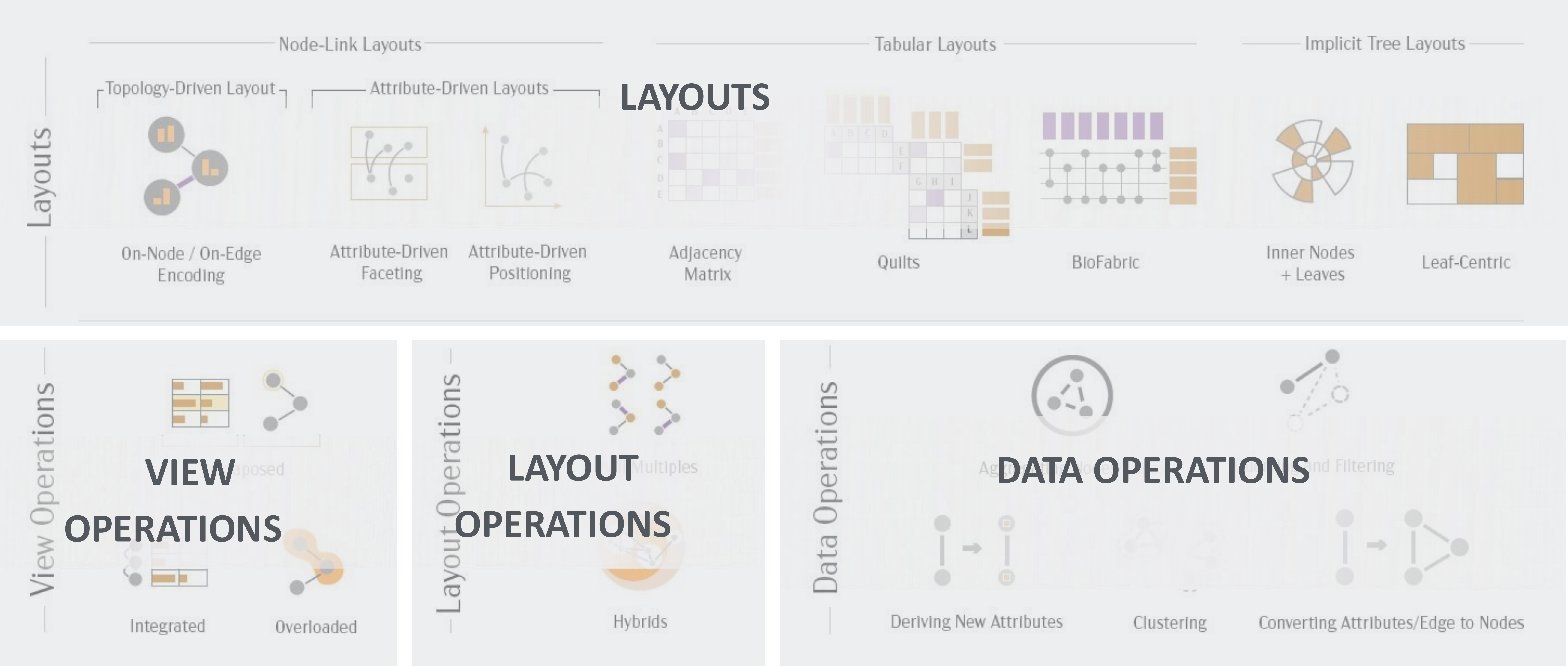


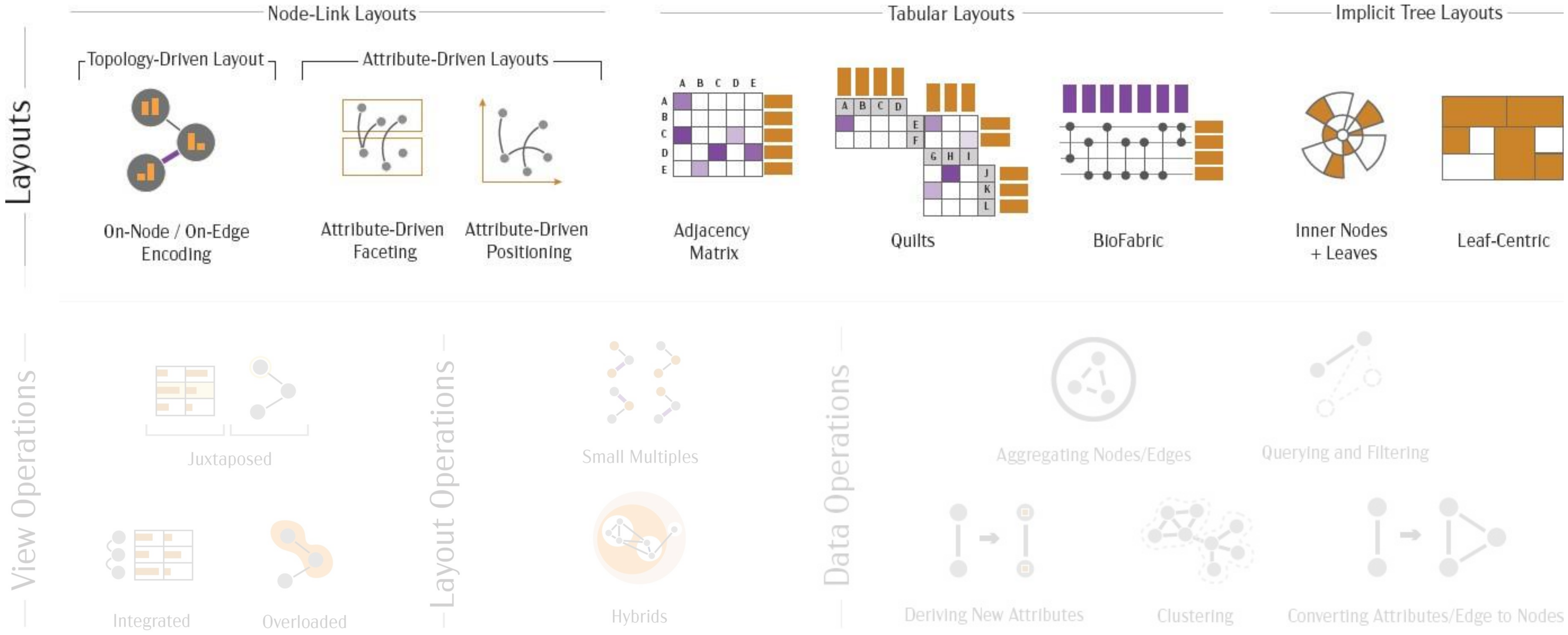
LAYOUTS

VIEW OPERATIONS

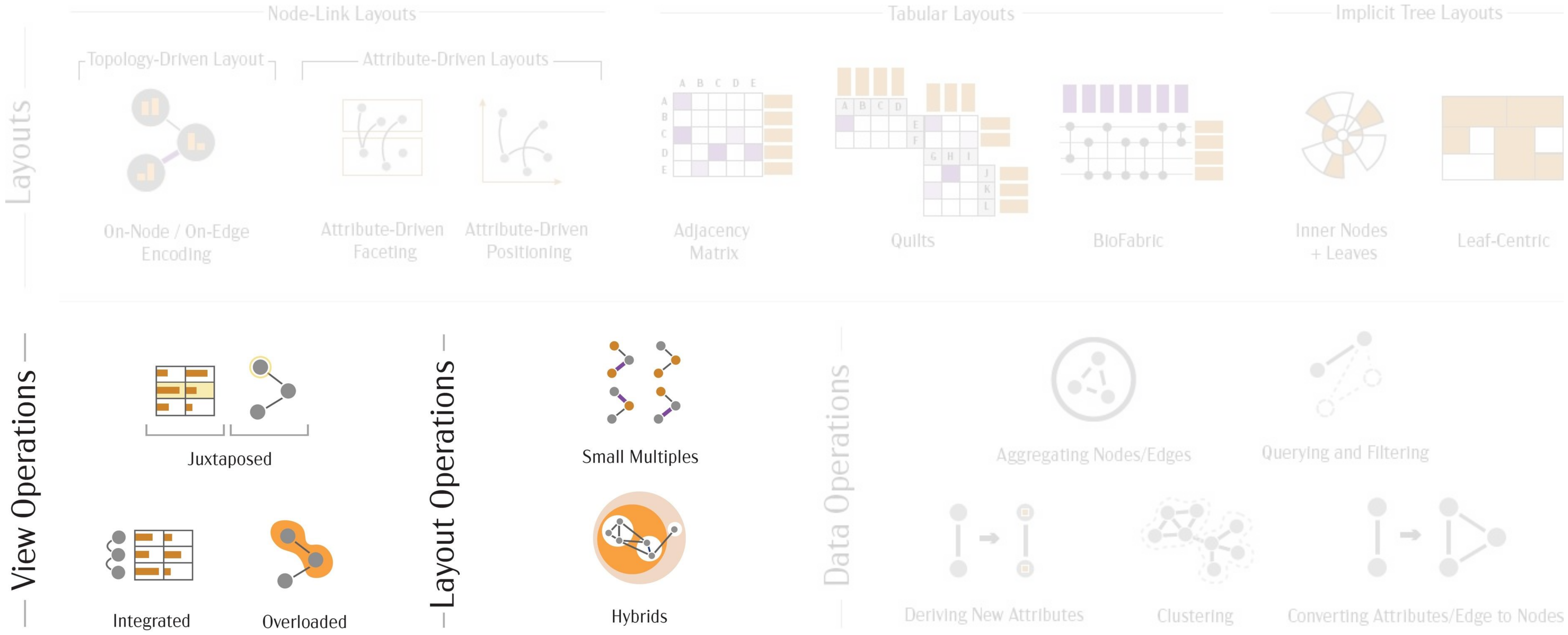
LAYOUT OPERATIONS

DATA OPERATIONS



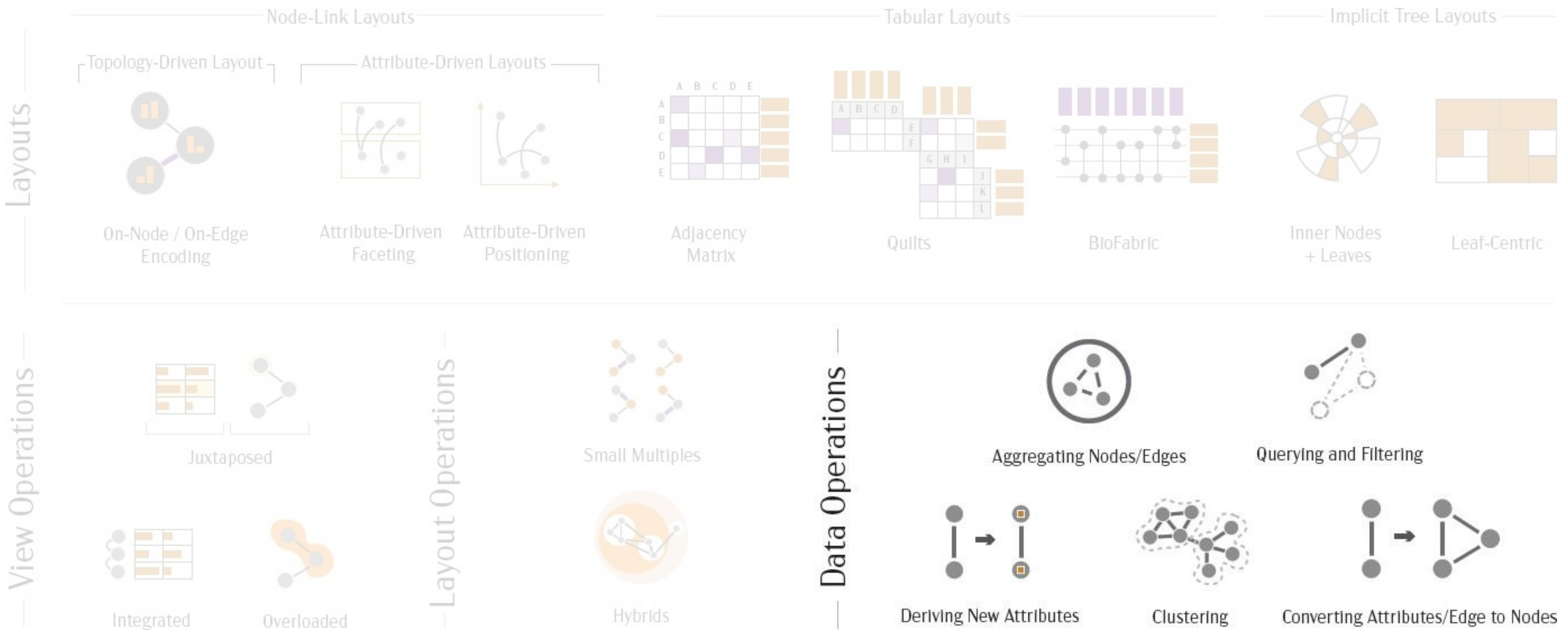


Layouts
Node-Link Layouts
Tabular Layouts
Implicit Tree Layouts
Topology-Driven Layout
Attribute-Driven Layouts
On-Node / On-Edge Encoding
Attribute-Driven Faceting
Attribute-Driven Positioning
Adjacency Matrix
Quilts
BioFabric
Inner Nodes + Leaves
Leaf-Centric
View Operations
Juxtaposed
Integrated
Overloaded
Layout Operations
Small Multiples
Hybrids
Data Operations
Aggregating Nodes/Edges
Querying and Filtering
Deriving New Attributes
Clustering
Converting Attributes/Edge to Nodes



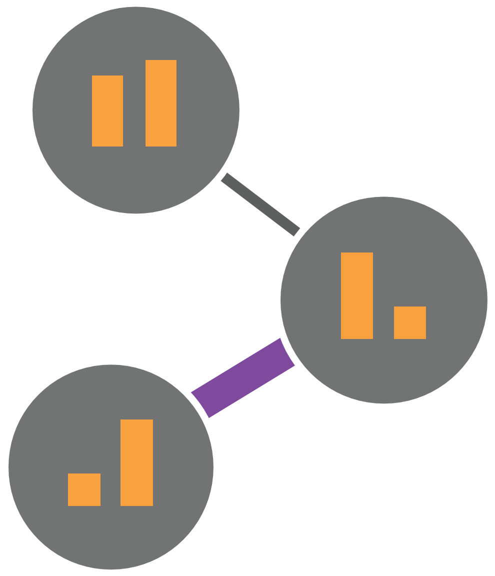
Separate Views for
Topology and Attributes

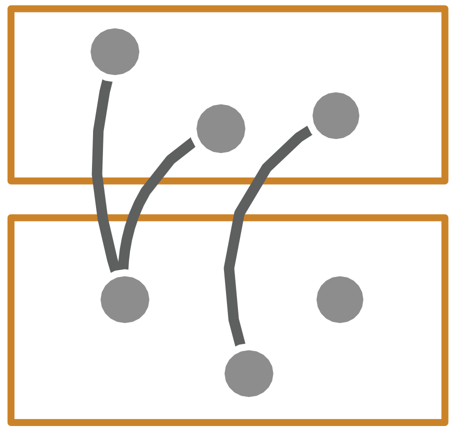
Multiple layouts for
Topology or Attributes

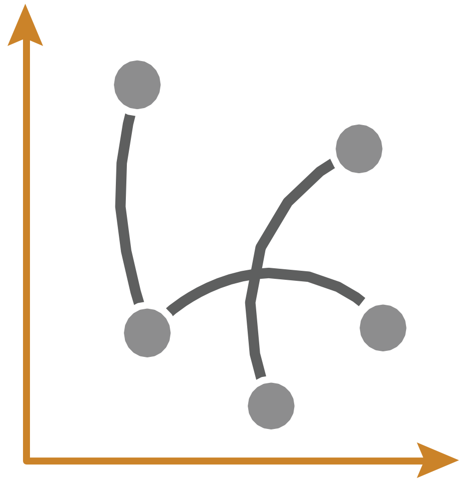


Node-Link Layouts
Tabular Layouts
Implicit Tree Layouts
Layouts
Topology-Driven Layout
Attribute-Driven Layouts
On-Node / On-Edge Encoding
Attribute-Driven Faceting
Attribute-Driven Positioning
Adjacency Matrix
Quilts
BioFabric
Inner Nodes + Leaves
Leaf-Centric
View Operations
Juxtaposed
Integrated
Overloaded
Layout Operations
Small Multiples
Hybrids
Data Operations
Aggregating Nodes/Edges
Querying and Filtering
Deriving New Attributes
Clustering
Converting Attributes/Edge to Nodes

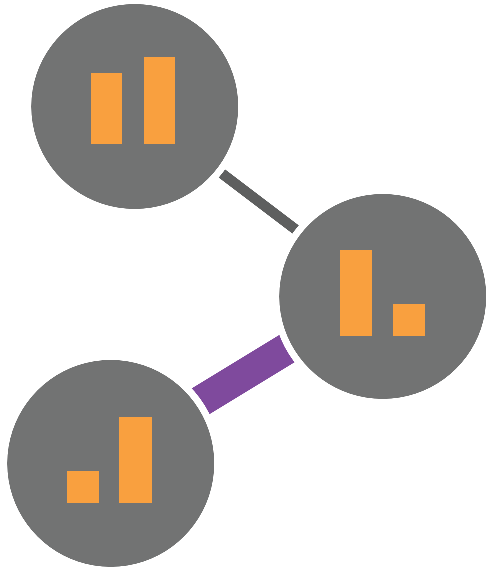
Node-Link Layouts





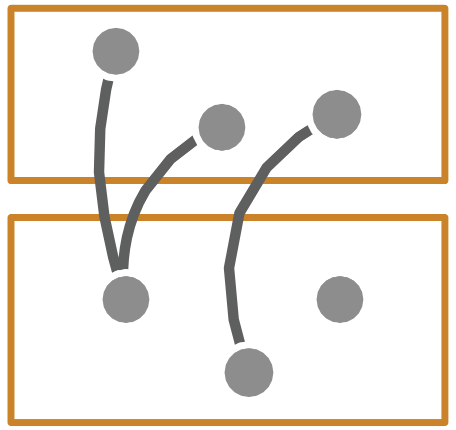


Topology Driven Layout

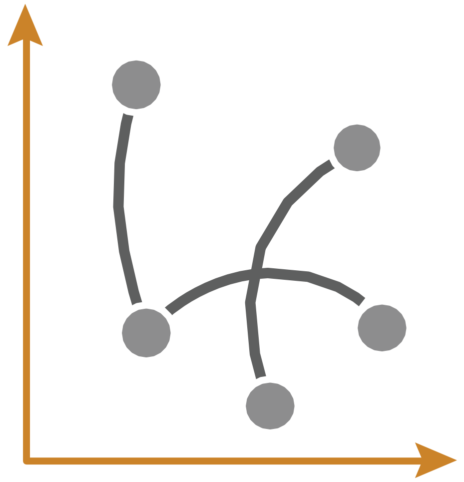


On-Node / On-Edge
Encoding

Attribute Driven Layouts

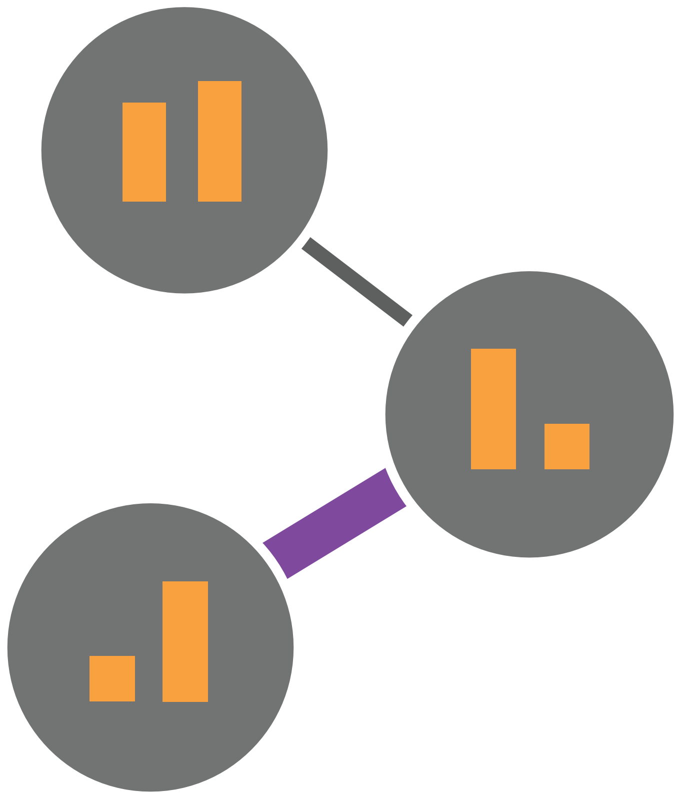


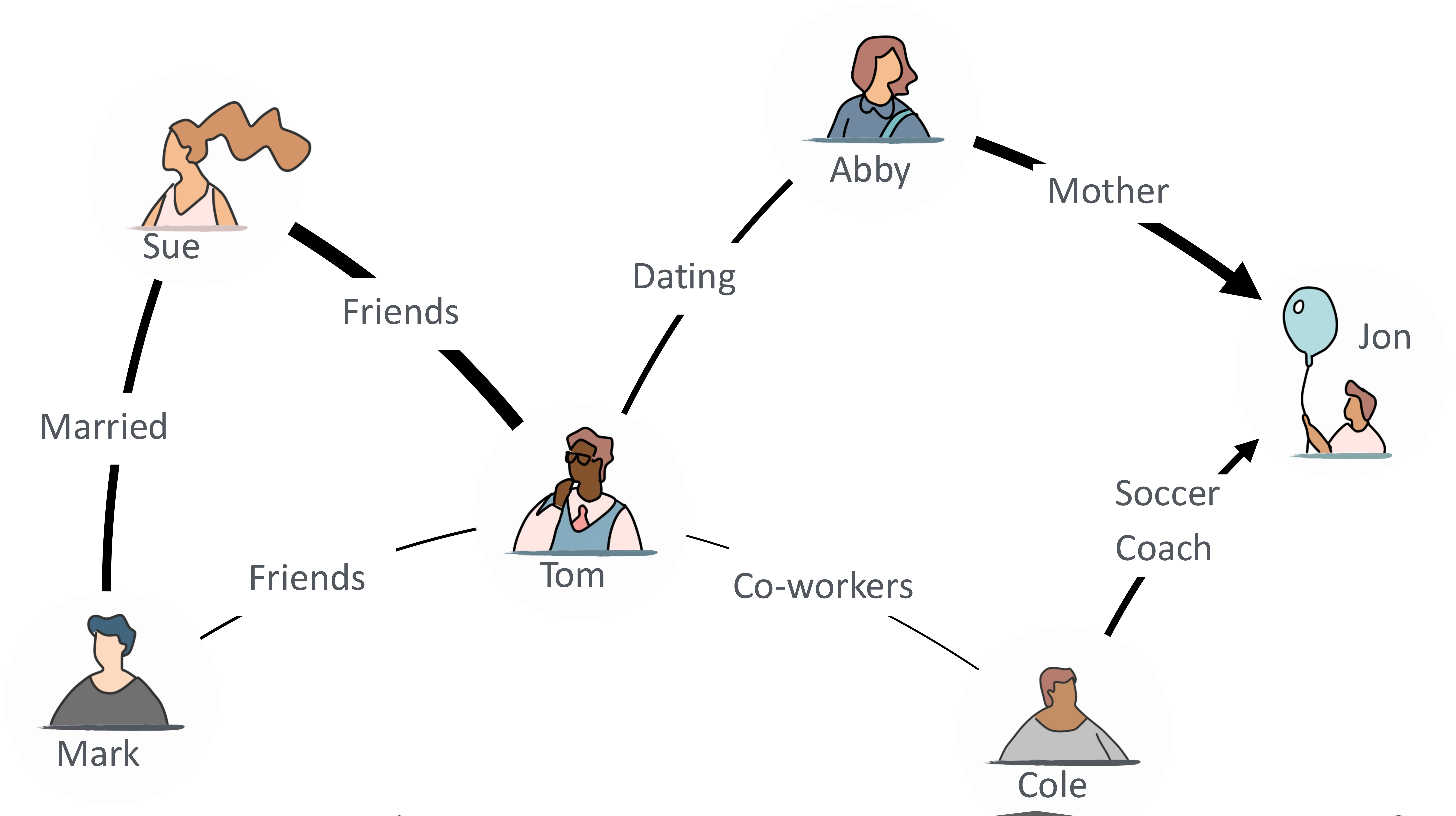
Attribute-Driven
Faceting



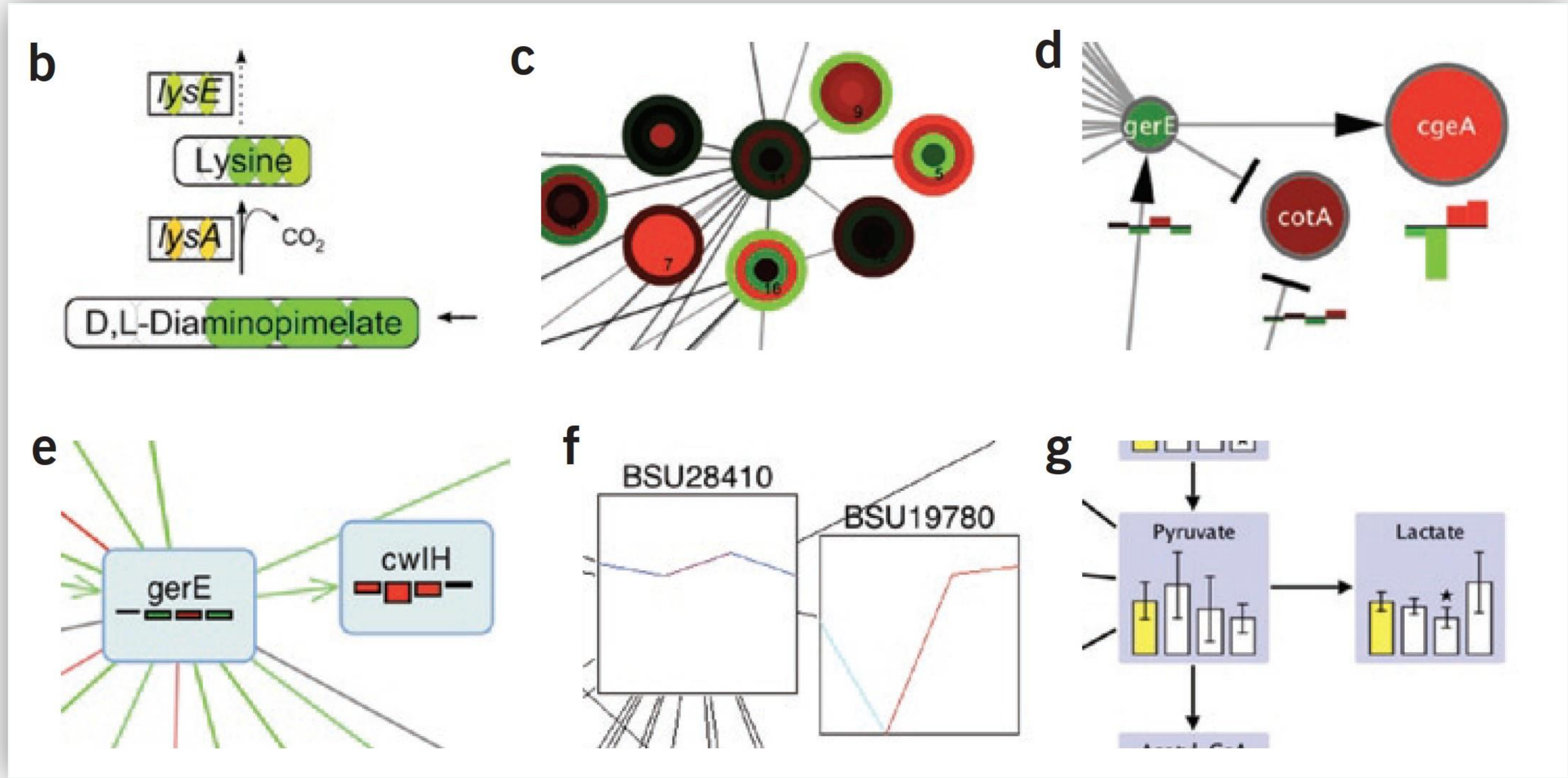
Attribute-Driven
Positioning

On-Node / On-Edge Encoding

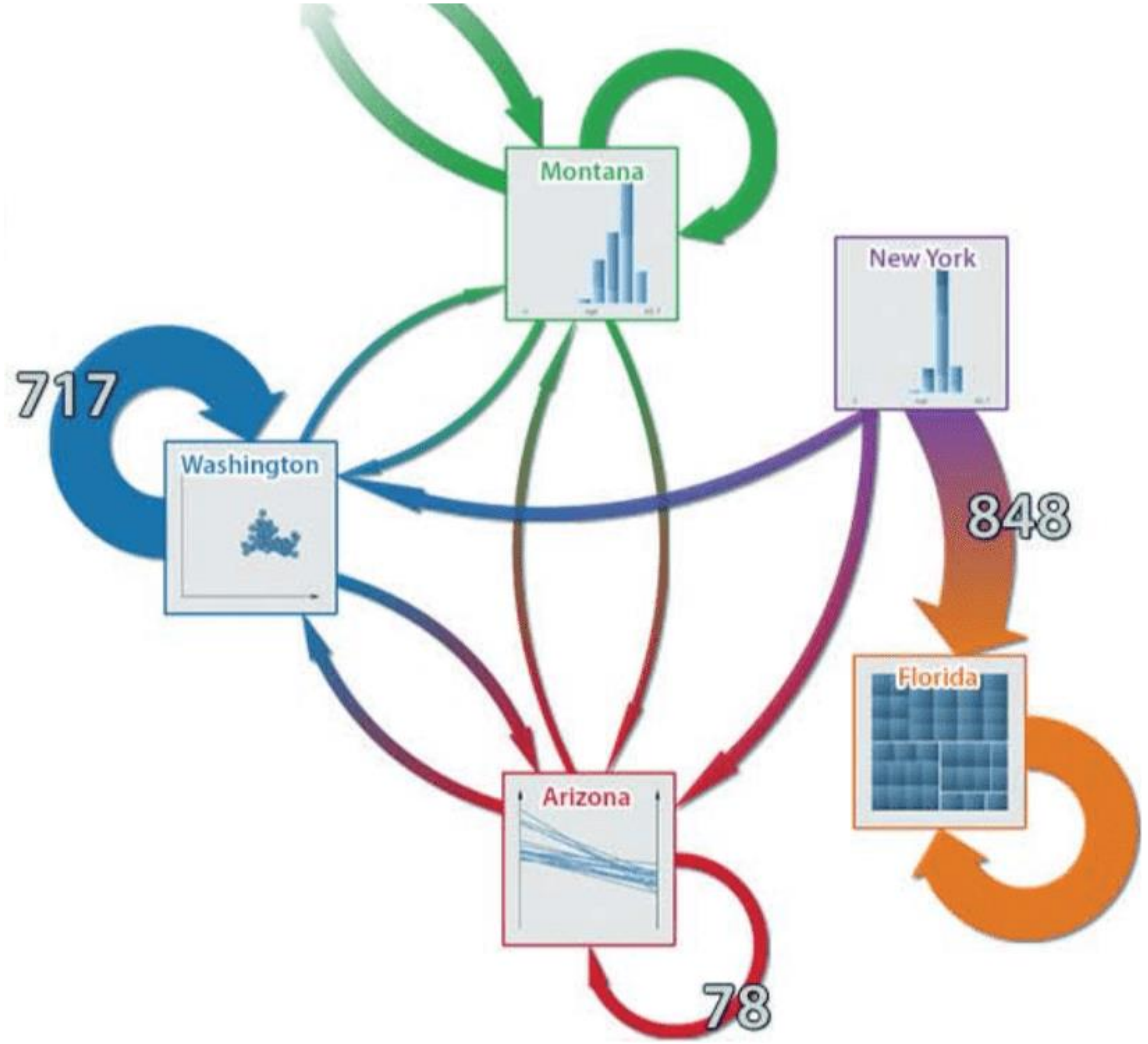



Sue
Abby
Mother
Jon
Dating
Friends
Married
Tom
Friends
Mark
Co-workers
Soccer
Coach
Cole

On-Node / On-Edge Encoding

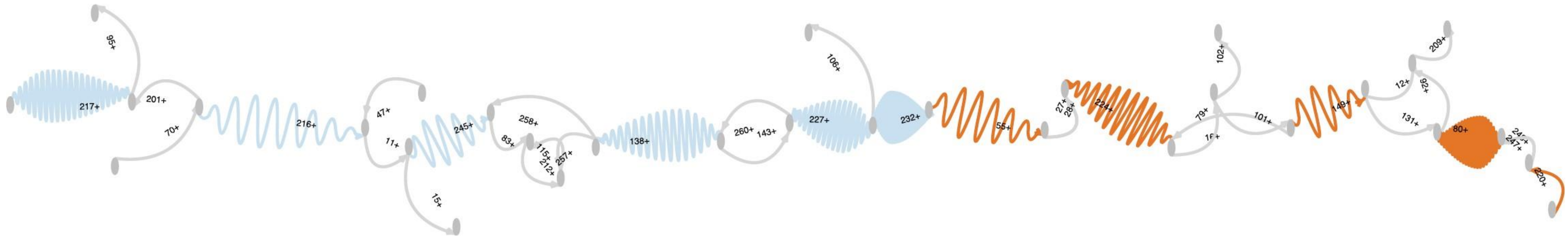


On-Node / On-Edge Encoding

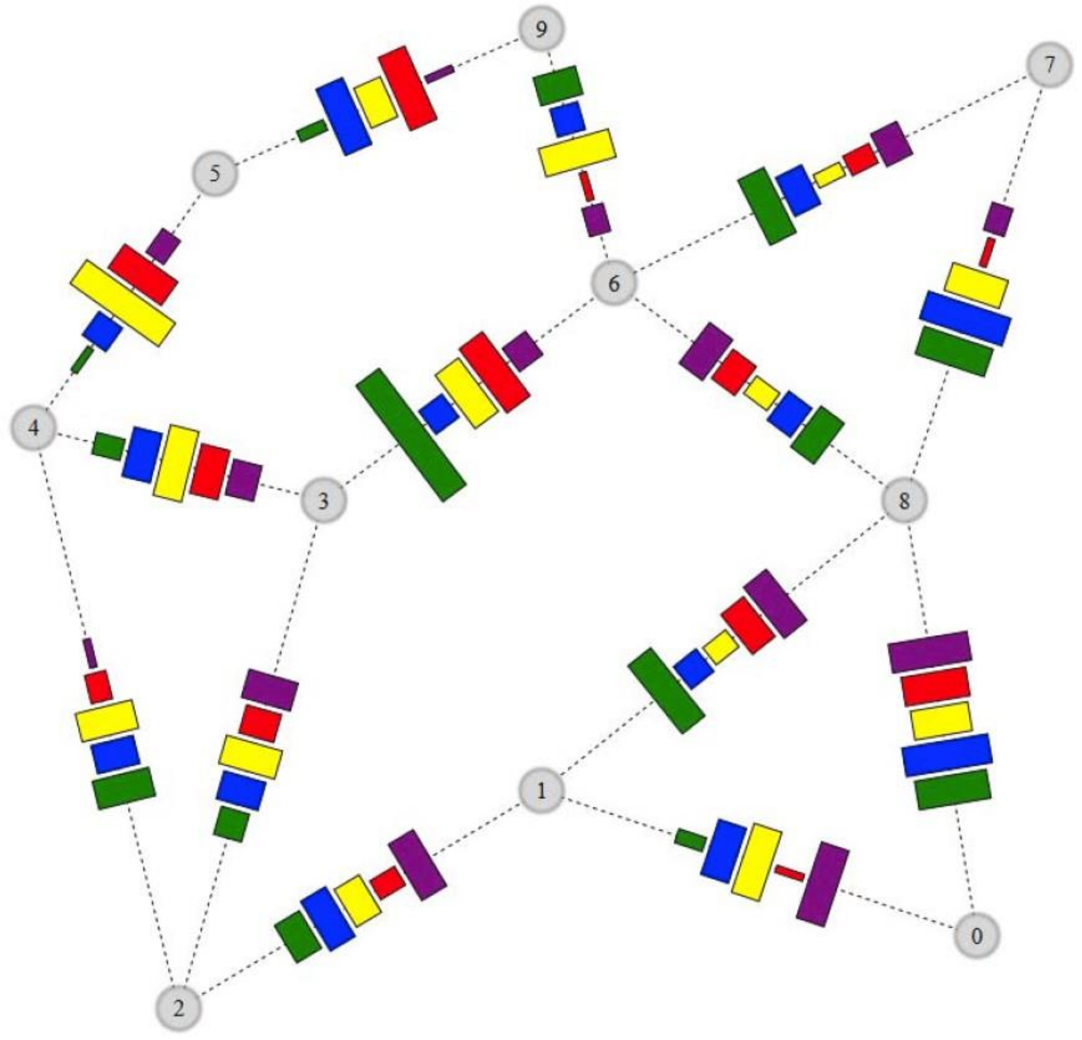


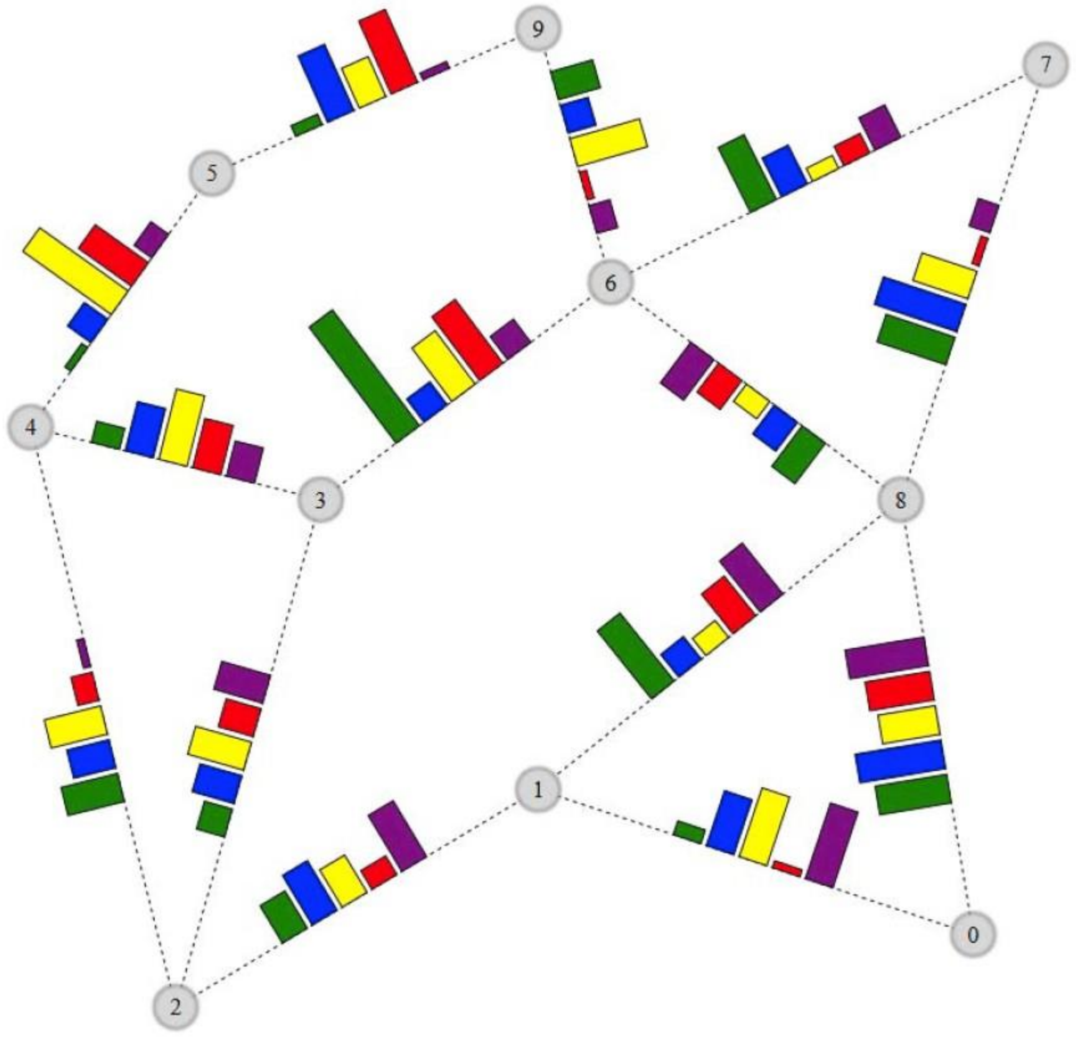
Elzen and Wijk, 2014

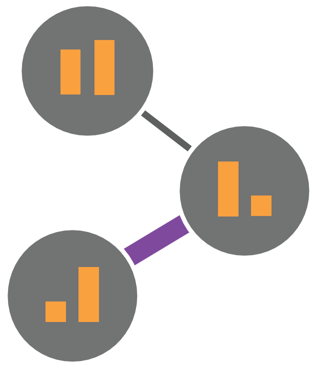
On-Node / On-Edge Encoding



On-Node / On-Edge Encoding







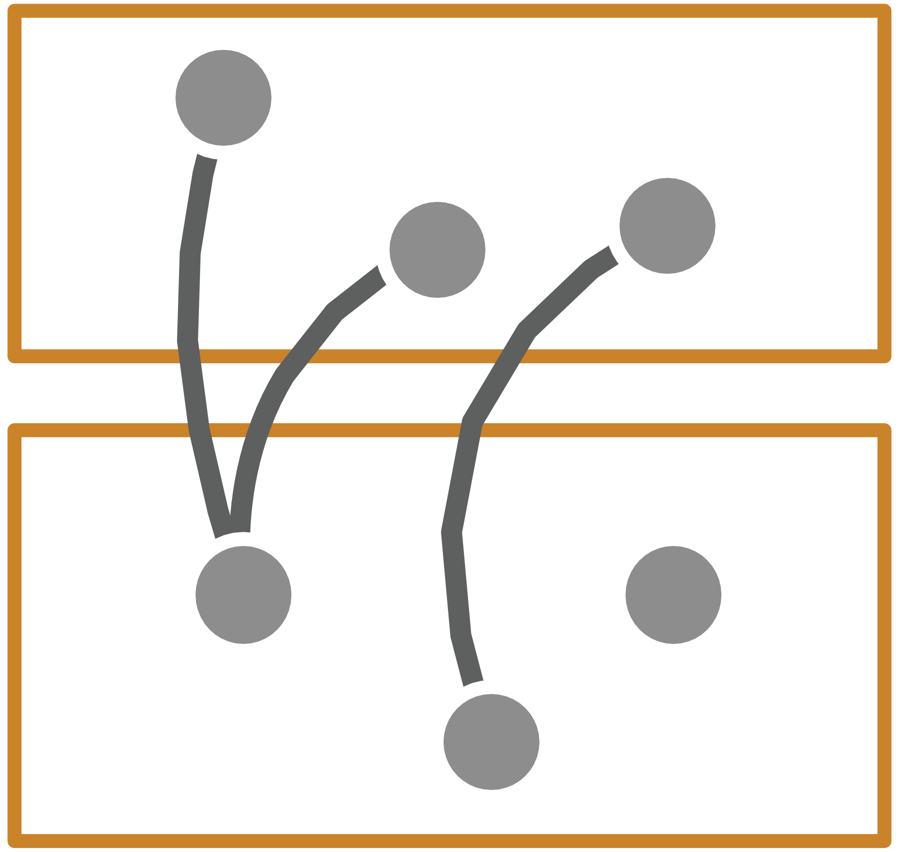
On-Node / On-Edge Encoding

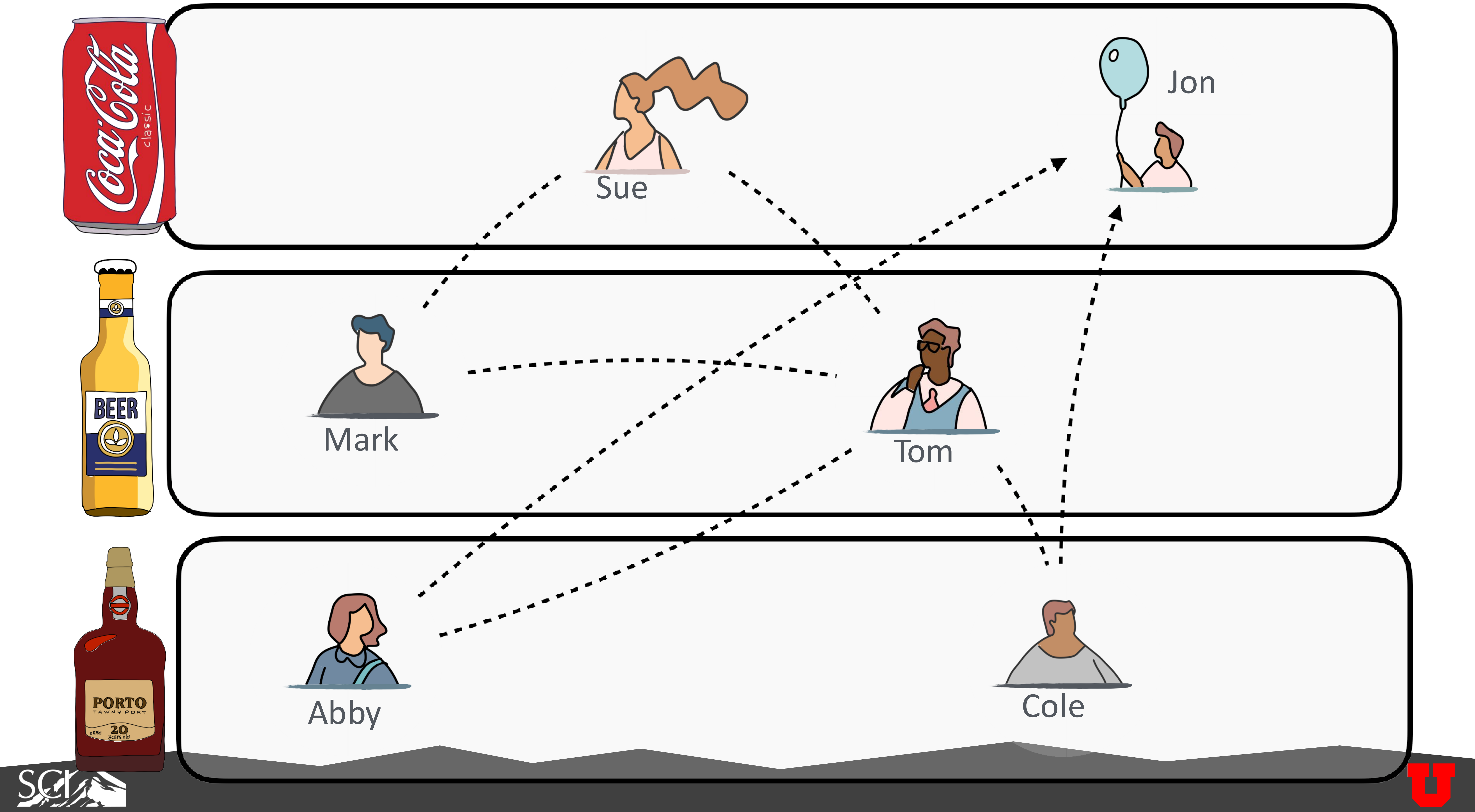
Is easily understood by most users
Works well for all types of networks

Scalability.
Node size leaves little space to encode attributes.

Recommended for small networks when only a few (usually under five) attributes on the nodes are shown, or in combination with a zooming/filtering strategy

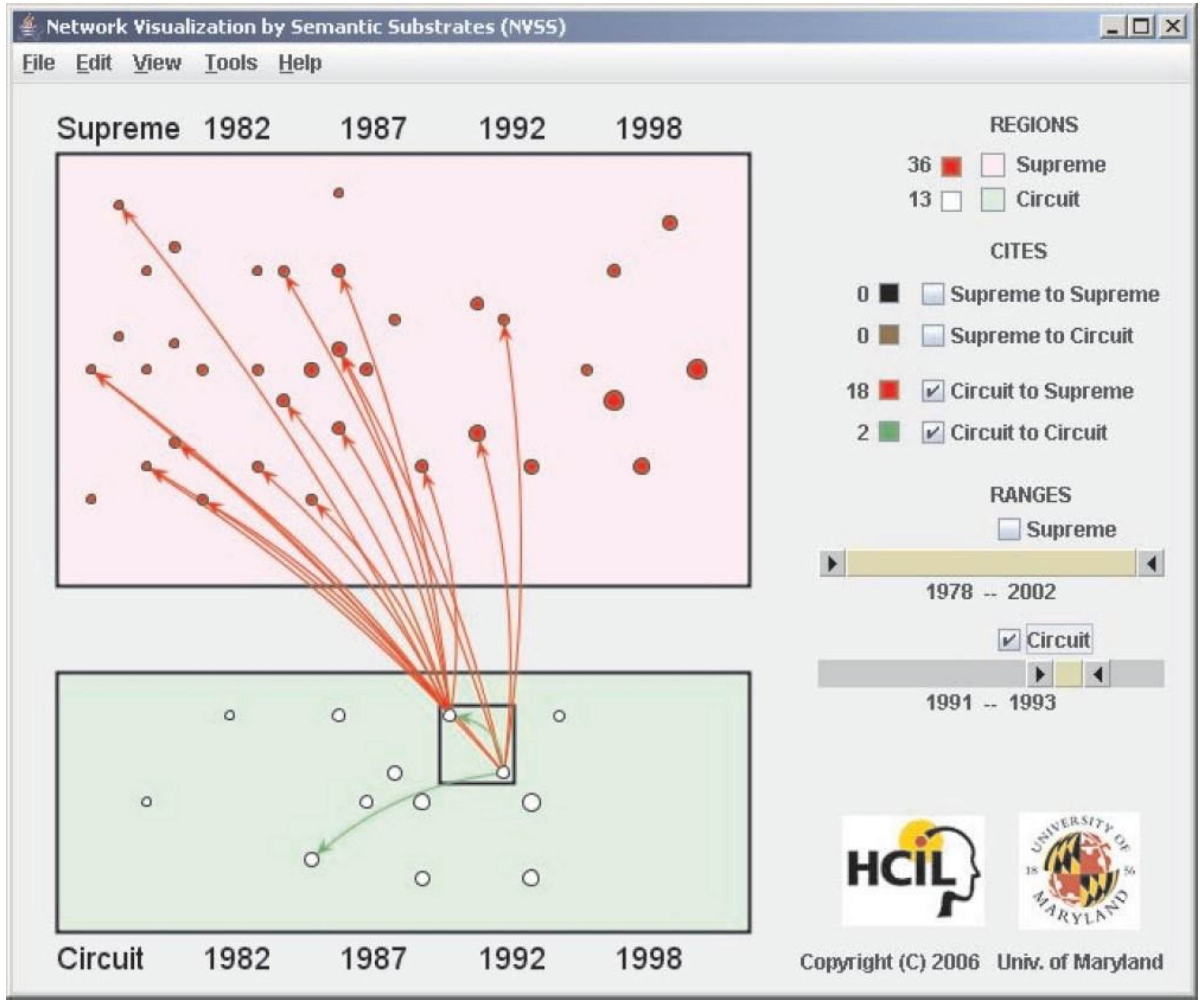
Attribute-Driven Faceting



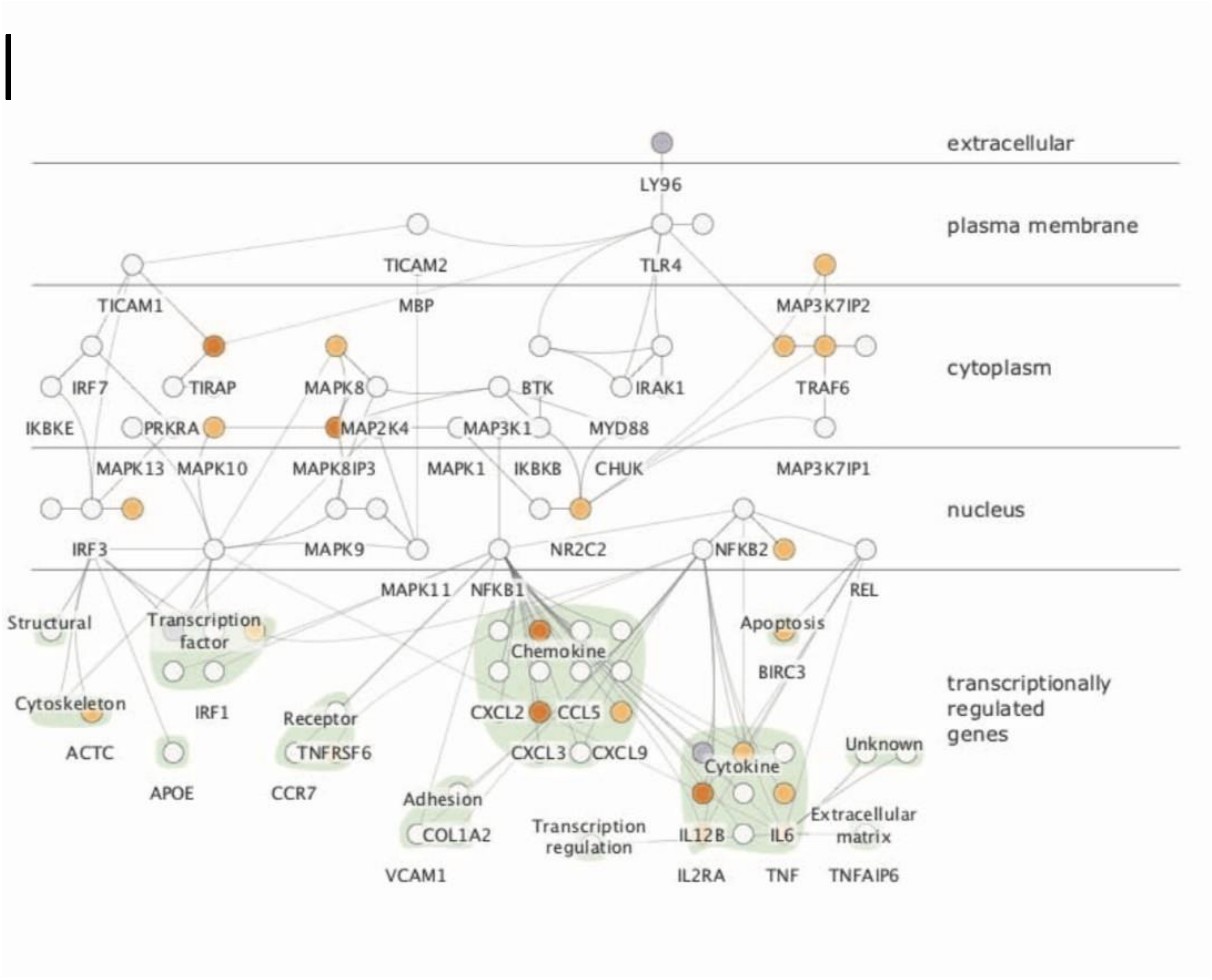


Coca-Cola classic
Sue
Jon
BEER
Mark
Tom
PORTO
TAWNY PORT
20 years old
Abby
Cole
SCI

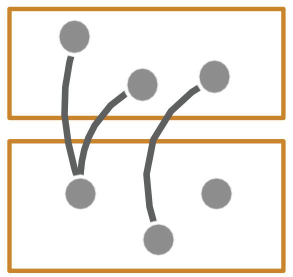
Semantic Substrates



Cerebral



Barskey et al. 2008



Attribute-Driven Faceting

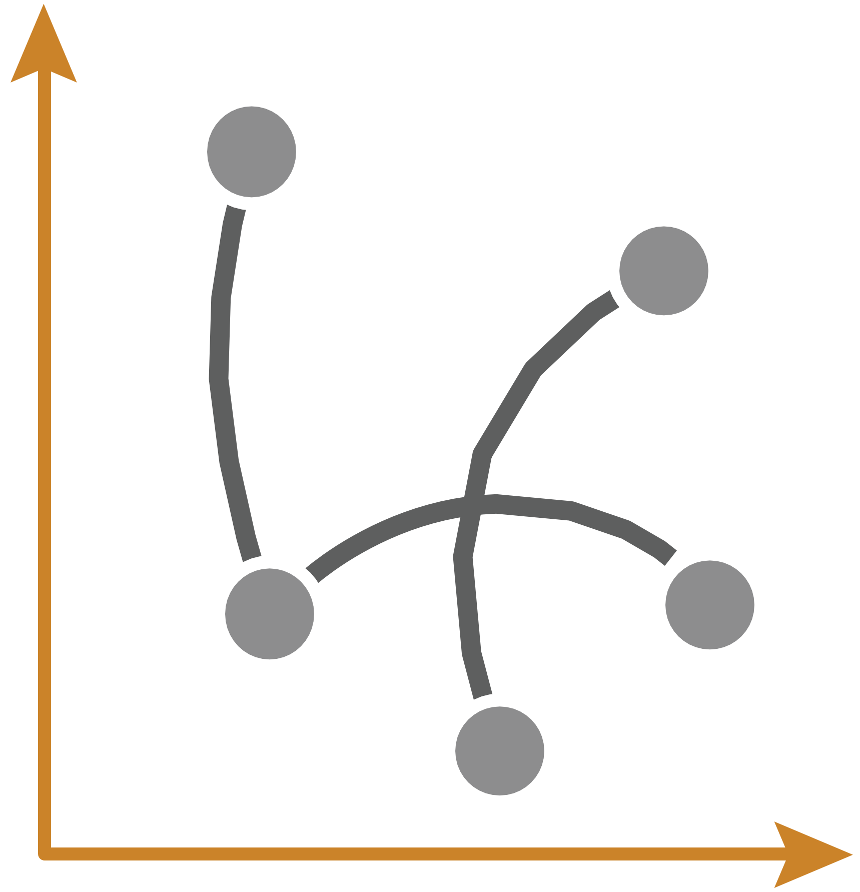
Well suited for networks with different node types or with an important categorical or set-like attribute.

Less scalable with respect to the number of nodes and network density than node-link layouts.

Neighborhoods, paths, and clusters are not easily visible if they span different facets.

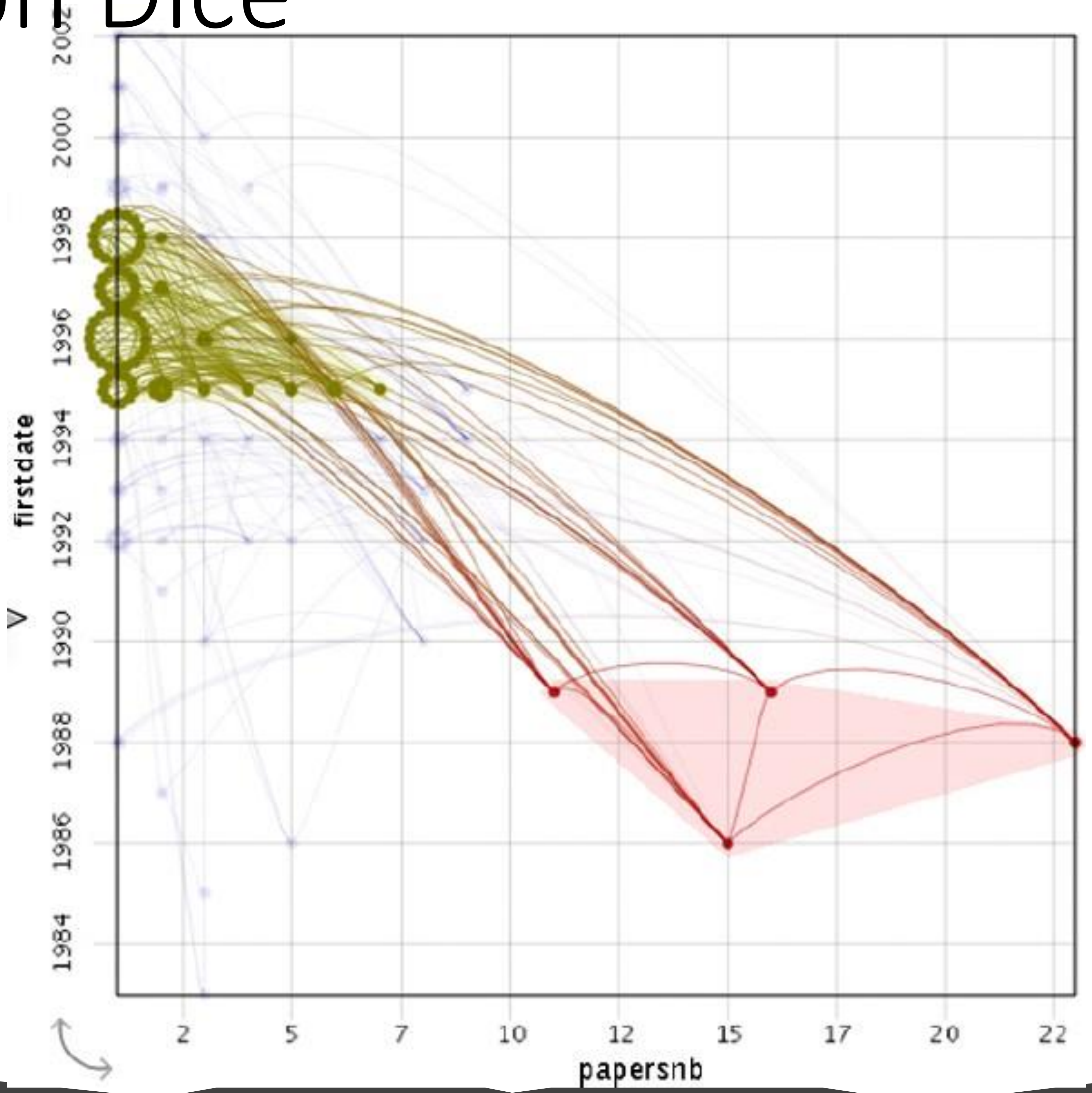
Recommended for networks where nodes can be separated into groups easily and where these groups are central to the analysis

Attribute-Driven Positioning



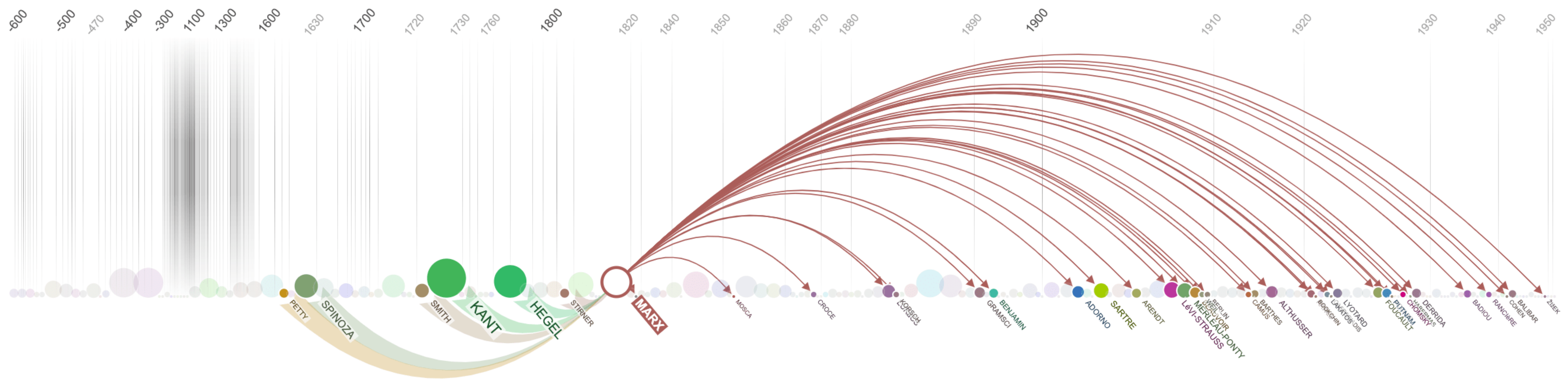
ANCHORAGE
VANCOUVER
EDMONTON
SEATTLE
PORTLAND
SAN FRANCISCO
DENVER
MINNEAPOLIS / ST. PAUL
CHICAGO
KANSAS CITY
DALLAS
TORONTO
CLEVELAND
MONTREAL
BALTIMORE
WASHINGTON D.C.
PHILADELPHIA
NEW YORK
JFK & NEWARK
BOSTON
HALIFAX
TAMPA BAY
ORLANDO
ICELAND
HELSINKI
STOCKHOLM
OSLO
GOTHENBURG
BERGEN
COPENHAGEN
BILLUND
BERLIN
HAMBURG
FRANKFURT
MUNICH
AMSTERDAM
GLASGOW
BRUSSELS
ZURICH
MANCHESTER
MILAN
LONDON
HEATHROW
& GATWICK
DUBLIN
PARIS
ORLY & CDG
GENEVA
MADRID

Graph Dice

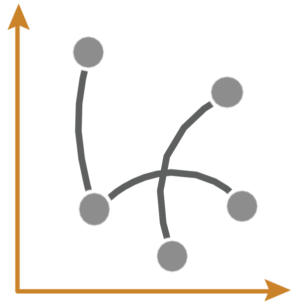


firstdate
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
papersnb
2
5
7
10
12
15
17
20
22

Edge Map



Dork et al. 2011



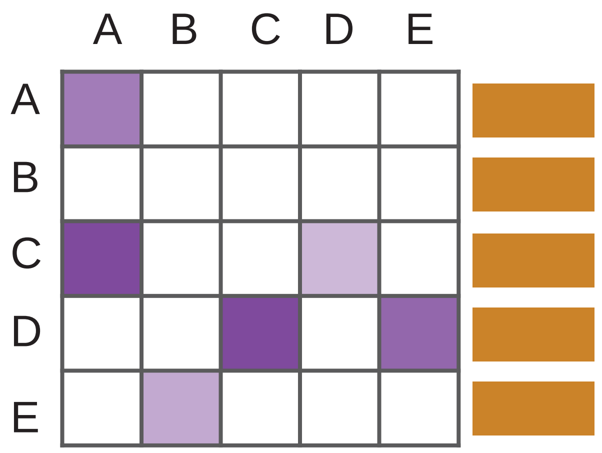
Attribute-Driven Positioning

Well suited for quantitative attributes

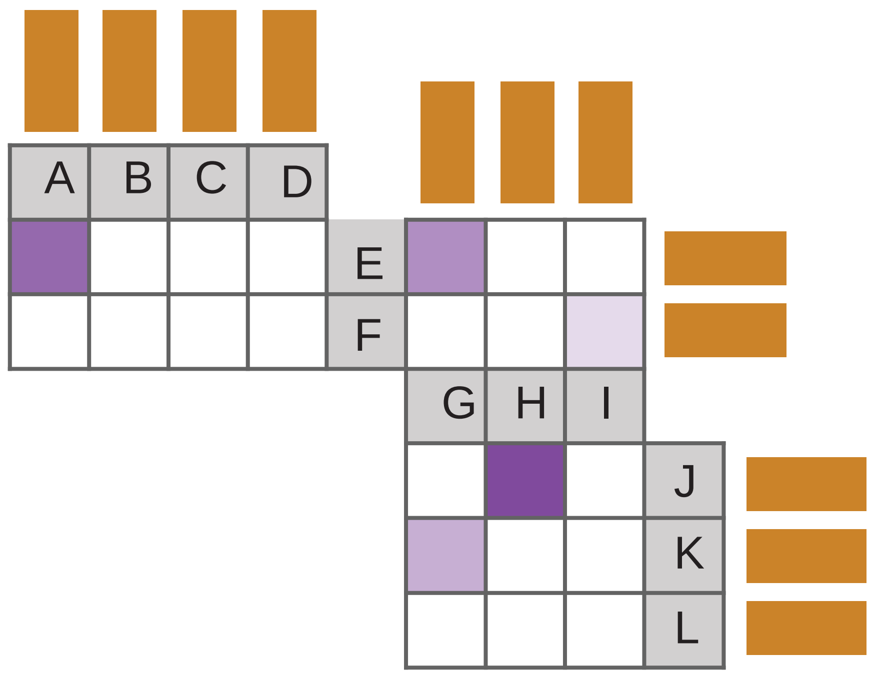
Does not lend itself well to visualizing the topology of the network.

Recommended for smaller, sparse networks where relationships between node attributes are paramount to the analysis task, and topological features only provide context

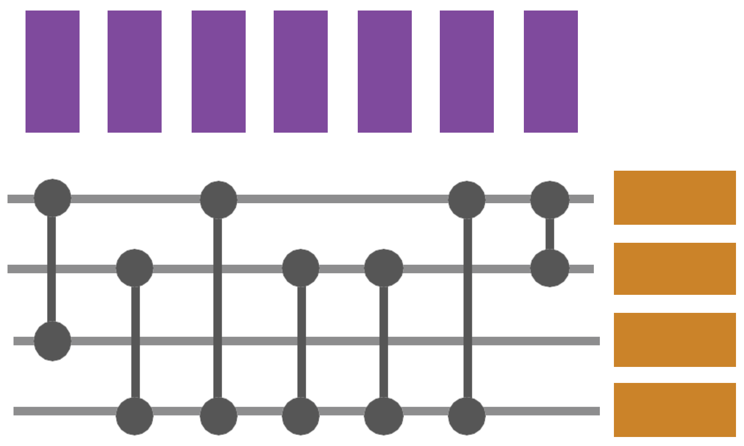
Tabular Layouts



Adjacency Matrix

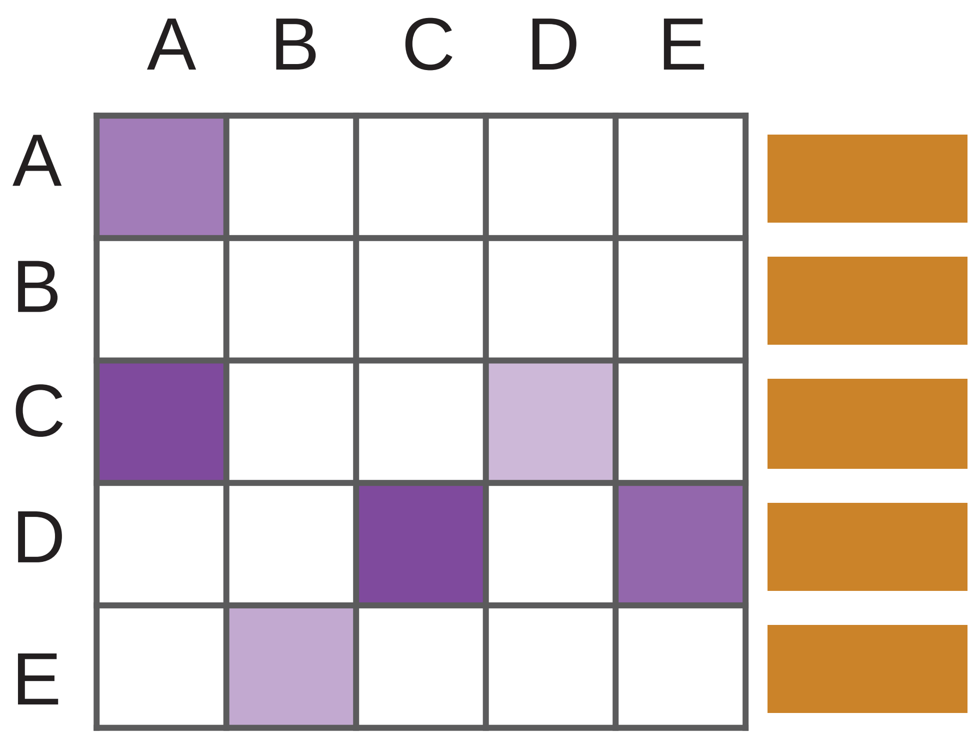


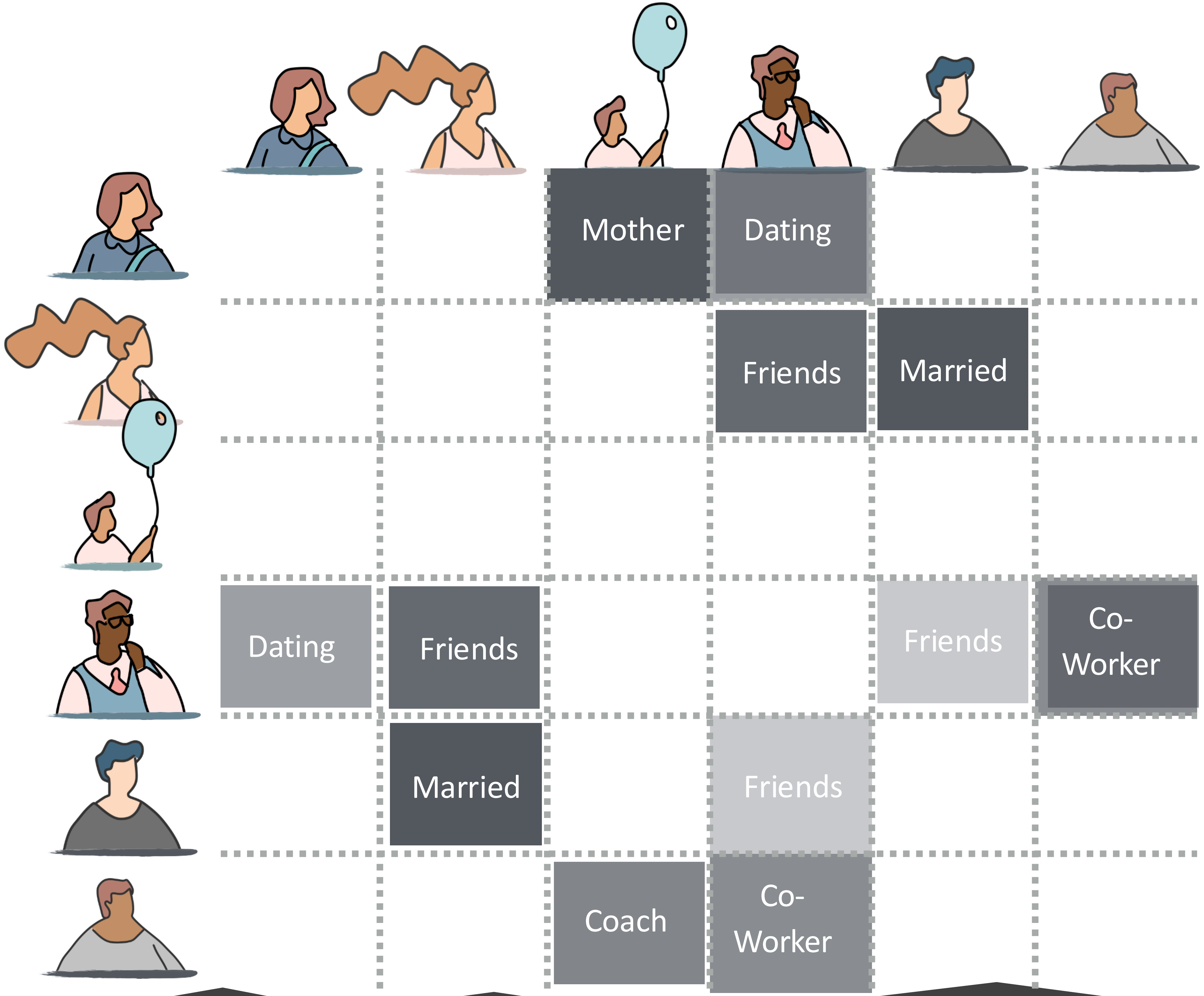
Quilts



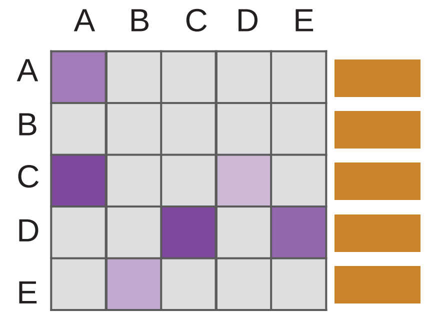
BioFabric

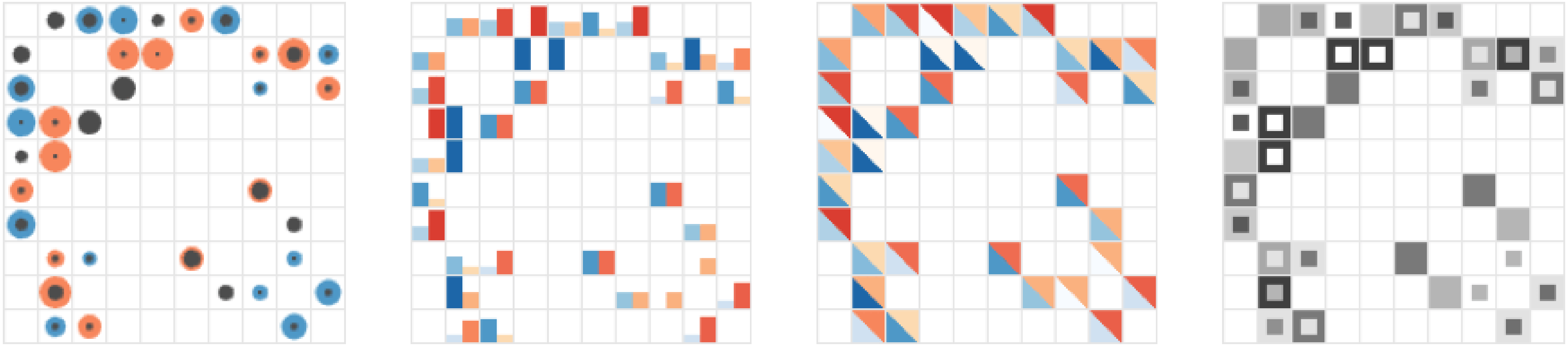
Adjacency Matrix




Mother
Dating
Friends
Married
Dating
Friends
Friends
Co-Worker
Married
Friends
Coach
Co-Worker

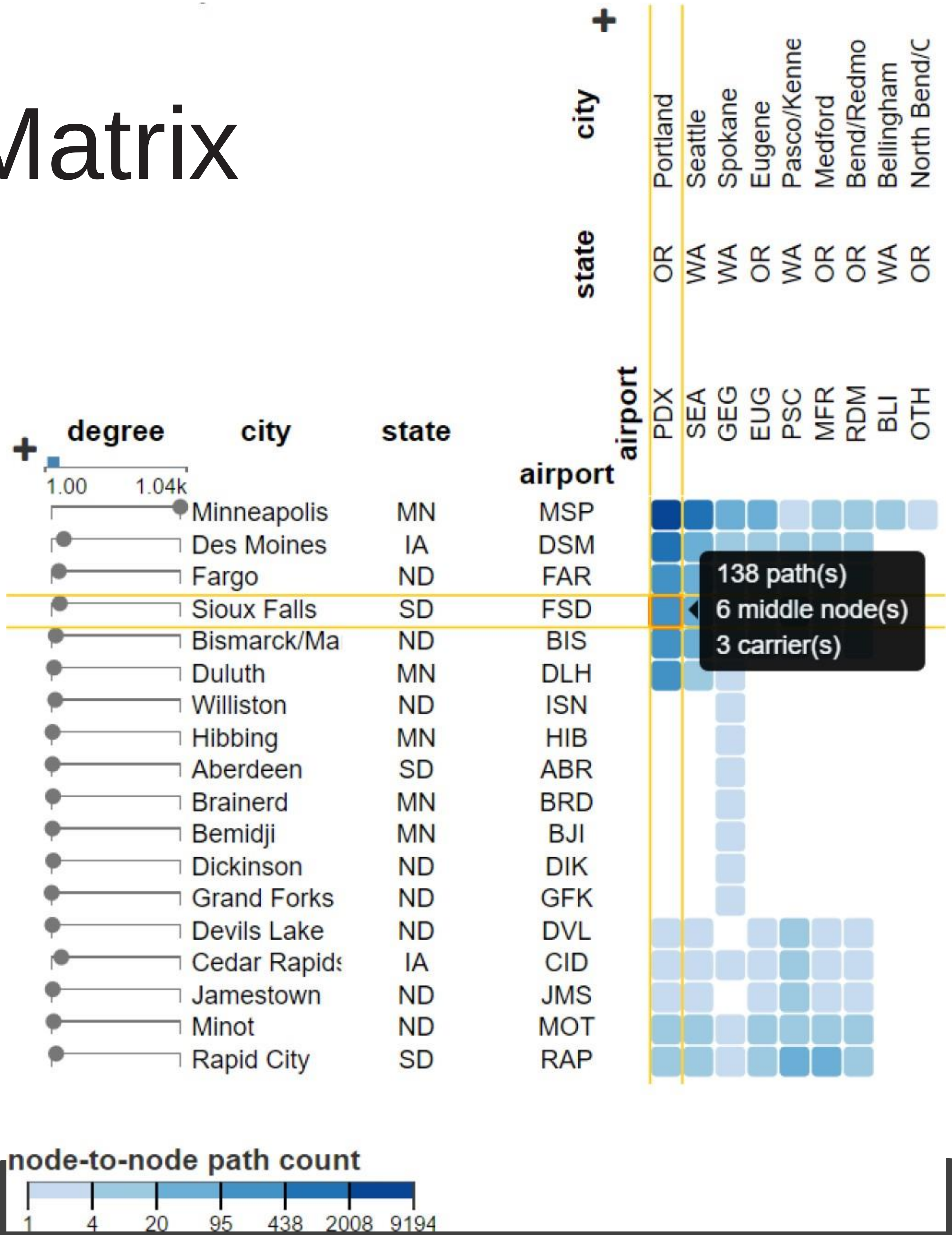
Adjacency Matrix

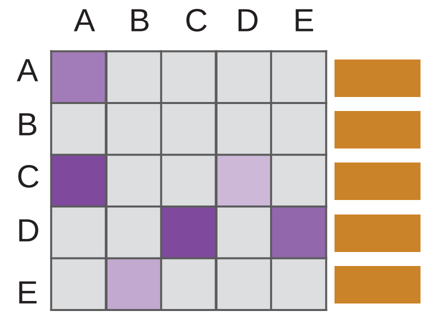




Alper et al, 2013

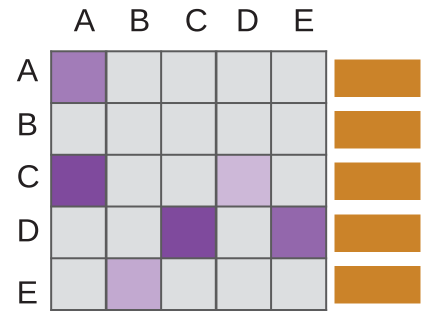
Adjacency Matrix

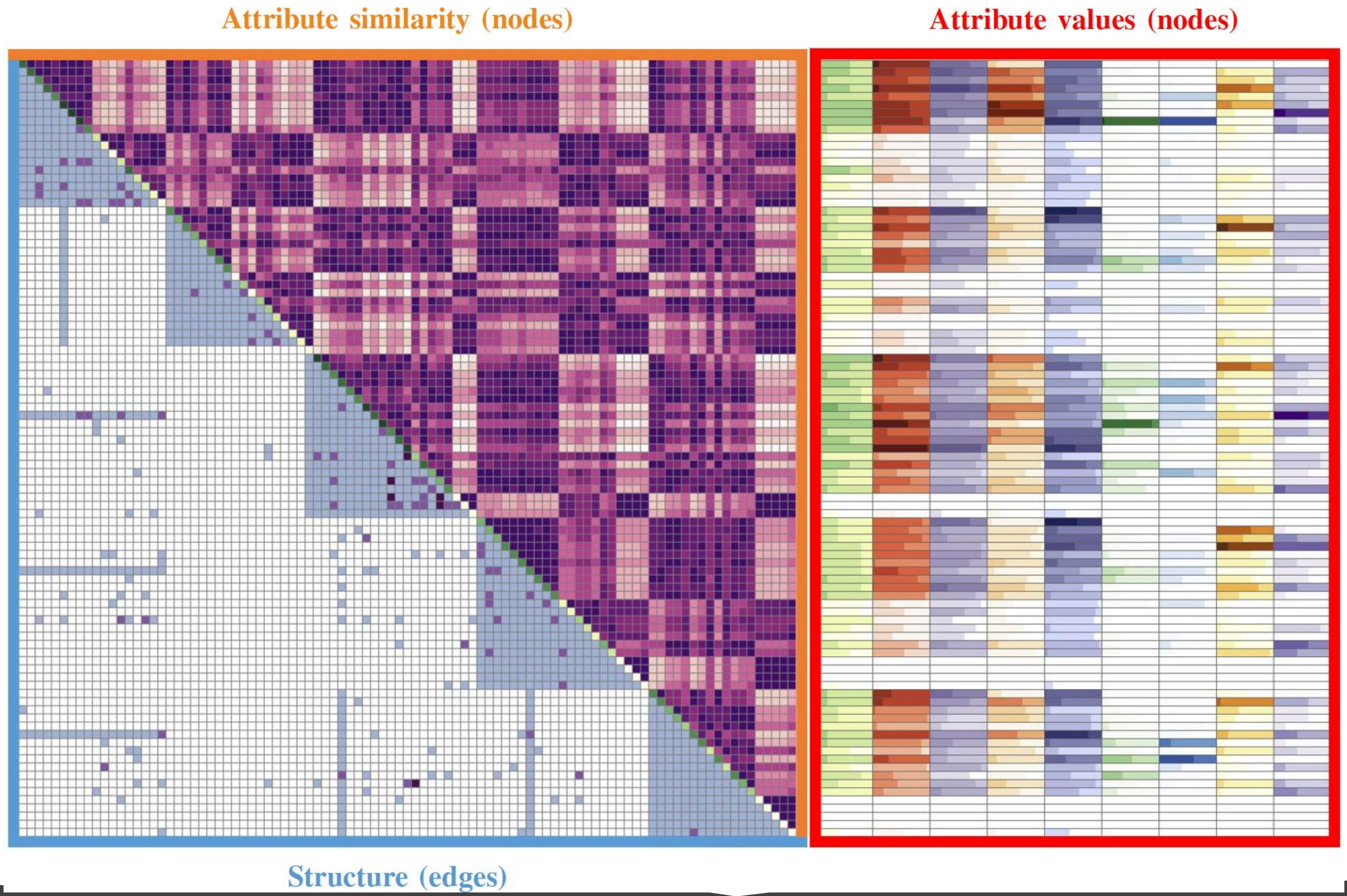


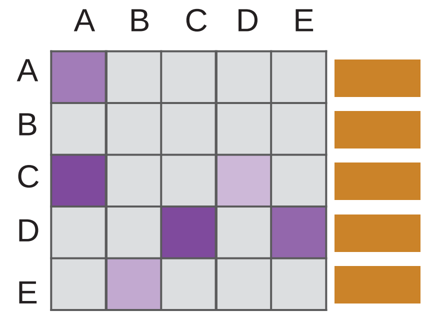


Kerzner et al, 2017

Adjacency Matrix







Adjacency Matrix

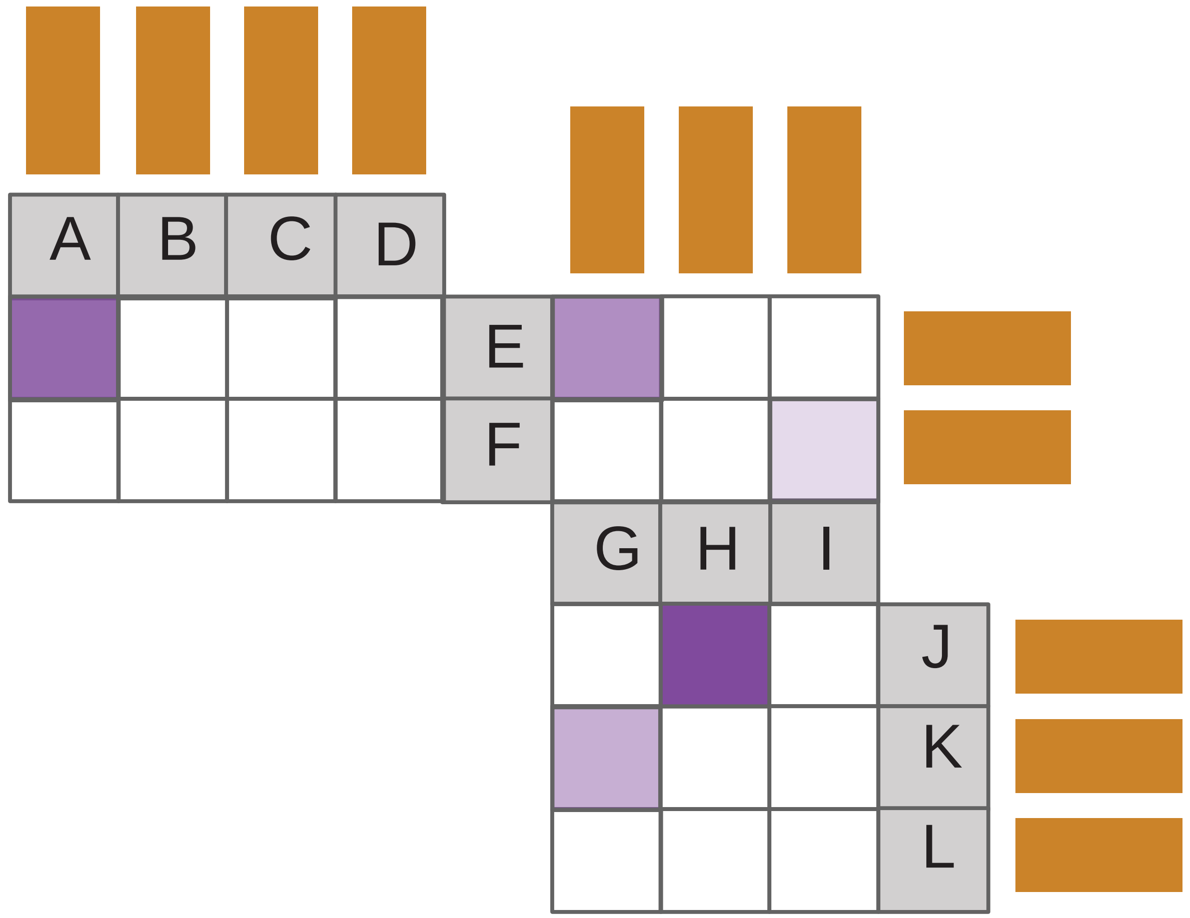
Ideal for dense and completely connected networks

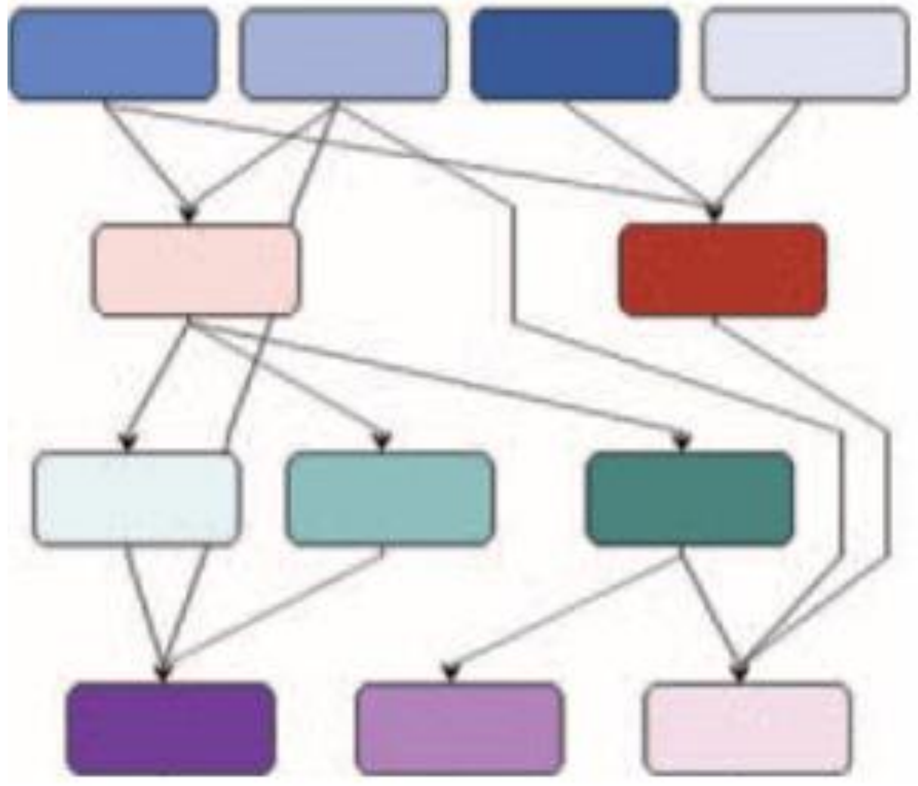
Requires quadratic space with respect to the number of nodes.

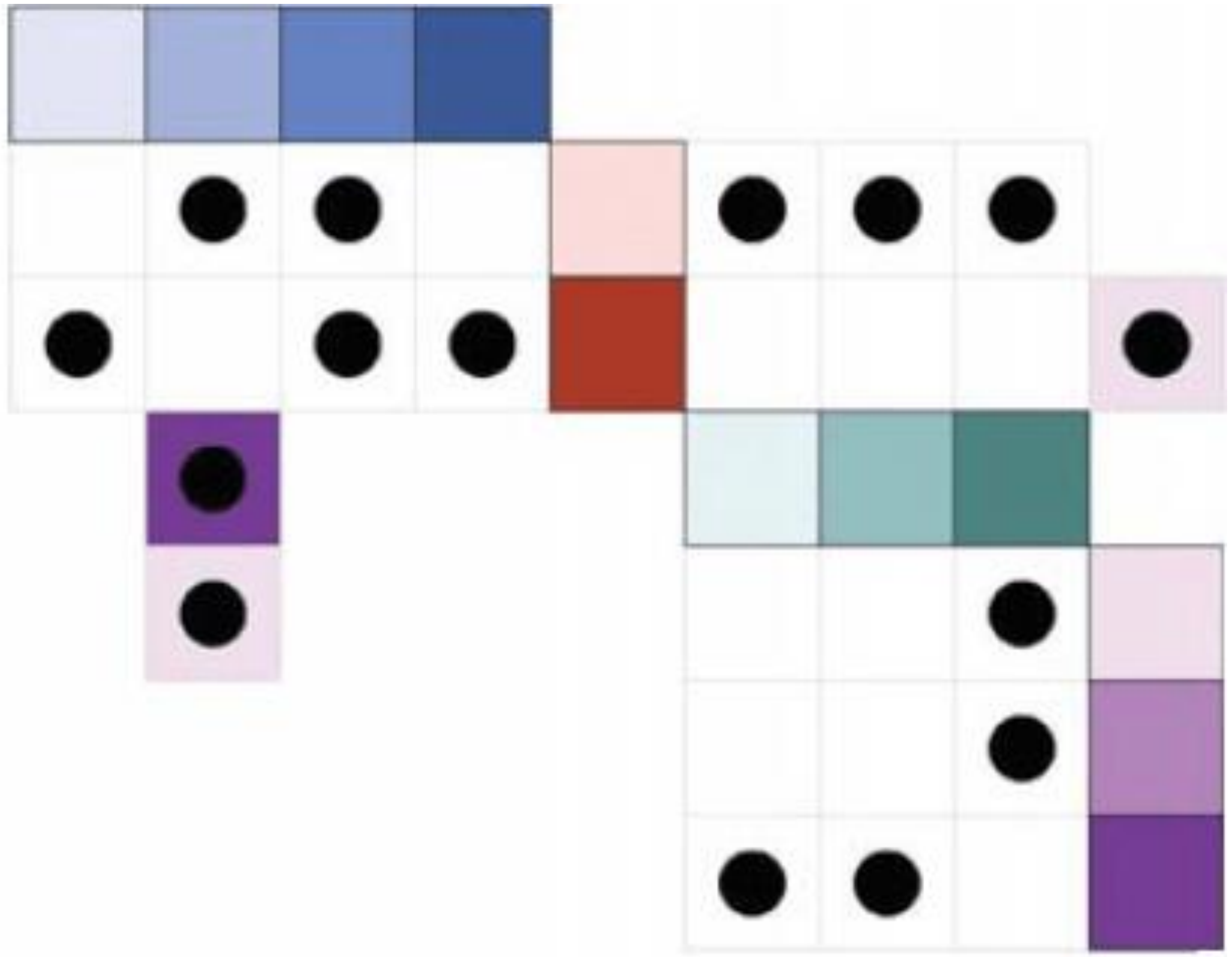
Complexity of choosing the right reordering algorithm

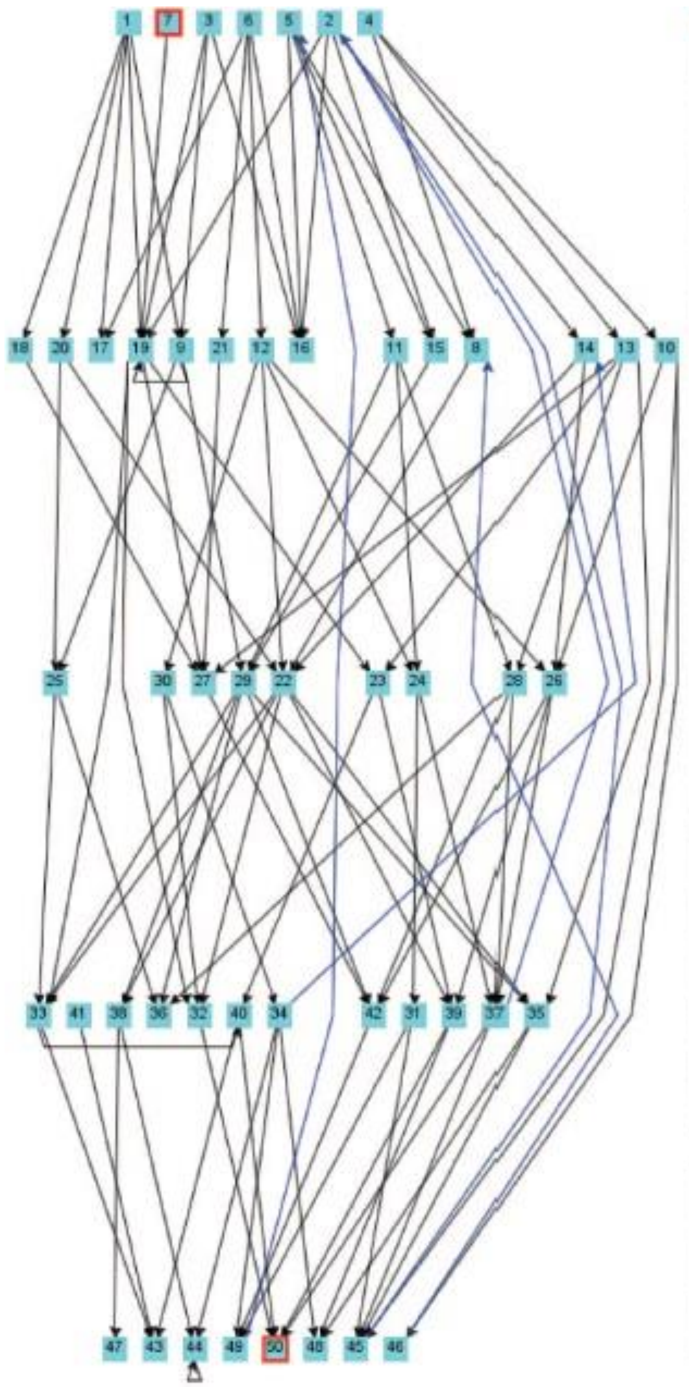
Recommended for smaller, complex and dense networks with rich node and/or edge attributes, for all tasks except for those involving paths

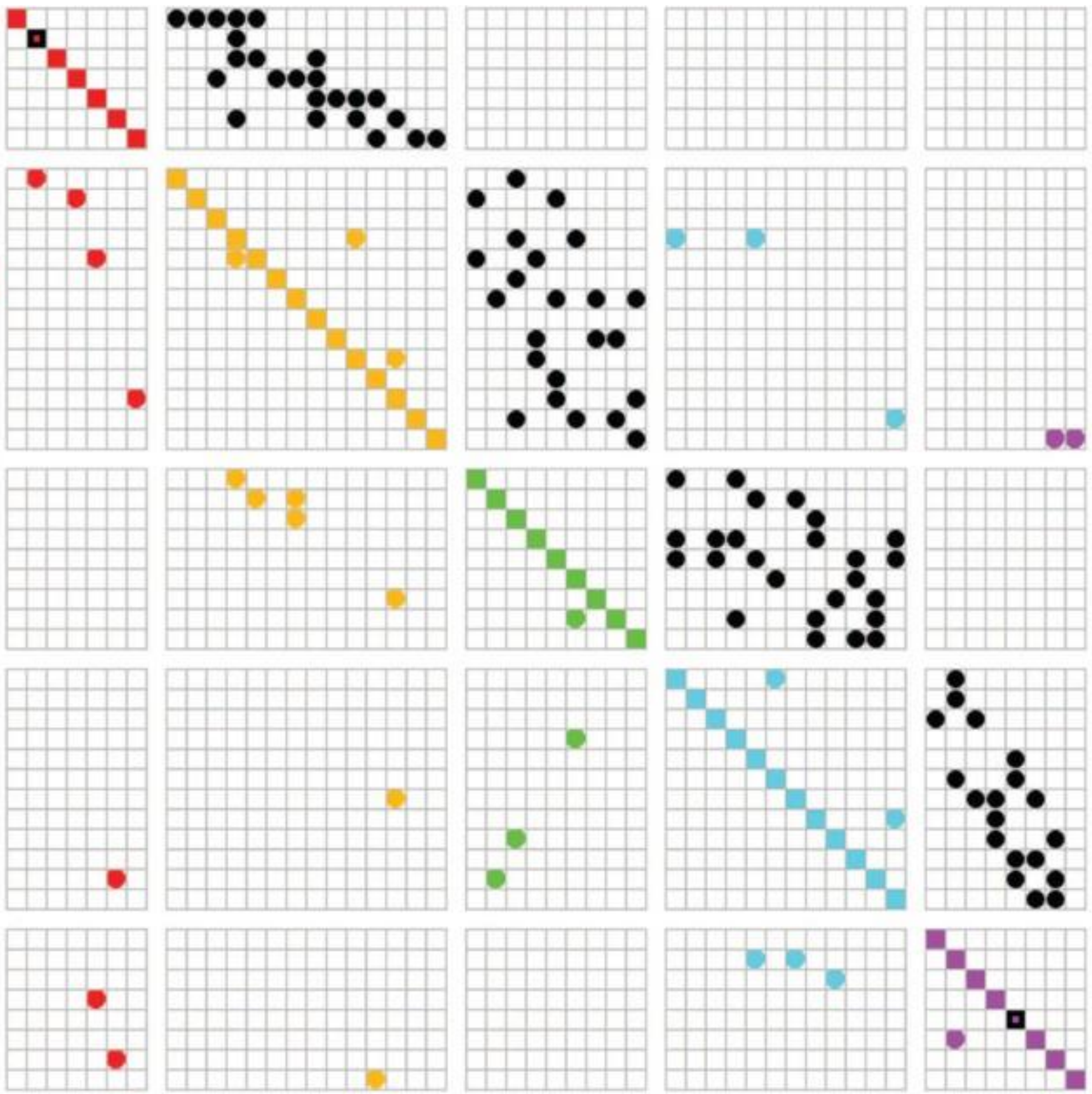
Quilts

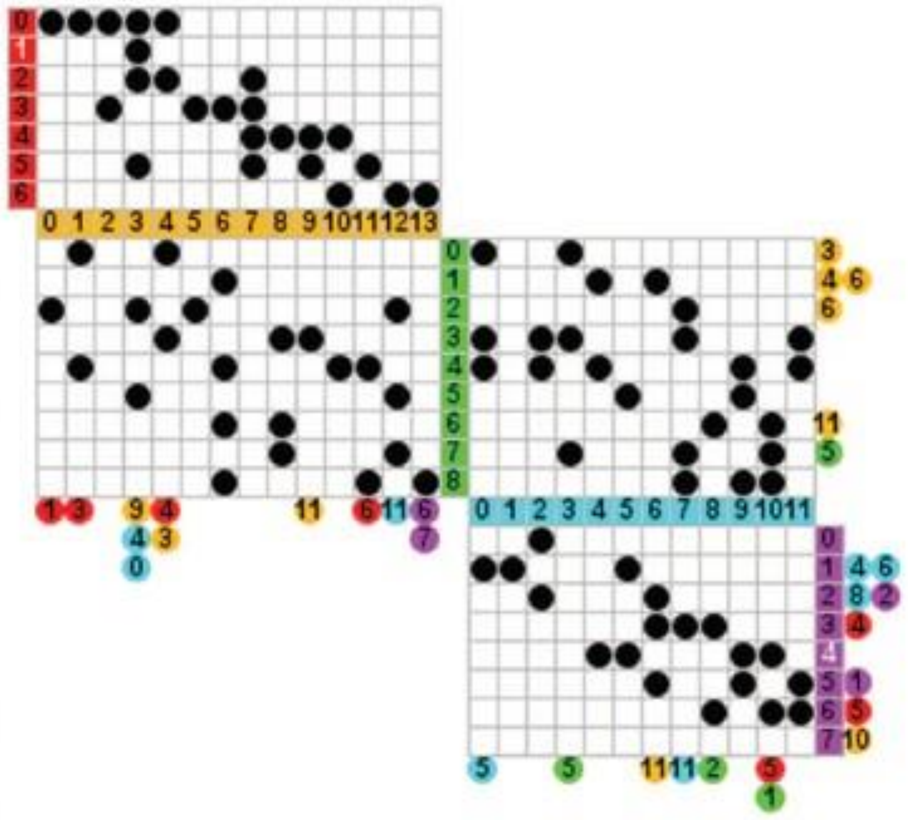


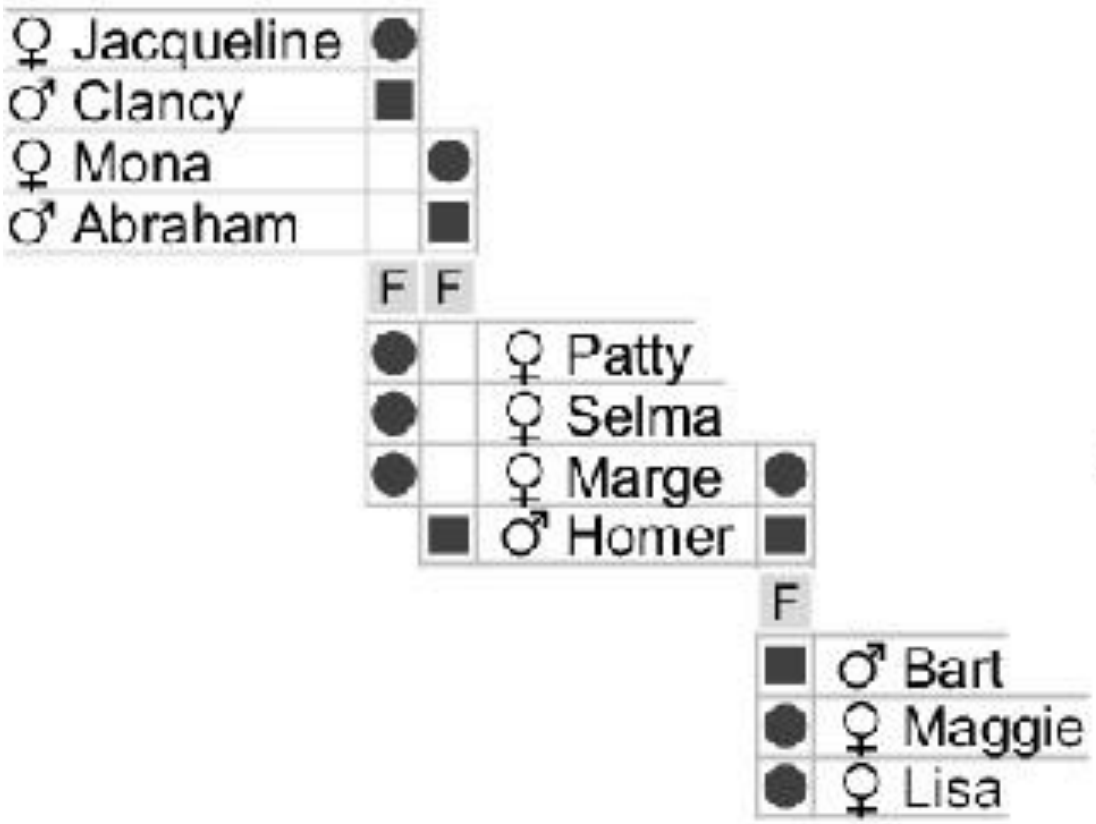


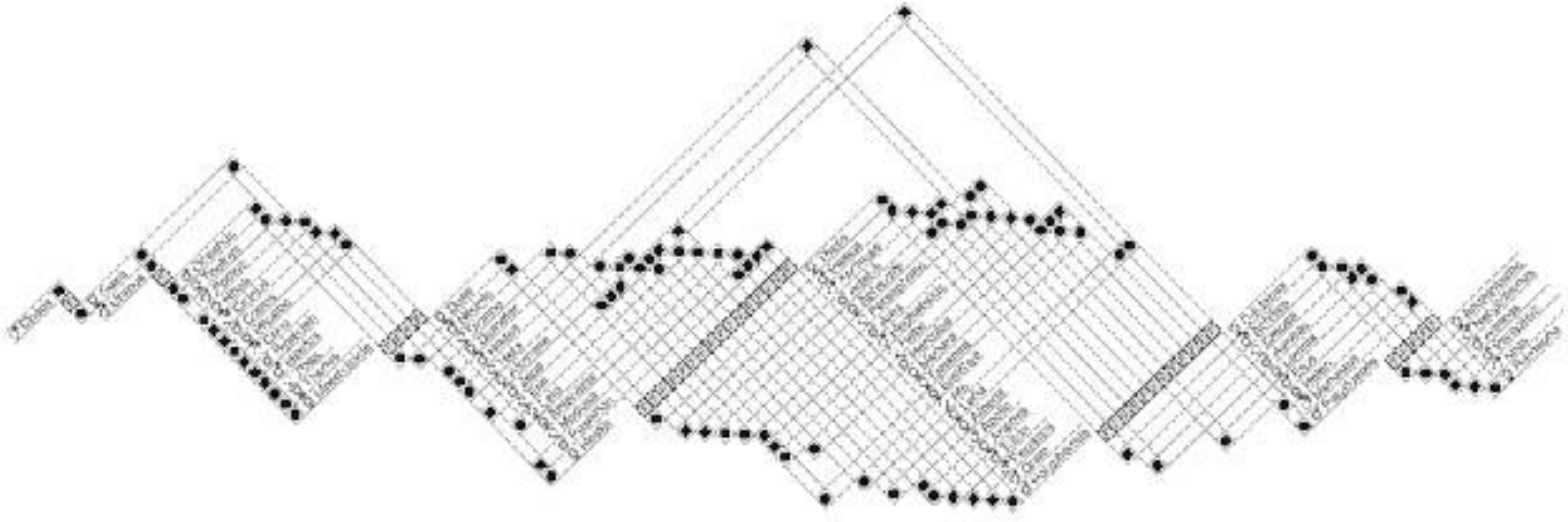









♀ Jacqueline
♂ Clancy
♀ Mona
♂ Abraham
F F
♀ Patty
♀ Selma
♀ Marge
♂ Homer
F
♂ Bart
♀ Maggie
♀ Lisa

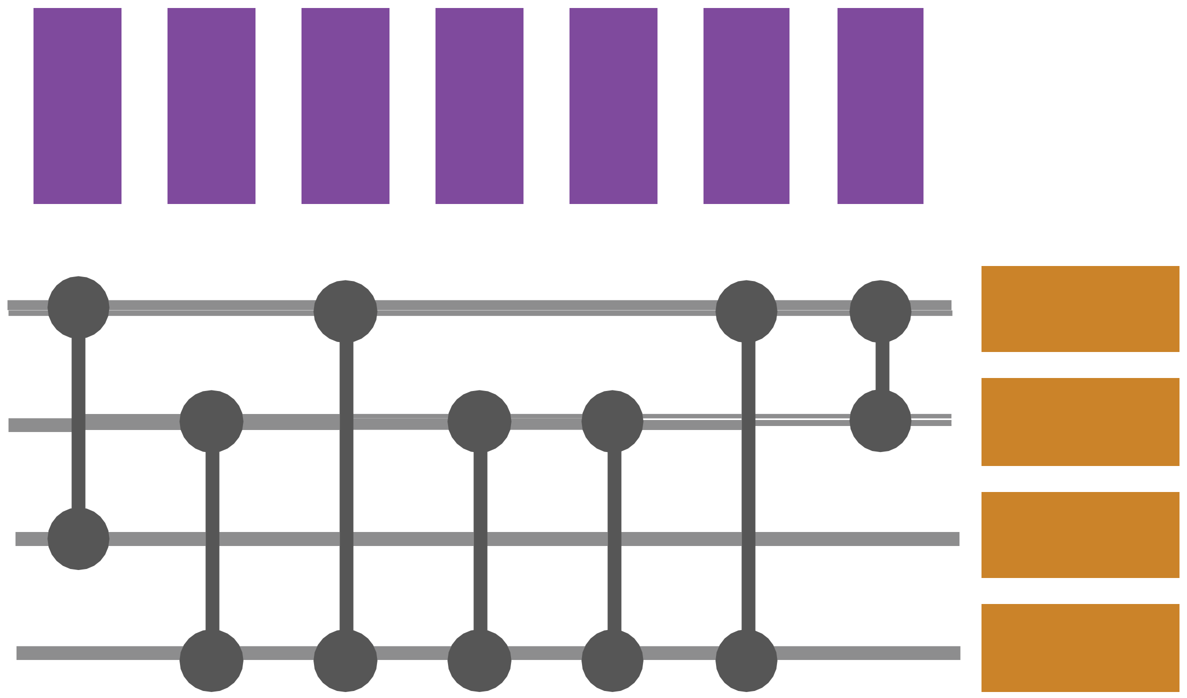


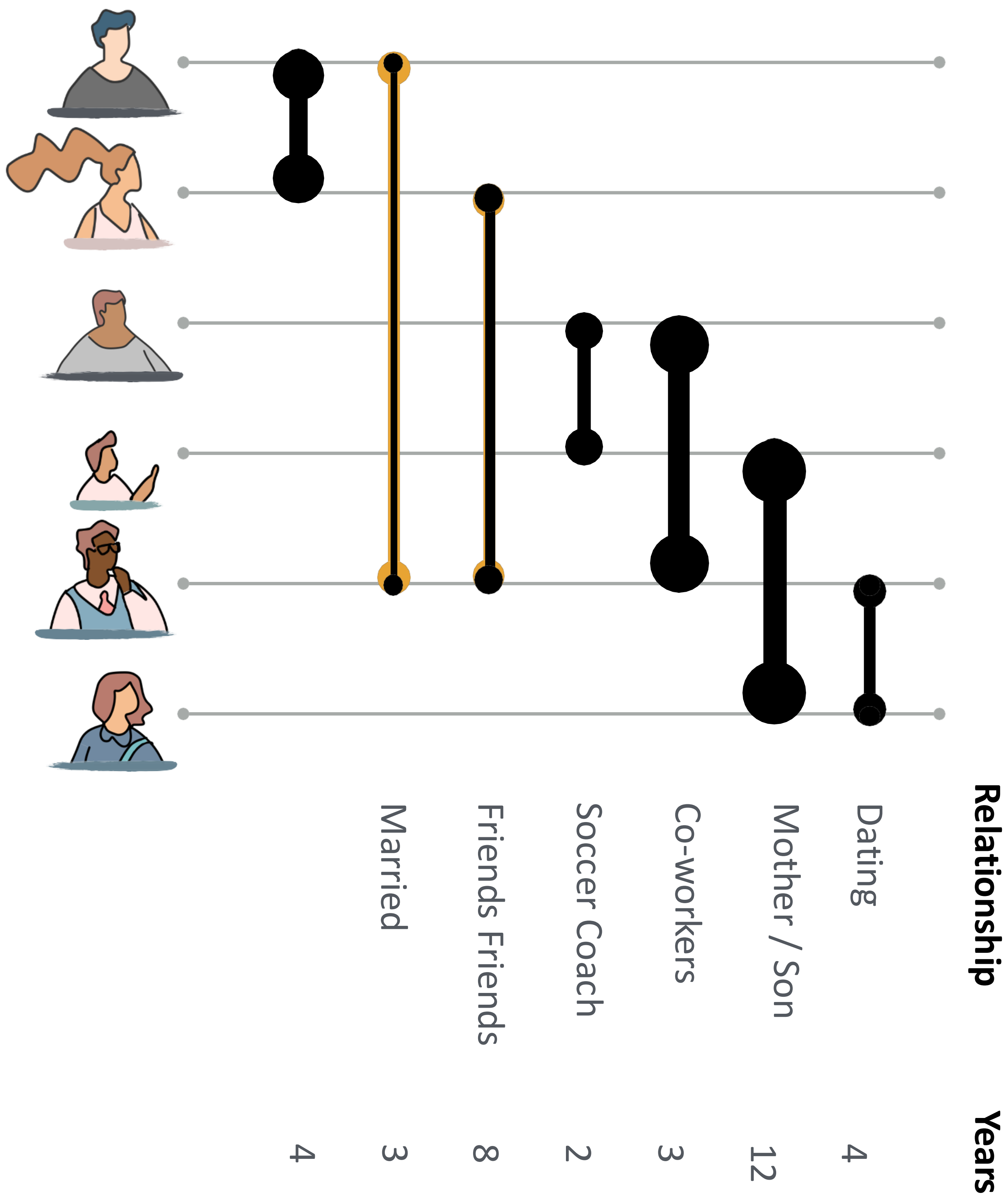
Well suited for layered networks

Links between nonconsecutive layers can be problematic to integrate and non-intuitive

Recommended for layered or k-partite networks with limited skiplinks.

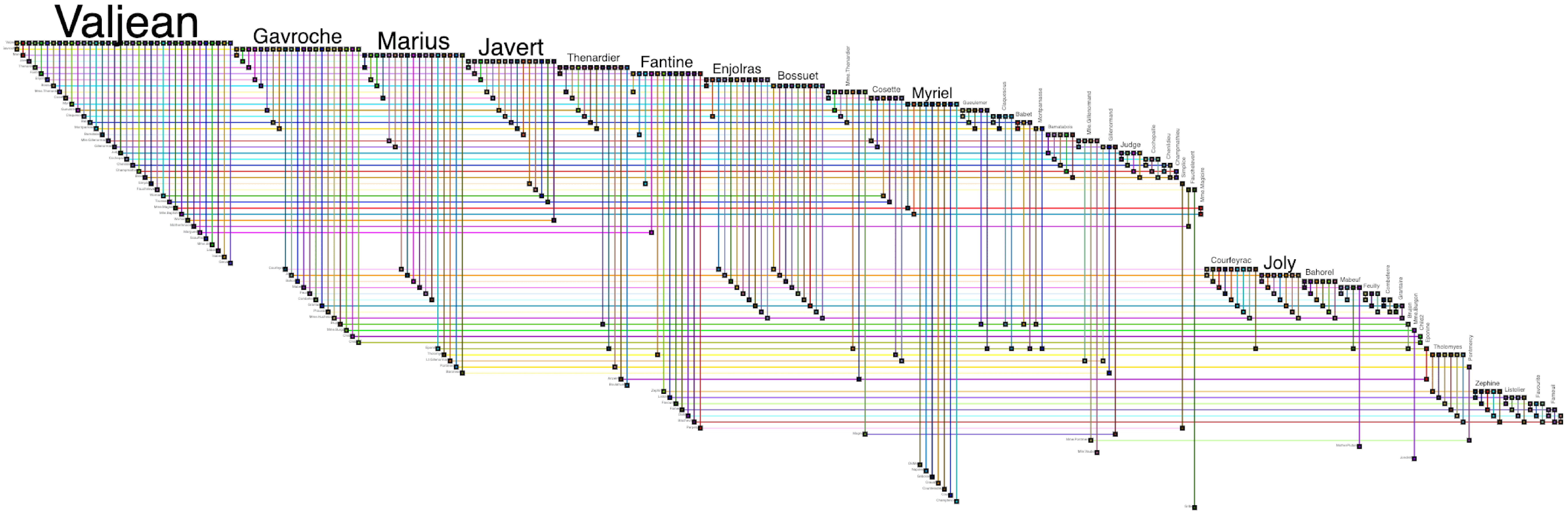
BioFabric



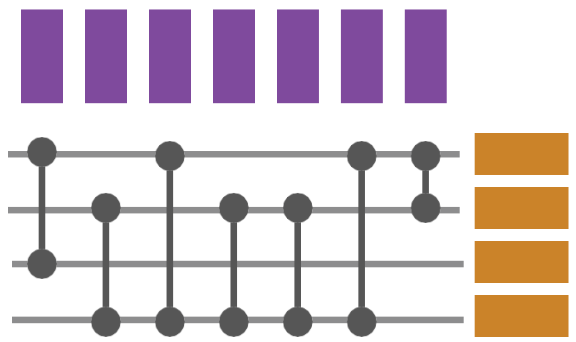


Name	Beverage	Day 1
Mark	Beer	1
Sue	Coke	0
Cole	Port	4
Jon	Coke	5
Tom	Beer	2
Abby	Port	3

BioFabric



Longabaugh, 2012



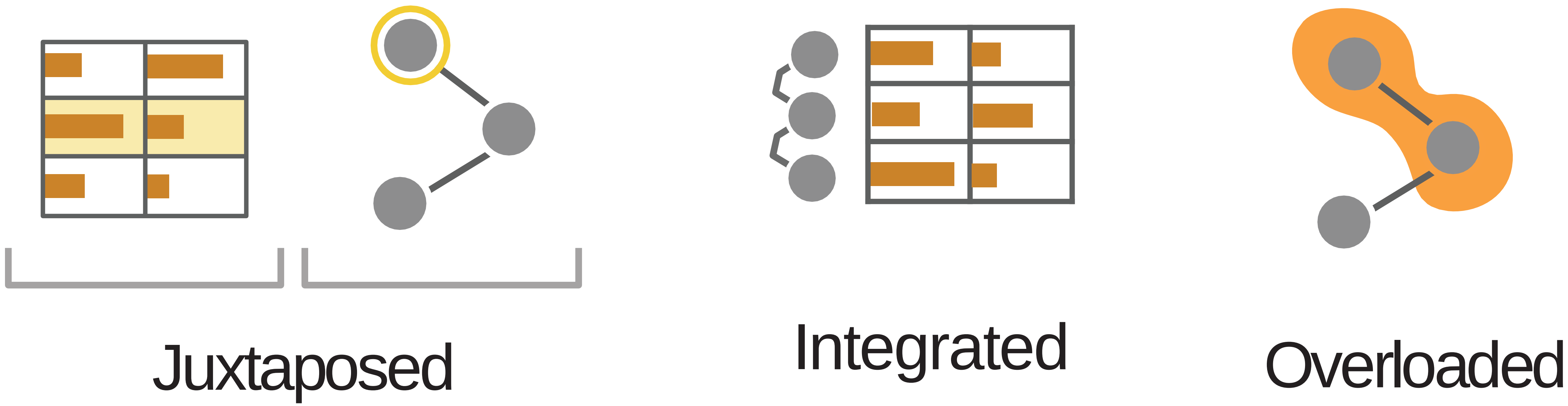
BioFabric

Can be used to visualize rich edge attributes and node attributes at the same time

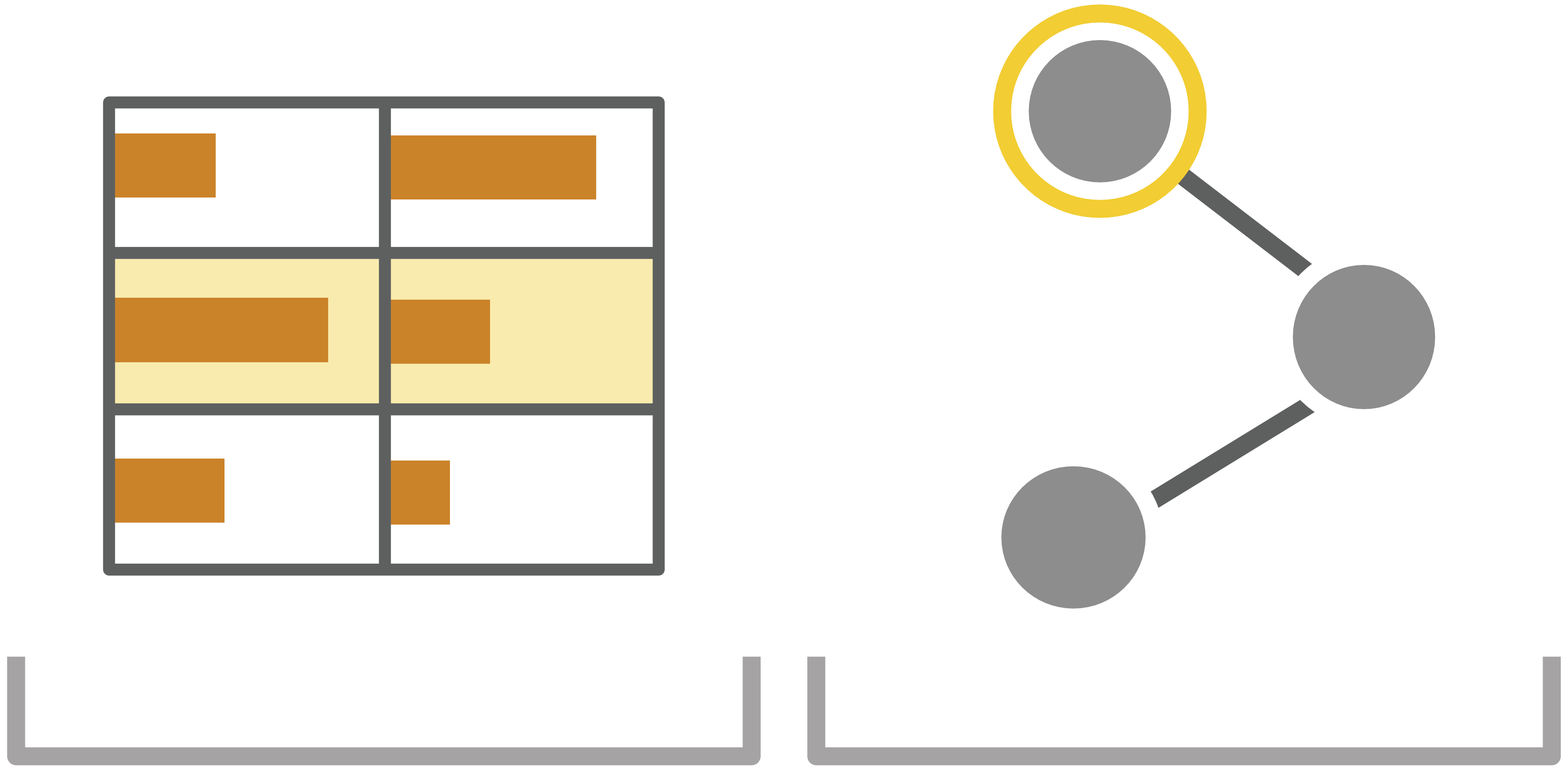
More difficult to discover neighbors and clusters in Biofabric compared to matrices.

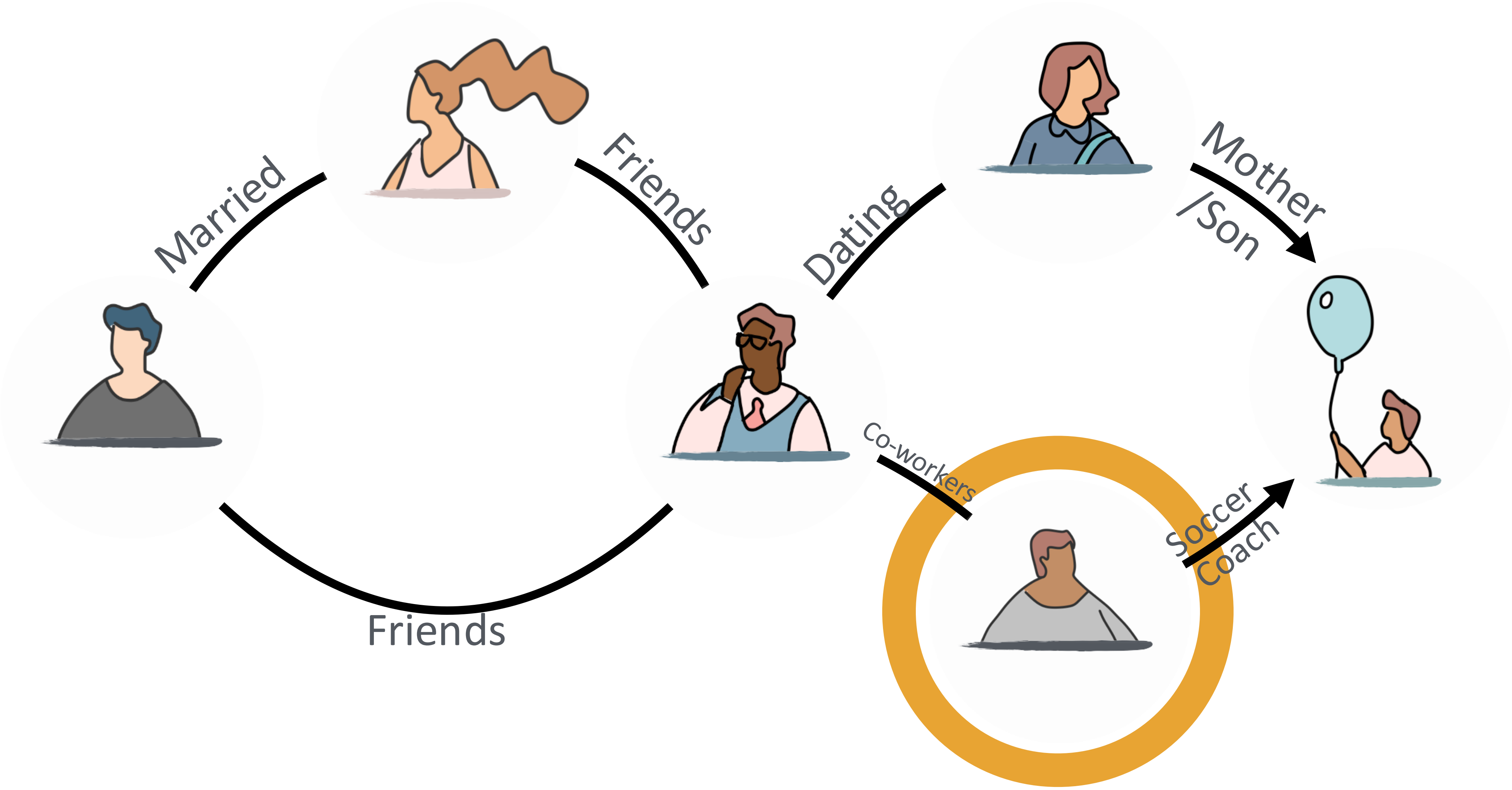
Recommended for small, sparse networks with many nodes and rich edge attributes

View Operations



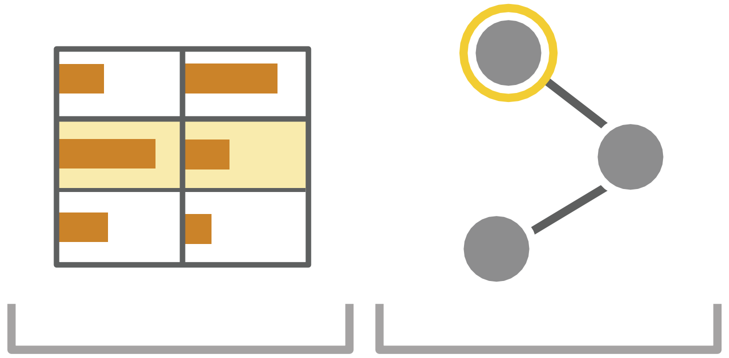
Juxtaposed

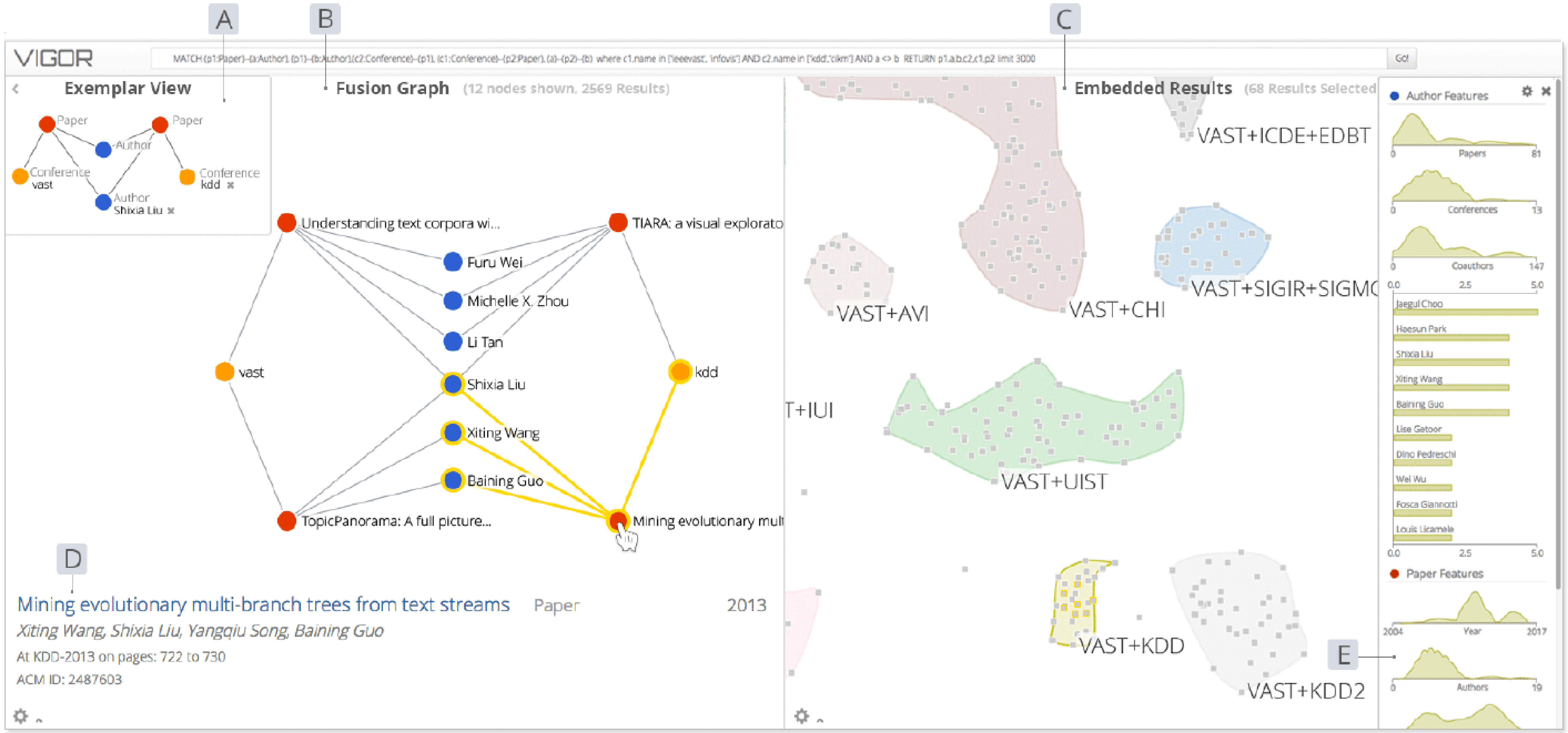




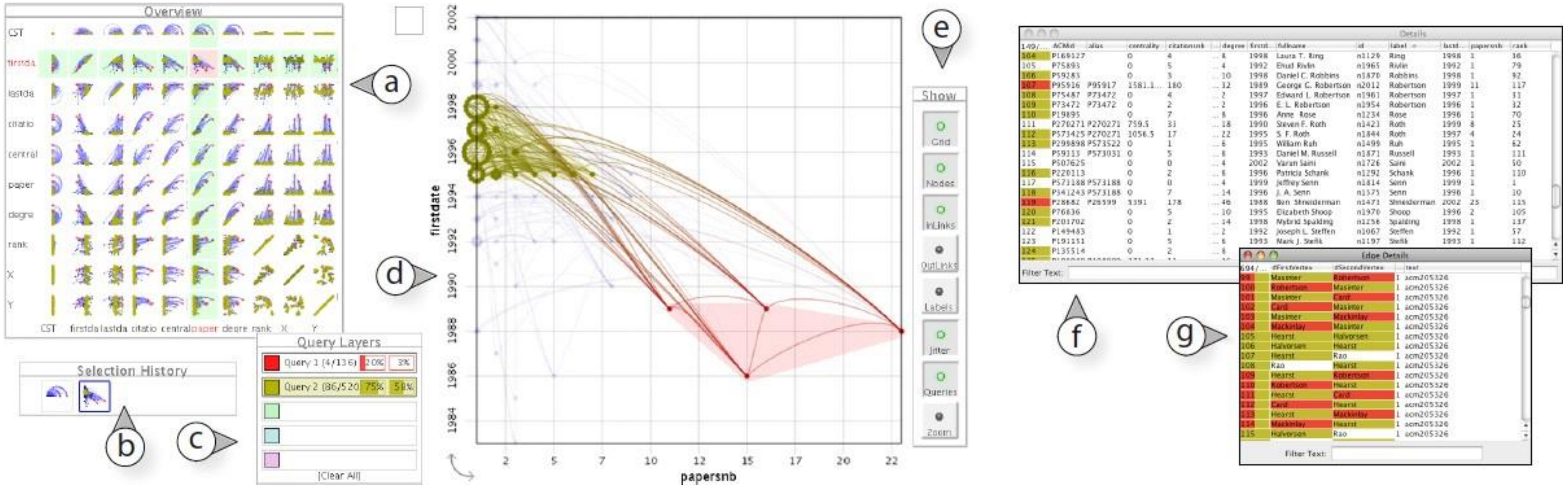
Name	Beverage	Day 1
Mark	Beer	1
Sue	Coke	0
Cole	Port	4
Jon	Coke	5
Tom	Beer	2
Abby	Port	3

VIGOR




A
B
C
D
E
VIGOR
Go!
Exemplar View
Paper
Paper
Author
Conference
vast
Conference
kdd
Author
Shixia Liu
Fusion Graph (12 nodes shown, 2569 Results)
Understanding text corpora wi...
TIARA: a visual explorato
Furu Wei
Michelle X. Zhou
Li Tan
vast
Shixia Liu
kdd
Xiting Wang
Baining Guo
TopicPanorama: A full picture...
Mining evolutionary mul
Mining evolutionary multi-branch trees from text streams
Paper
2013
Xiting Wang, Shixia Liu, Yangqiu Song, Baining Guo
At KDD-2013 on pages: 722 to 730
ACM ID: 2487603
Embedded Results (68 Results Selected
VAST+ICDE+EDBT
VAST+AVI
VAST+CHI
VAST+SIGIR+SIGMC
T+IUI
VAST+UIST
VAST+KDD
VAST+KDD2
Author Features
Papers
Conferences
Coauthors
Jaegul Choo
Haesun Park
Shixia Liu
Xiting Wang
Baining Guo
Lise Getoor
Dino Pedreschi
Wei Wu
Fosca Giannotti
Louis Licamele
Paper Features
Year
Authors

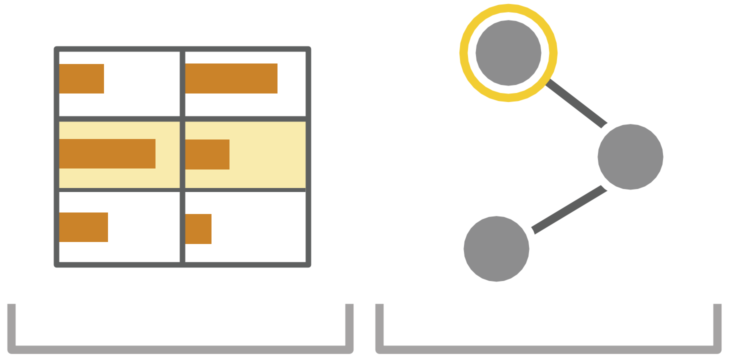
Graph Dice



Overview
Selection History
Query Layers
Query 1
Query 2
[Clear All]
firstdate
papersnb
Show
Grid
Nodes
InLinks
OutLinks
Labels
Jitter
Queries
Zoom
Details
Edge Details
Filter Text:
a
b
c
d
e
f
g

Independent views can optimize for topology and attribute independently.

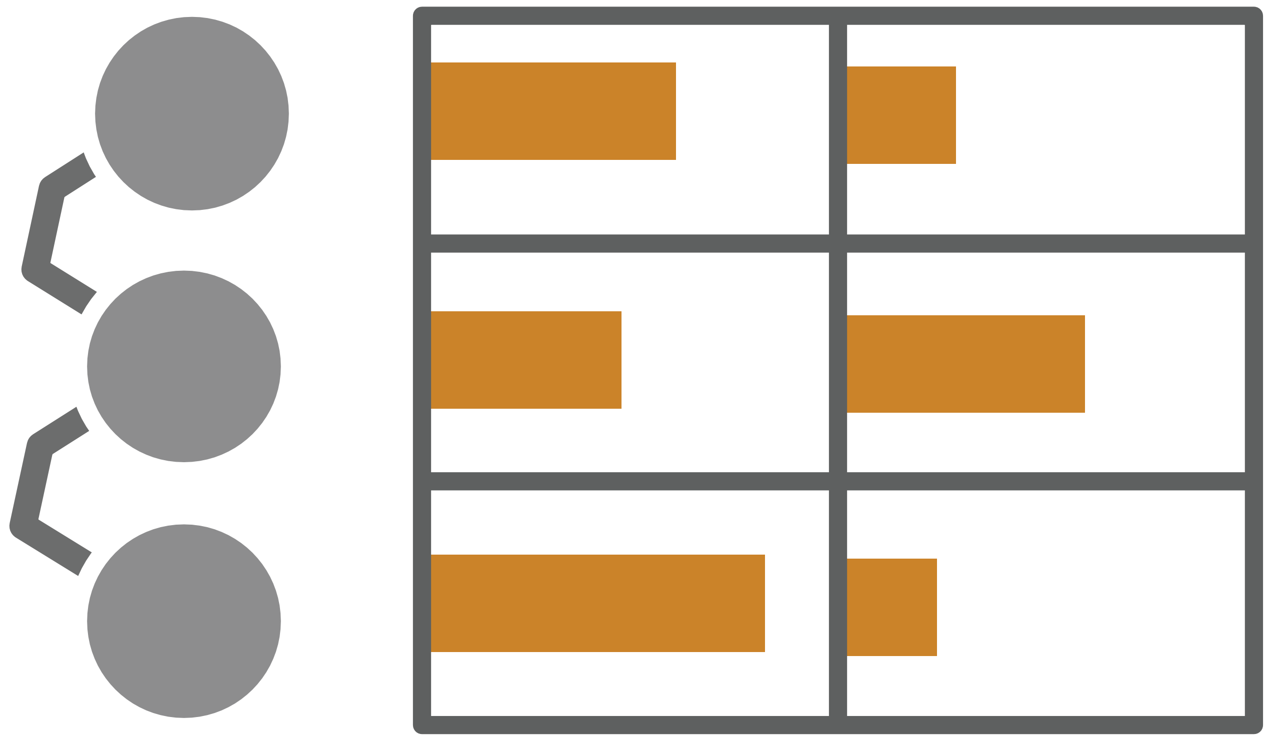
Not great for tasks on topological structures beyond a single node or edge.

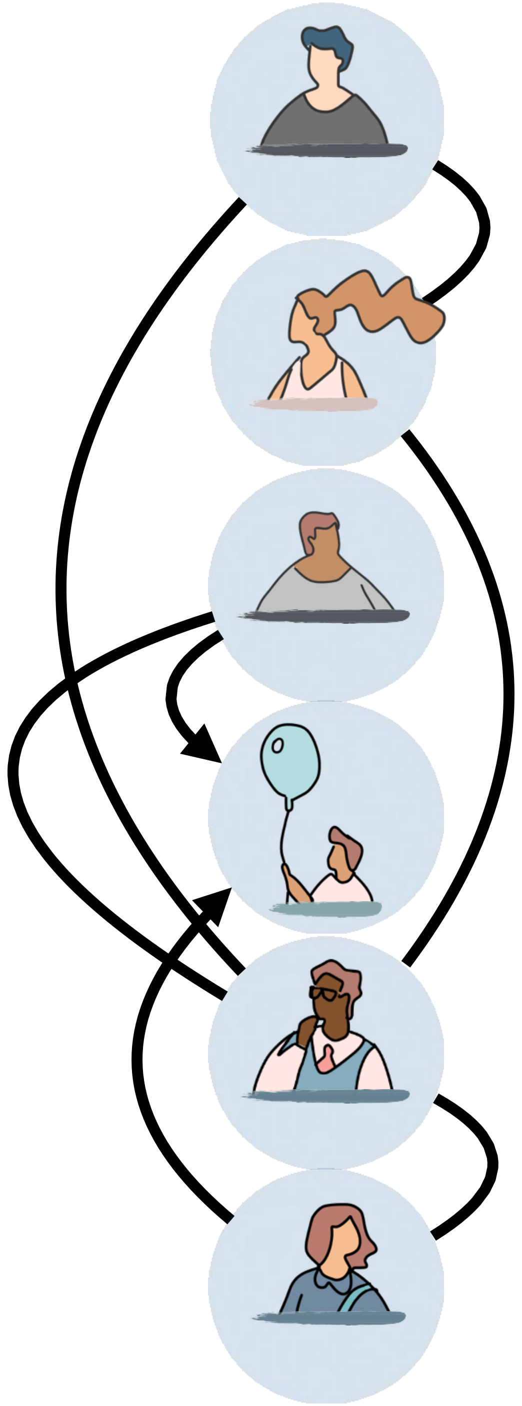


Juxtaposed

Recommended for large networks and/or very large numbers or heterogeneous types of node and link attributes

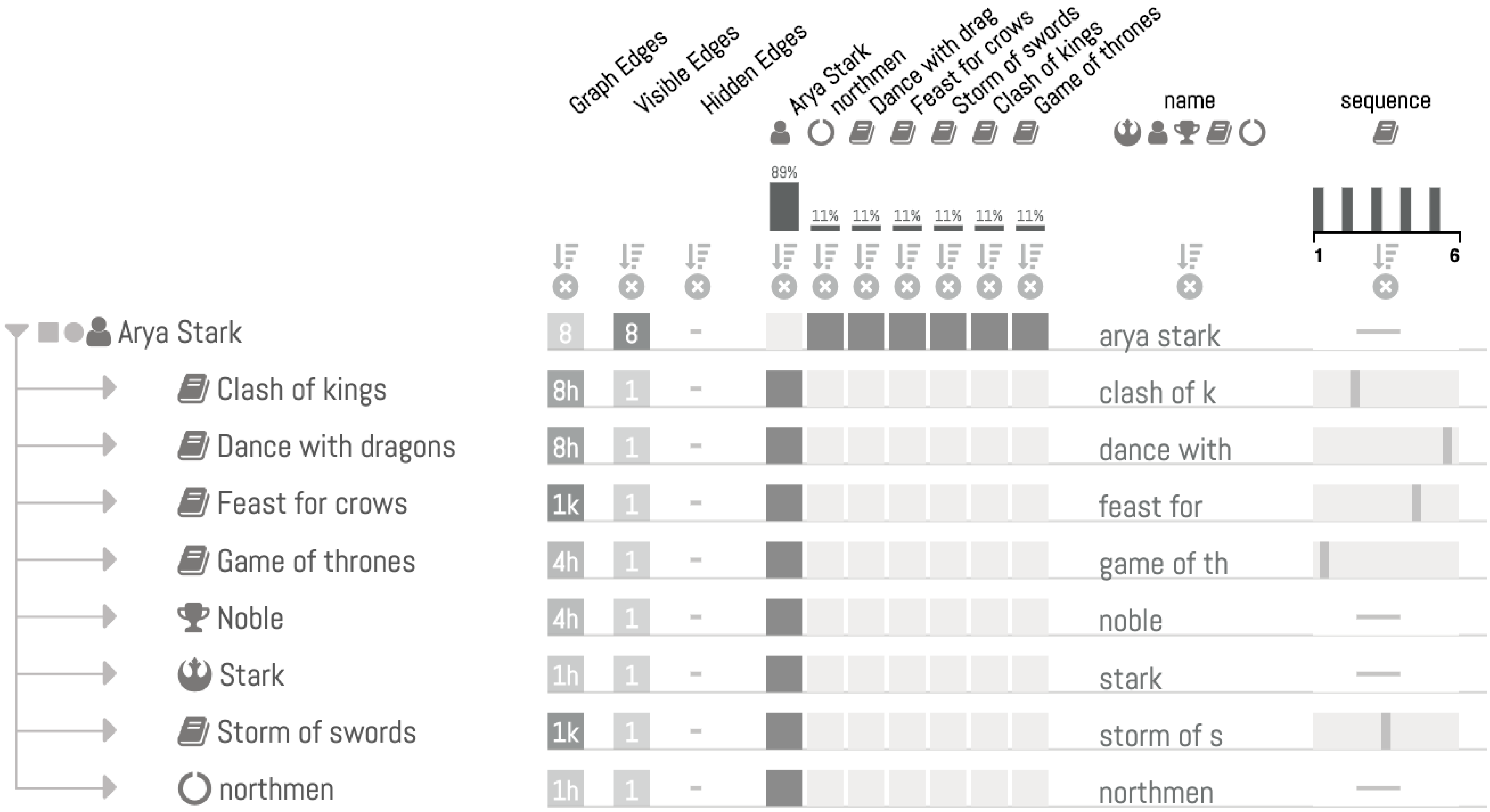
Integrated





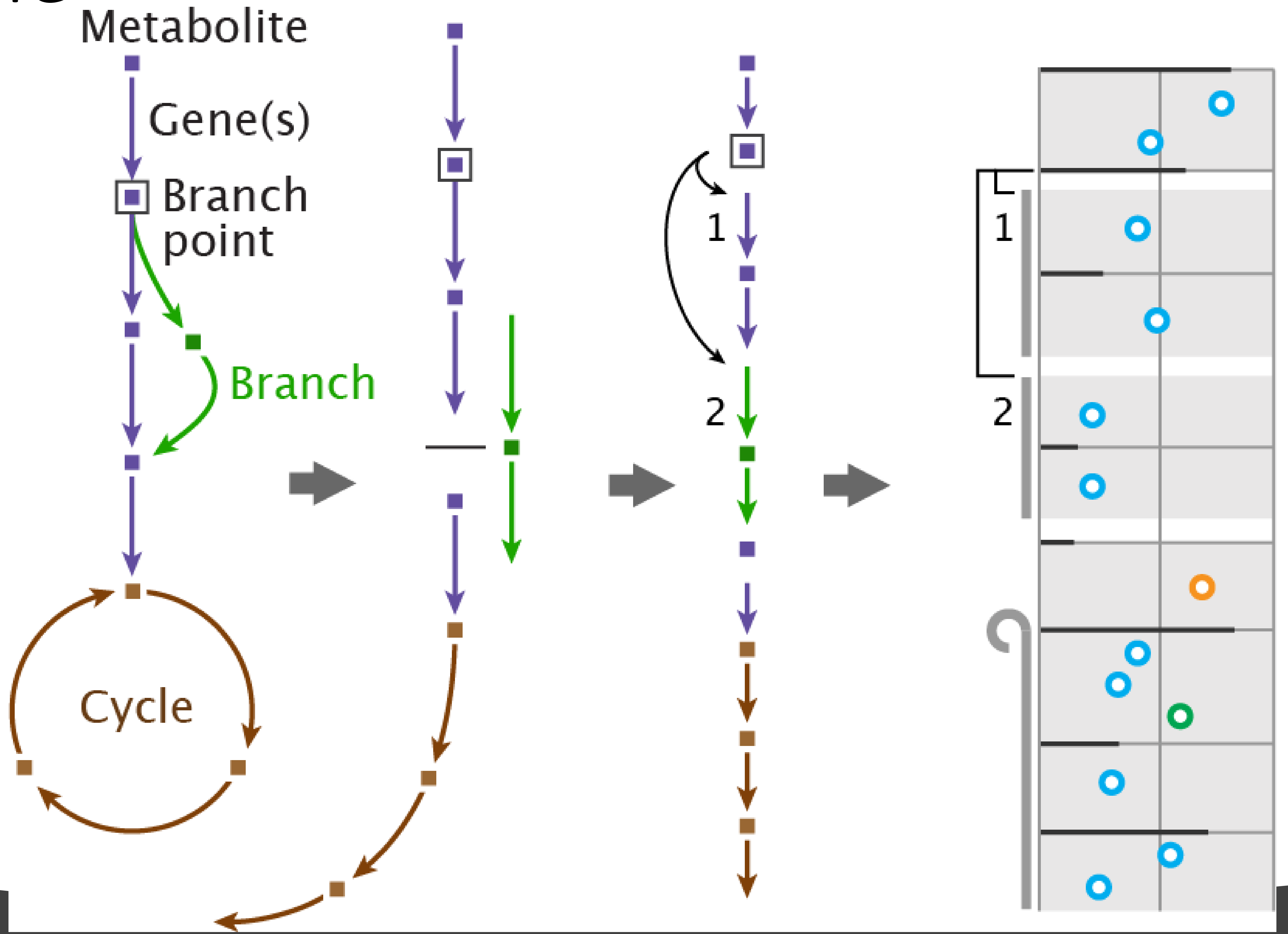
Name	Beverage	Day 1
Mark	Beer	1
Sue	Coke	0
Cole	Port	4
Jon	Coke	5
Tom	Beer	2
Abby	Port	3

Juniper



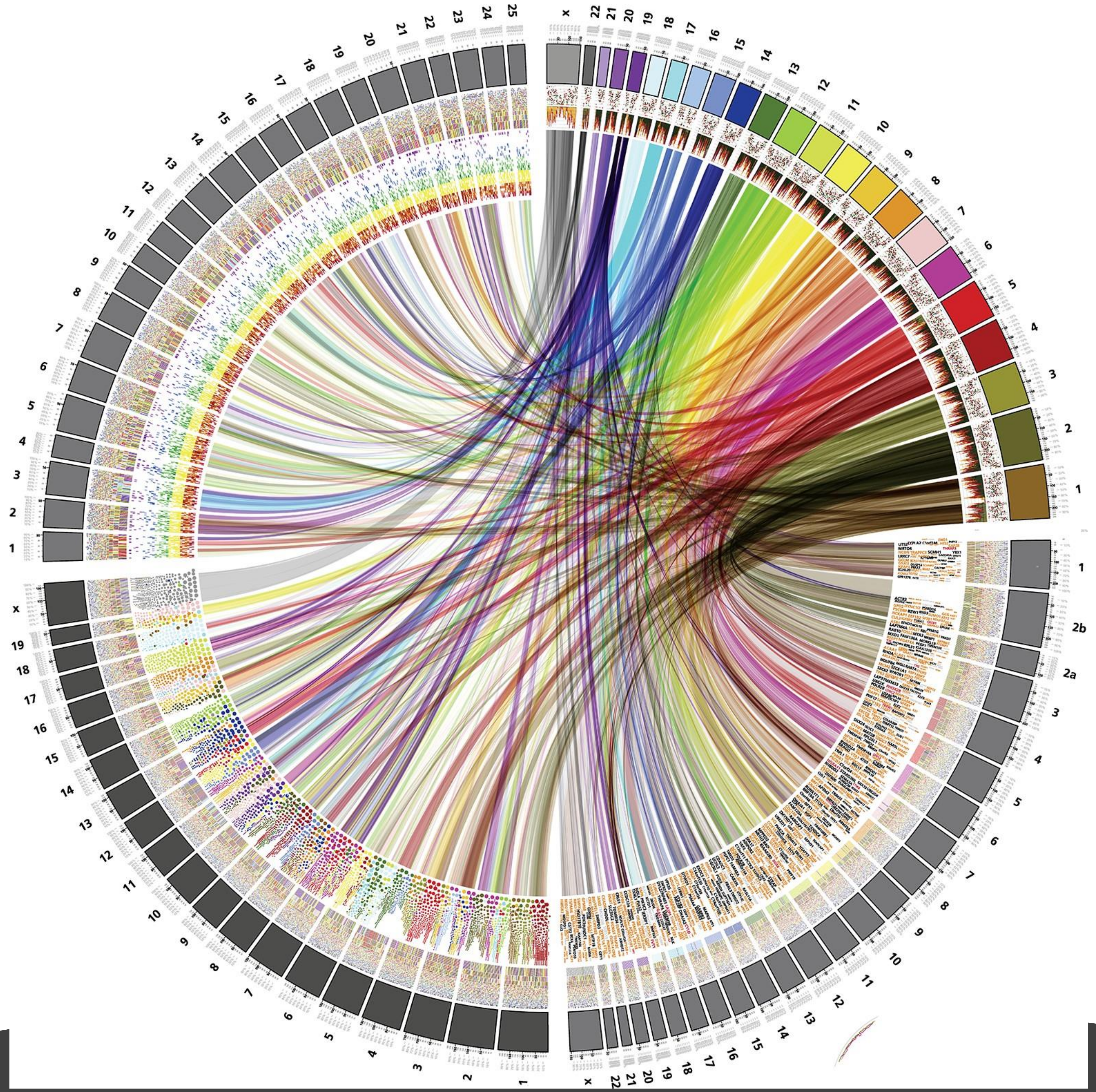
Graph Edges
Visible Edges
Hidden Edges
Arya Stark
northmen
Dance with drag
Feast for crows
Storm of swords
Clash of kings
Game of thrones
name
sequence
89%
11%
11%
11%
11%
11%
11%
1
6
Arya Stark
Clash of kings
Dance with dragons
Feast for crows
Game of thrones
Noble
Stark
Storm of swords
northmen
8
8h
8h
1k
4h
4h
1h
1k
1h
8
1
1
1
1
1
1
1
1
arya stark
clash of k
dance with
feast for
game of th
noble
stark
storm of s
northmen

Pathline

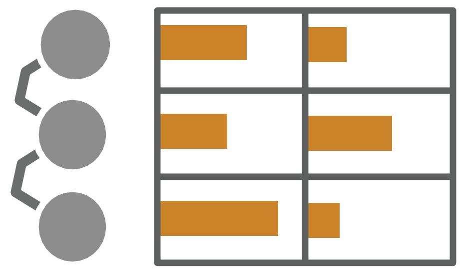


Meyer et al. 2010

Circos



Krzywinski et al. 2009



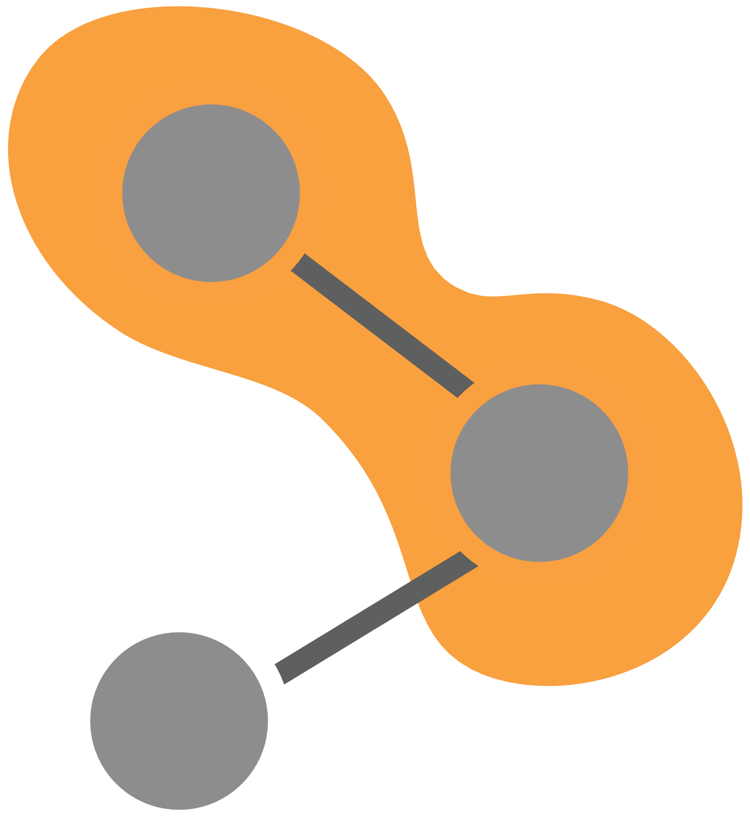
Integrated

good at integrating attributes with topology, if the topology can be represented in a linear layout.

Not suitable for networks that can not be sensibly linearized.

Recommended for networks with several, heterogenous, node attributes and well suited for tasks on single nodes, neighbors, and paths

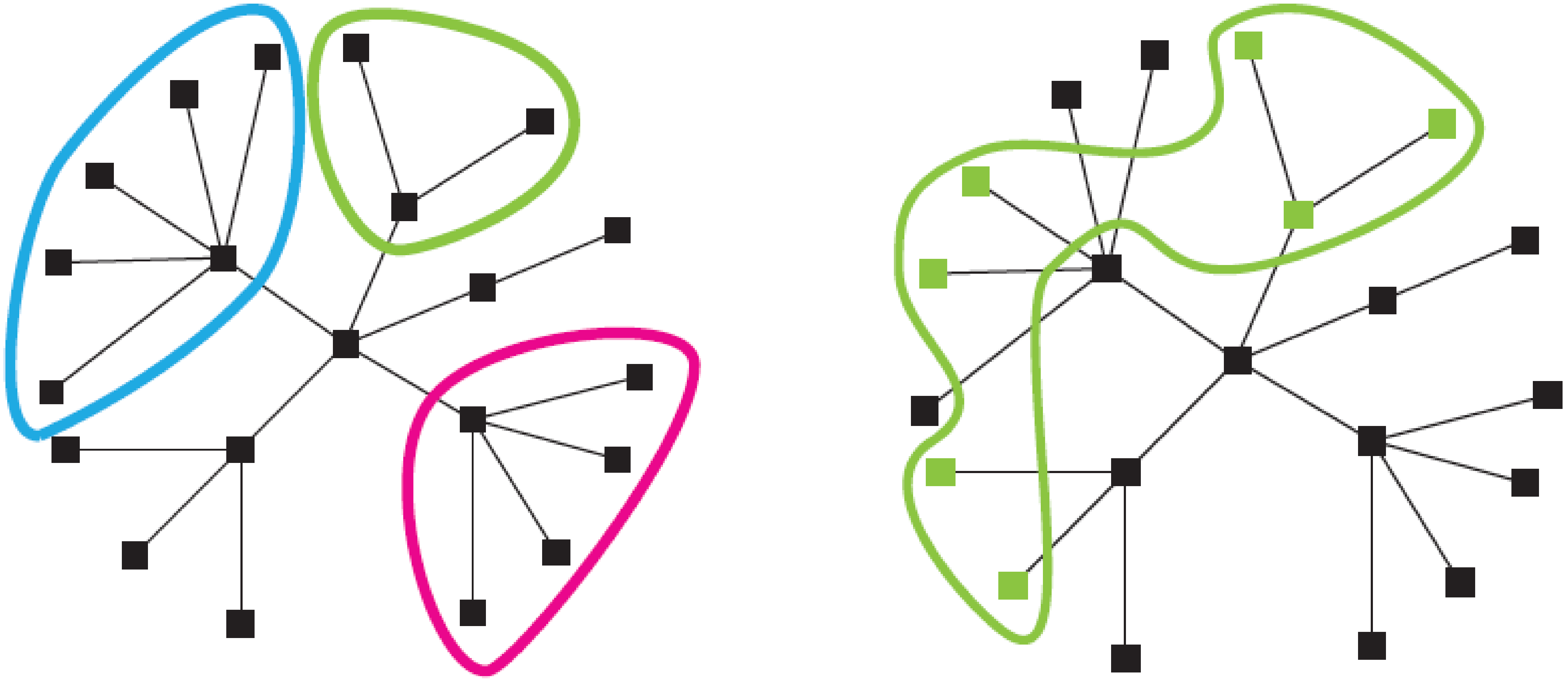
Overloaded



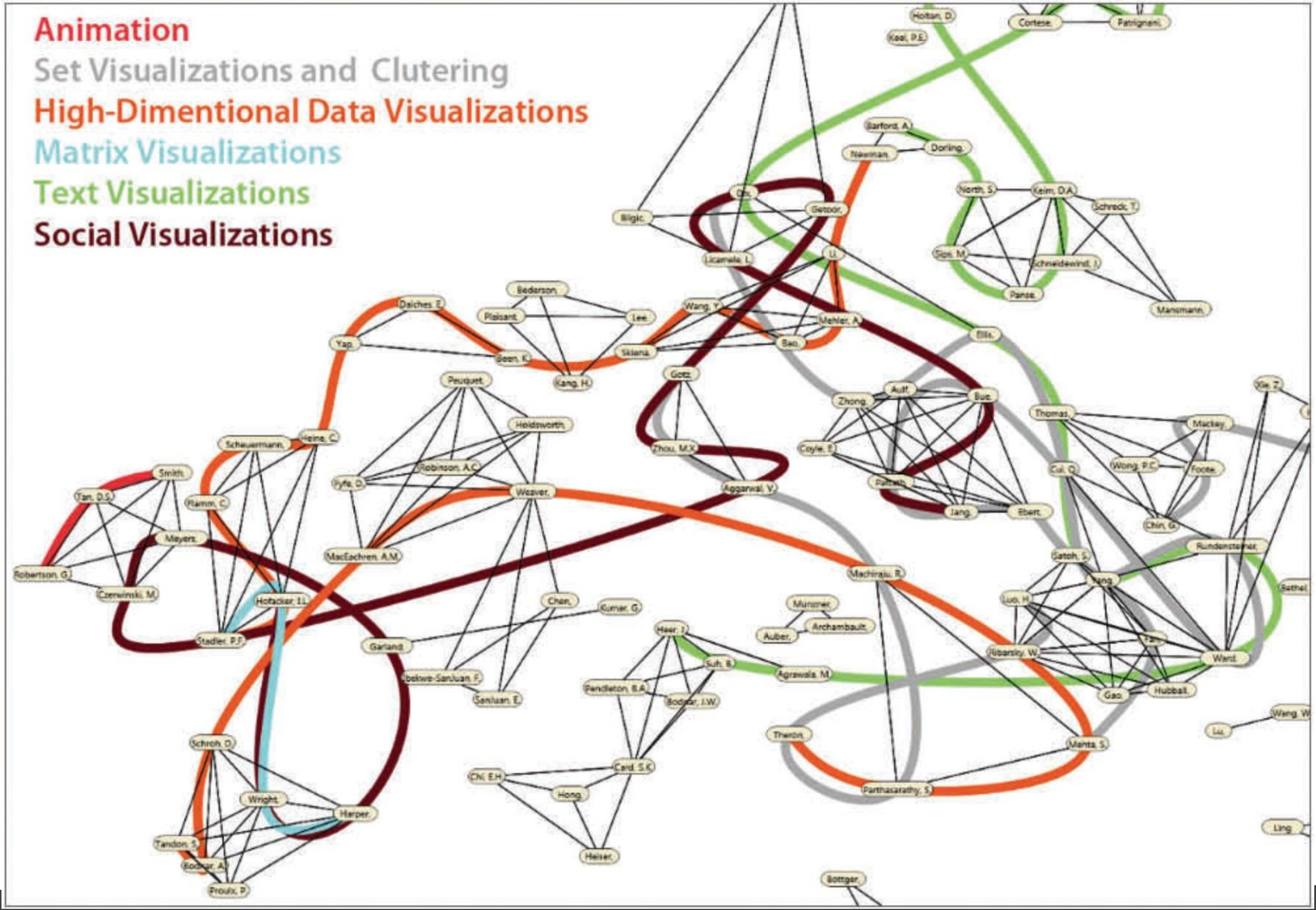
GMaps

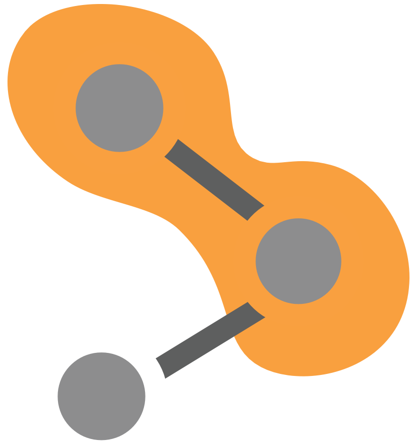
She &amp Him
Freelance Whales
Metric
Snow Patrol
The Dead Weather
Andrew W.K.
Electric Six
Eagles of Death Metal
Helmet
Queens of the Stone Age
The Shins
Empire of the Sun
Behemoth
Mudvayne
Mastodon
Radiohead
Beach House
Tool
Pink Floyd
Kalmah
Death
Gojira
Animal Collective
Isis
Dark Tranquillity
Moonspell
Mono
The Album Leaf
Efterklang
Porcupine Tree
Battles
Akron/Family
Pain of Salvation
Les Savy Fav
Tim Hecker
Tortoise
Jaga Jazzist

Bubble Sets



LineSets


Animation
Set Visualizations and Clutering
High-Dimentional Data Visualizations
Matrix Visualizations
Text Visualizations
Social Visualizations



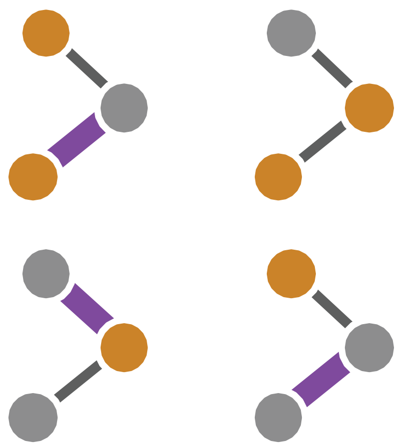
Overloaded

good at displaying sets and clusters

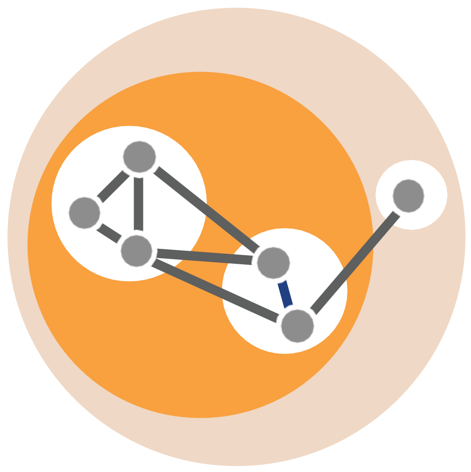
Not suitable for displaying more than one or two attributes at a time.

Recommended for recommend overloading for the particular use case of visualizing set-memberships or clusters on top of node-link diagrams

Layout Operations

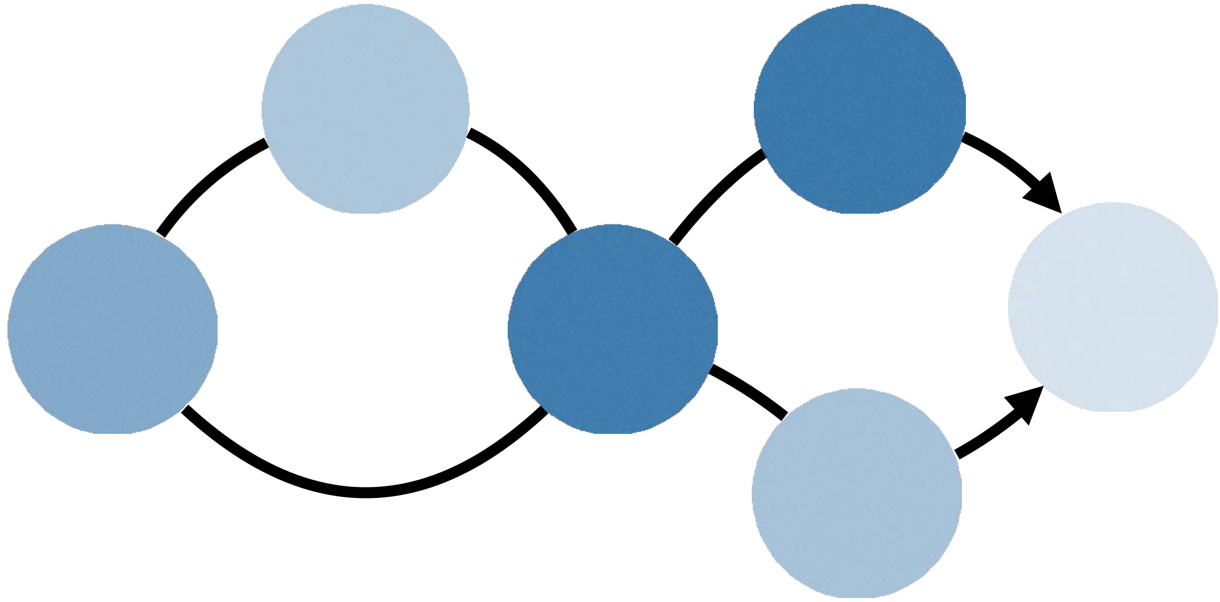


Small Multiples

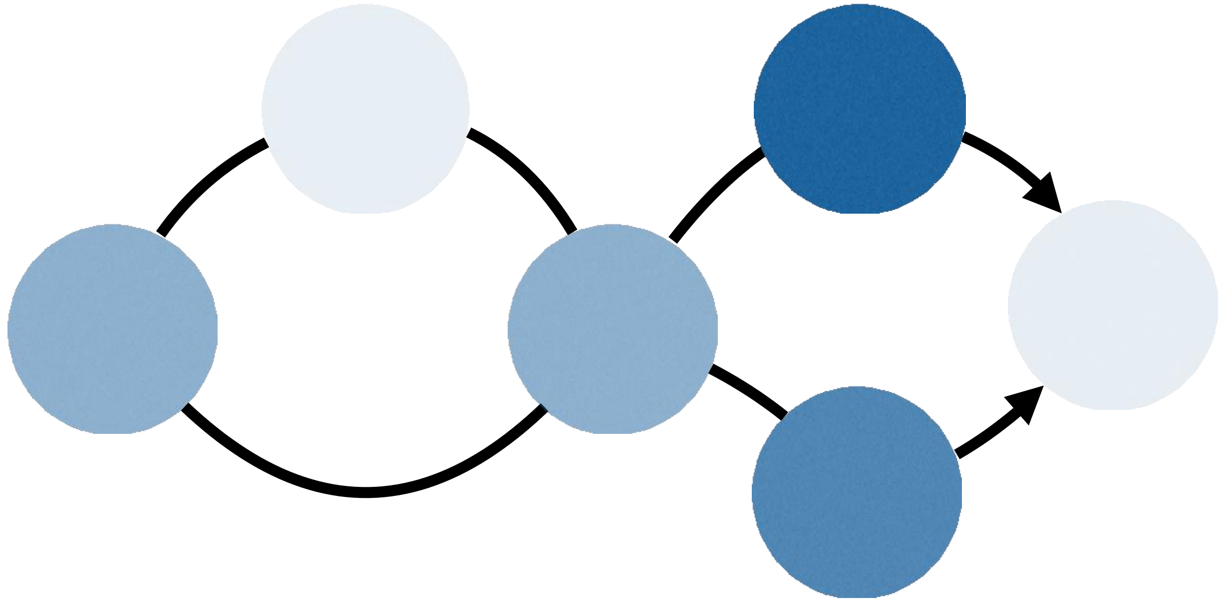


Hybrids

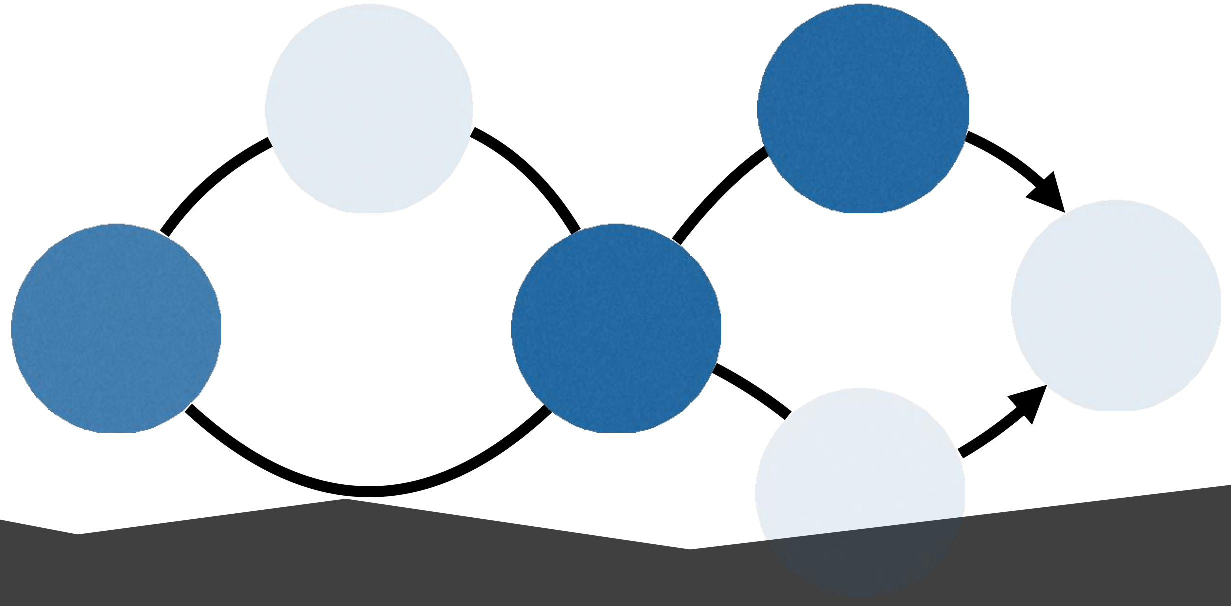
Day 1



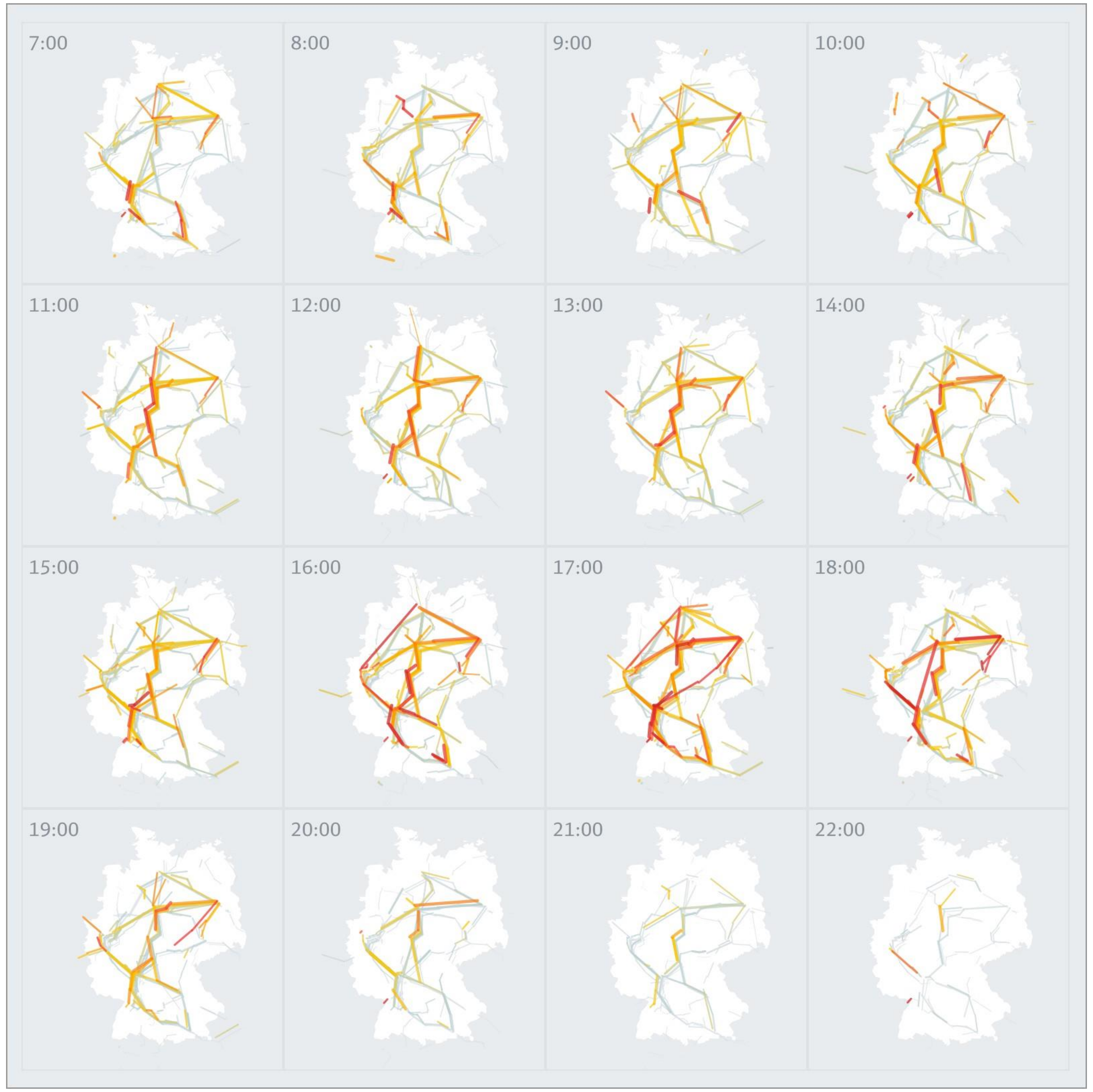
Day 2



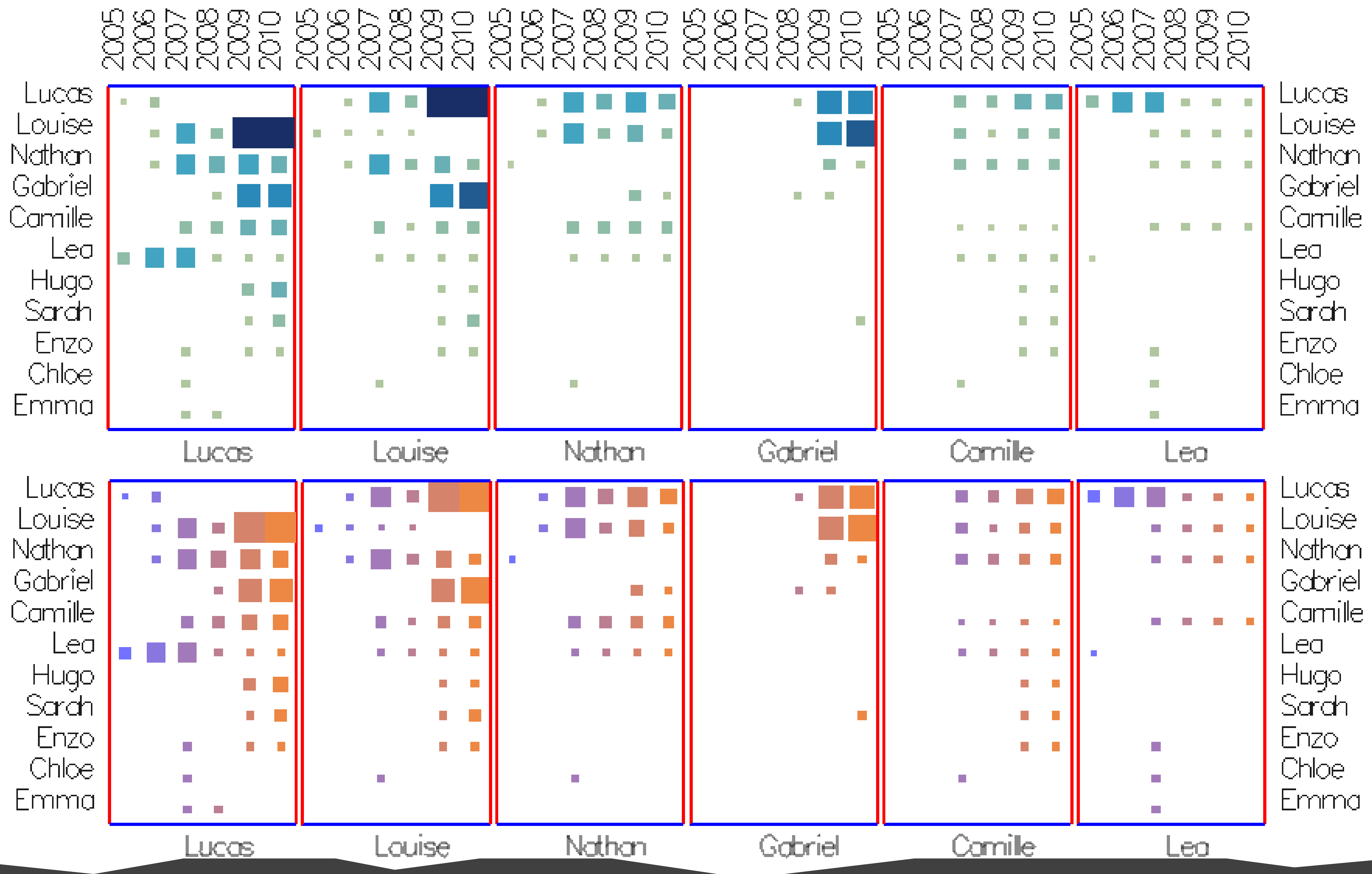
Day 3

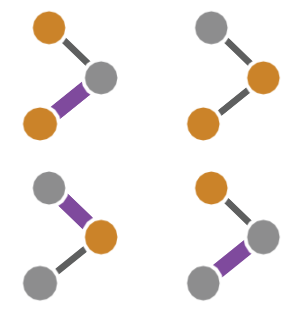


Peakspotting



Small Multiples





Small Multiples

Common layout facilitates attribute comparisons in specific topological features

Not ideal for large networks, or tasks on clusters

Recommended for small networks where the tasks are focused on attribute comparison

More at http://vdl.sci.utah.edu/mvnv/