The Sprawlter Graph Readability Metric: Combining Sprawl and Area-aware Clutter

Zipeng Liu
University of British Columbia

Takayuki Itoh
Ochanomizu University

Jessica Q. Dawson
University of British Columbia

Tamara Munzner
University of British Columbia

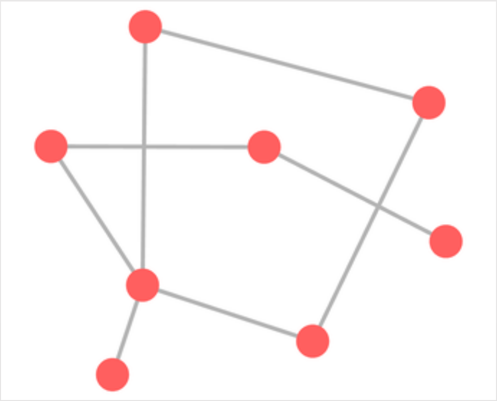
DESIGNING for PEOPLE

Project link: http://www.cs.ubc.ca/labs/imager/tr/2020/sprawlter/

Presented at PacificVis 2020, on June 3 2020

Node-link graph readability metrics

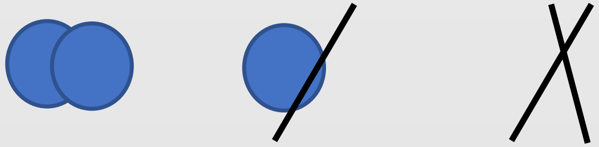
- Measure ability of a human to read a graph
  - Synonym: aesthetic criteria
  - Example: # edge-edge crossing

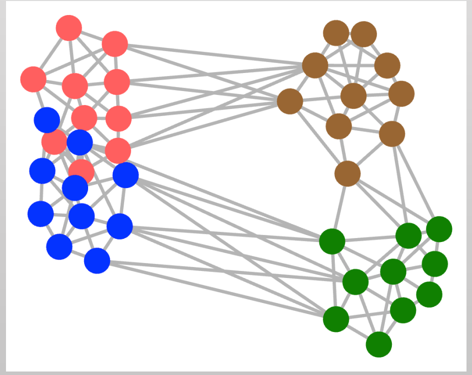


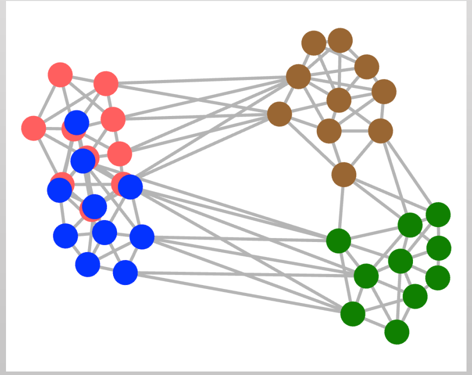
- Why care about metric?
  - Evaluate layout
  - Make good layout

Readability problems

- Integer crossing counts
  - Lack precision

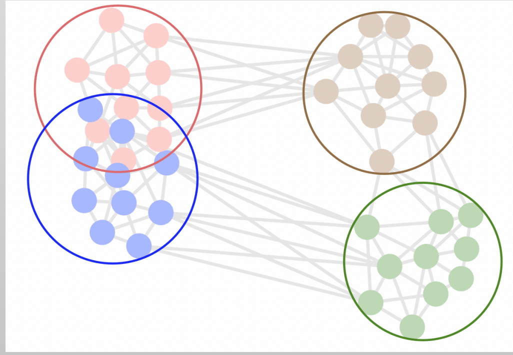


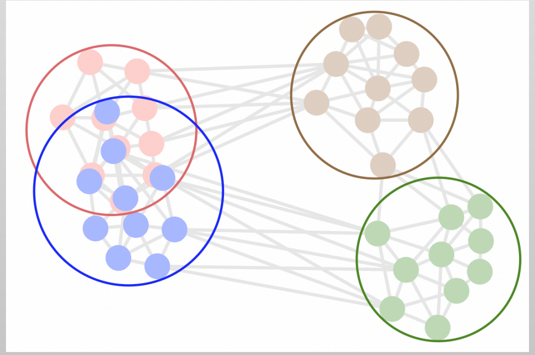




Readability problems

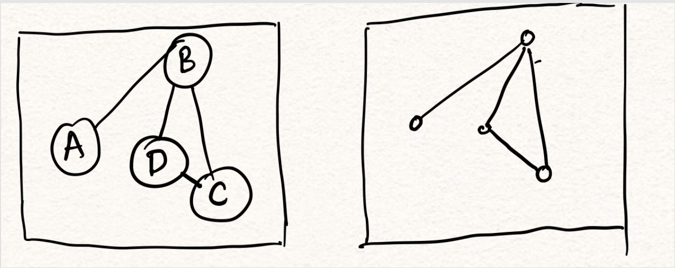
- Integer crossing counts
  - Lack precision
- Single-level structure
  - Metanodes ignored but more salient

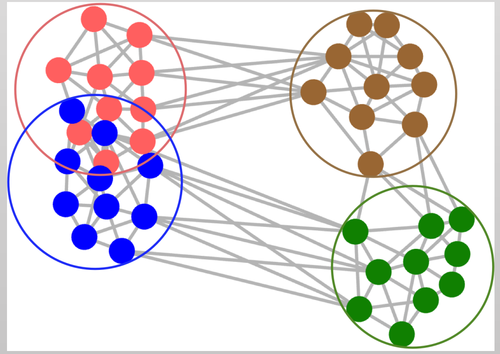


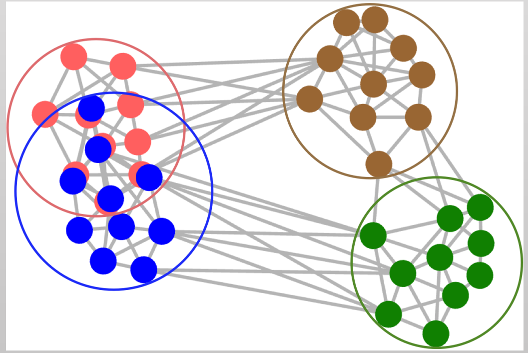


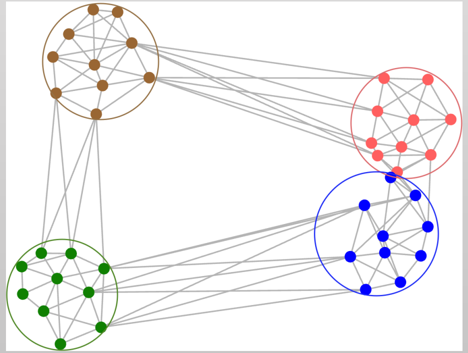
Readability problems

- Integer crossing counts
  - Lack precision
- Single-level structure
  - Metanodes ignored but more salient
- Clutter-only metrics
  - *Sprawl* (geometric sparseness) ignored









Solutions for

- More precise clutter
- Account for metanodes
- Combine clutter and sprawl

Readability problems

- Integer crossing counts
- Single-level structure
- Clutter-only metrics

Contributions

- Propose new **area-aware clutter** metric
  - Area / angles of overlap → penalty
    - Penalty mapping function
  - Implemented for node-node, node-edge, edge-edge
- Propose new **sprawlter** (sprawl - clutter) metric
  - Trade off sprawl & area-aware clutter with geometric mean
- Evaluate both metrics against baselines
  - 56 graph layouts
    - 13 datasets
    - 7 layout approaches: 6 algorithms + manual

Readability problems

- Integer crossing counts
- Single-level structure
- Clutter-only metrics

Related work: readability metrics

- Single-purpose metrics
  - Single-level clutter:
    - Edge-edge / node-edge crossings [Purchase 03]

      Does not account for geometric overlap

    - Edge crossing angles [Dunne 15]

      Does not account for just-touching cases

  - Multi-level clutter: much less work
    - Group overlap [Dunne 15]

      Does not account for clutter between metanodes

  - Sprawl
    - Compactness [Kieffer 15], visualization coverage [Dunne 15]

      Does not integrate sprawl and clutter

- Compound metrics
  - Implicitly used by layout algorithms (e.g. GEM [Frick 94], LinLog [Noack 27], NEATO [Gansner 04])
    - Physics quantity (e.g. force, energy, stress)

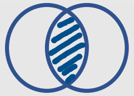
      Does not provide explicit quantity

  - Explicitly combined
    - Weighted sum of single-purpose metrics [Huang 16]

      Does not say what weights to use

      Our **sprawlter** metric uses geometric mean

Sprawlter metric computational pipeline

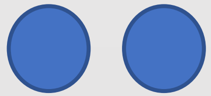
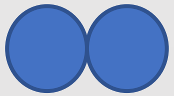
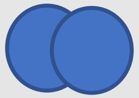
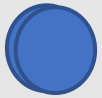
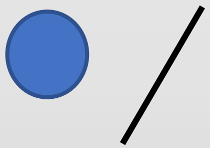
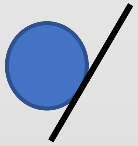
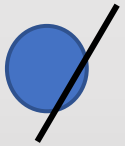
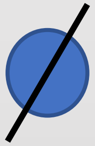
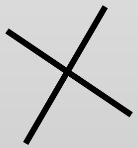
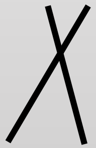
1. Measure geometric overlaps  x

2. Map measurements to penalties $f(x)$

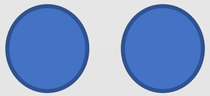
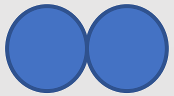
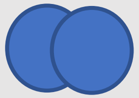
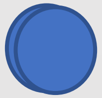
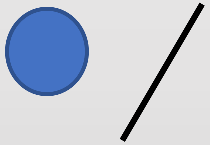
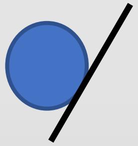
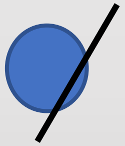
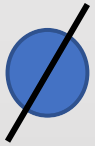
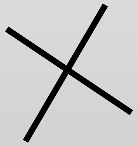
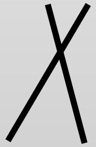
3. Sum up penalties $AreaMetric = \sum f(x)$

4. Combine clutter with sprawl $Sprawlter = \sqrt{Sprawl * AreaMetric}$

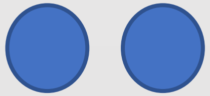
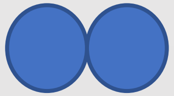
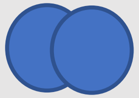
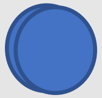
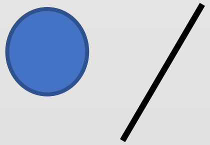
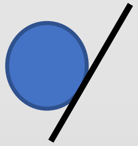
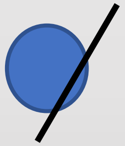
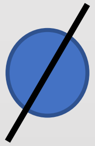
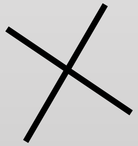
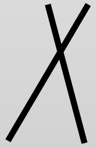
1. Measure geometric overlaps

Amount of geometric overlap x		None	Near-min	Some	Near-max
Node-node	Area				
Node-edge	Length				
Edge-edge	Angle				

2. Map measurements to penalties

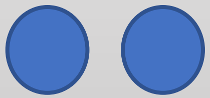
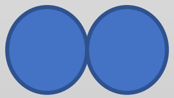
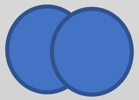
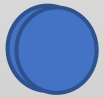
Amount of geometric overlap x		None	Near-min	Some	Near-max
Node-node	Area				
Node-edge	Length				
Edge-edge	Angle				
Count (baseline)		0	1	1	1

2. Map measurements to penalties

Amount of geometric overlap $\boldsymbol{x}$		None	Near-min	Some	Near-max
Node-node	Area				
Node-edge	Length				
Edge-edge	Angle				
Count (baseline)		0	1	1	1
Penalty $f(x)$		0	?	?	?

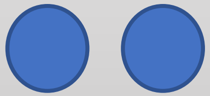
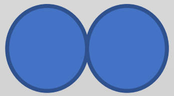
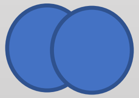
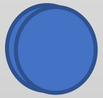
Penalty mapping function: general requirements

- Substantial touching penalty

Amount of geometric overlap $\boldsymbol{x}$		None	Near-min	Some	Near-max
Node-node	Area				
Count (baseline)		0	1	1	1
Penalty $\boldsymbol{f}(\boldsymbol{x})$		0	0 < min		

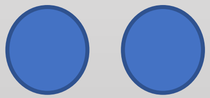
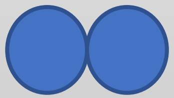
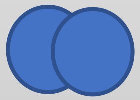
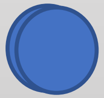
Penalty mapping function: general requirements

- Substantial touching penalty
- Increasing penalty

Amount of geometric overlap $\boldsymbol{x}$		None	Near-min	Some	Near-max
Node-node	Area				
Count (baseline)		0	1	1	1
Penalty $\boldsymbol{f(x)}$		0	0 < min		min < max

Penalty mapping function: general requirements

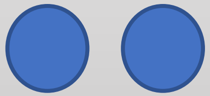
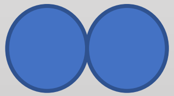
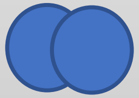
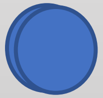
- Substantial touching penalty
- Increasing penalty
- Count calibration: min < 1 < max

Amount of geometric overlap $\boldsymbol{x}$		None	Near-min	Some	Near-max
Node-node	Area				
Count (baseline)		0	1	1	1
Penalty $\boldsymbol{f(x)}$		0	0 < min < 1	1	1 < max

Penalty mapping function: general requirements

- Substantial touching penalty
- Increasing penalty
- Count calibration: min < 1 < max
- Local function:
  - Pairs of nodes with different sizes should have different $f(x)$

f  $(x) \neq f$  (x)

Amount of geometric overlap $\boldsymbol{x}$		None	Near-min	Some	Near-max
Node-node	Area				
Count (baseline)		0	1	1	1
Penalty $\boldsymbol{f}(\boldsymbol{x})$		0	0 < min < 1	1	1 < max

Penalty mapping function: specific choices

$$f(x) = (\beta - \alpha)x^{\gamma} + \alpha M^{\gamma} \ (0 \leq x \leq M)$$

Penalty mapping function: specific choices

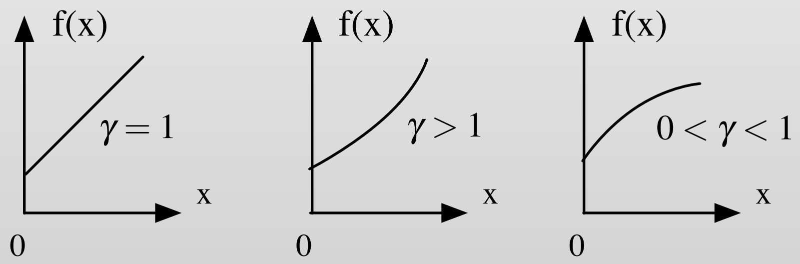
$$f(x) = (\beta - \alpha)x^{\gamma} + \alpha M^{\gamma} \ (0 \leq x \leq M)$$

- Constant M: max possible overlap

Penalty mapping function: specific choices

$$f(x) = (\beta - \alpha)x^{\gamma} + \alpha M^{\gamma} \quad (0 \le x \le M)$$

- Constant M: max possible overlap
- Parameters:
  - γ: curve shape of the power function, $f(x) \sim x^{\gamma}$
    - Inspired by Stevens' psychophysical power law



Penalty mapping function: specific choices

$$f(x) = (\beta - \alpha)x^{\gamma} + \alpha M^{\gamma} \ (0 \le x \le M)$$

- Constant M: max possible overlap
- Parameters:
  - γ: curve shape of the power function, $f(x) \sim x^{\gamma}$
    - Inspired by Stevens' psychophysical power law
  - α and β: control factor of min and max penalty

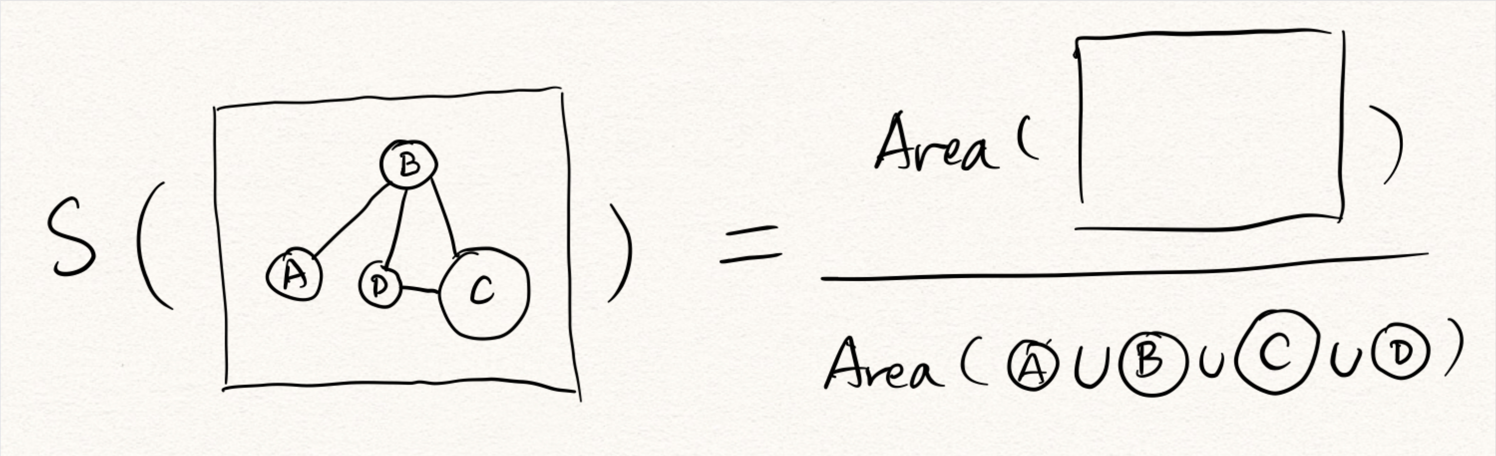
- More in paper and supplemental

3. Sum up penalties

$$AreaMetric = \sum f(x)$$

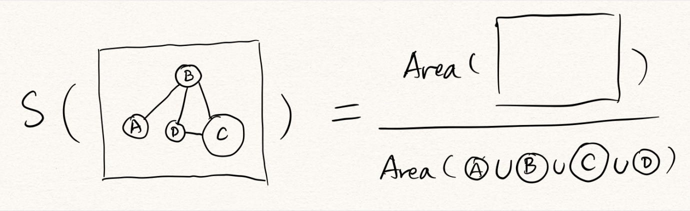
4. Combine clutter with sprawl

Sprawl $S = \frac{Total\ drawing\ area}{Area\ of\ all\ nodes}$



4. Combine clutter with sprawl

Sprawl $S = \frac{Total\ drawing\ area}{Area\ of\ all\ nodes}$



Sprawlter metric properties

- High penalty for bad cases:
  - Increase sprawl to reduce clutter
  - Compact space usage but high clutter
- Low penalty for good case:
  - low sprawl and low clutter

Combine sprawl and clutter with geometric mean

$$Sprawlter = \sqrt{Sprawl \cdot max\{\ AreaMetric, 1\ \}}$$

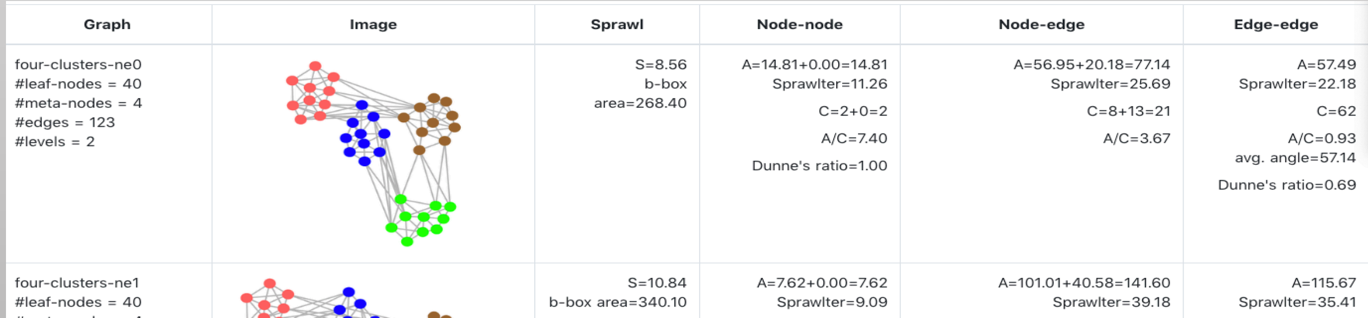
- Normalize different ranges
- Clutter with floor of 1 to retain sprawl in no-clutter case

Evaluation

- Quantitative
  - Compare computed values: our metrics vs. previous
- Qualitative
  - Discuss matches and mismatches: layout pictures vs. metric values
    - Subjective judgement of authors

Evaluation procedure

- Generate 56 graph layouts
    - different datasets and layout algorithms
- Compute area, sprawl, sprawlter and baseline previous metrics
- Supplemental table
    - layout pictures and metrics for all
- Paper
    - detailed assessment of metrics w.r.t. layout characteristics
- Talk
    - selected examples

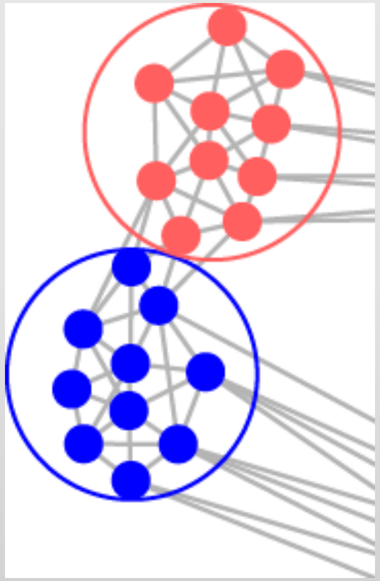
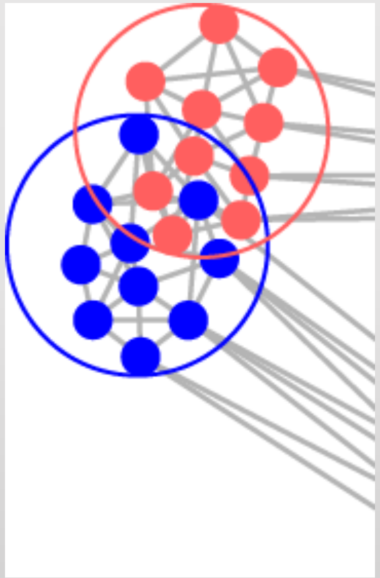
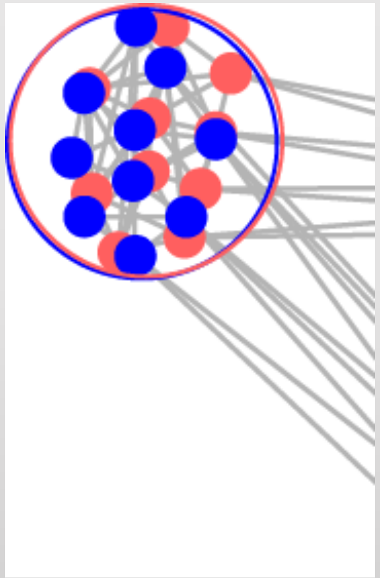


Graph	Image	Sprawl	Node-node	Node-edge	Edge-edge
four-clusters-ne0 #leaf-nodes = 40 #meta-nodes = 4 #edges = 123 #levels = 2		S=8.56 b-box area=268.40	A=14.81+0.00=14.81 Sprawlter=11.26 C=2+0=2 A/C=7.40 Dunne's ratio=1.00	A=56.95+20.18=77.14 Sprawlter=25.69 C=8+13=21 A/C=3.67	A=57.49 Sprawlter=22.18 C=62 A/C=0.93 avg. angle=57.14 Dunne's ratio=0.69
four-clusters-ne1 #leaf-nodes = 40		S=10.84 b-box area=340.10	A=7.62+0.00=7.62 Sprawlter=9.09	A=101.01+40.58=141.60 Sprawlter=39.18	A=115.67 Sprawlter=35.41

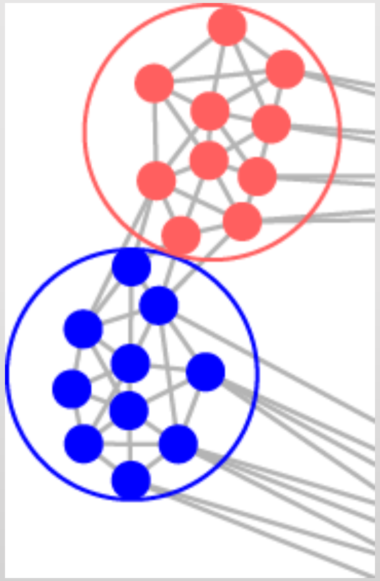
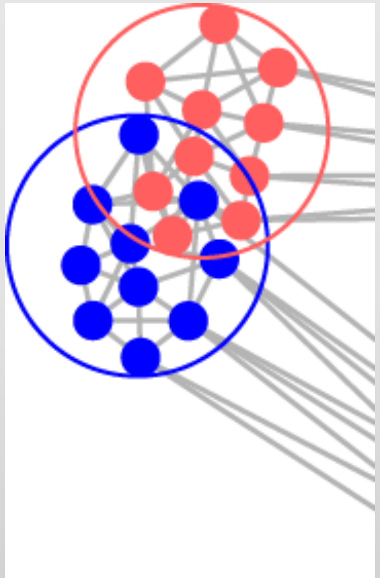
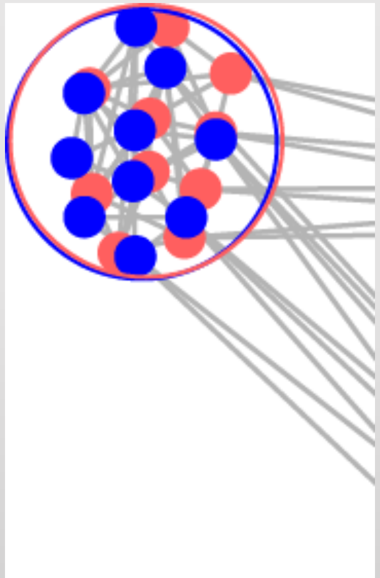
Graph layouts

- 56 graph layouts = 38 synthetic + 18 real-world
  - #nodes: 50 - 5K
  - #edges: 150 - 10K
  - Layout algorithm:
    - Manual position
    - Single-level: GEM [Frick 94], FME [Gronemann 09], Davidson-Harel [Davidson 96], Stress Majorization [Gansner 04]
    - Multi-level: Koala [Itoh 15], GrouseFlocks [Archambault 08]

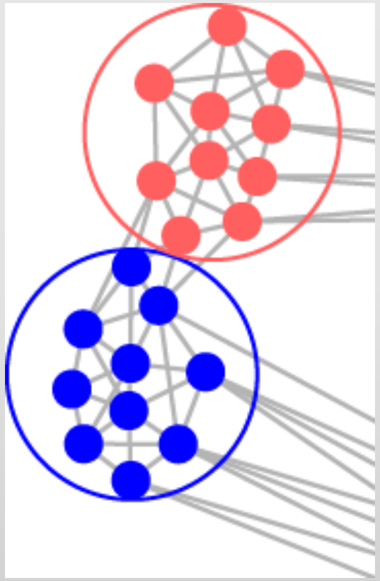
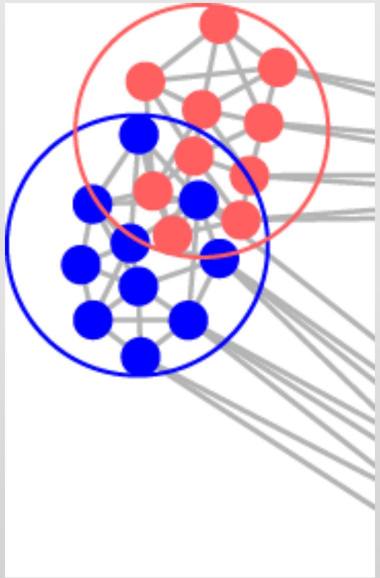
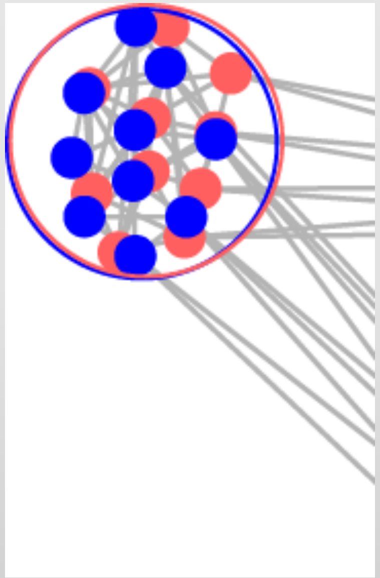
Node-node: increasing overlap

	Near-min	Some	Near-max
Layout			

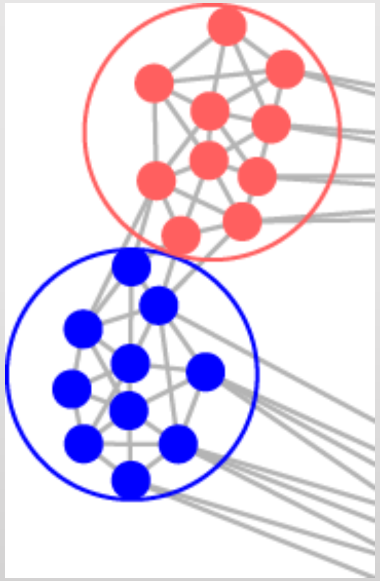
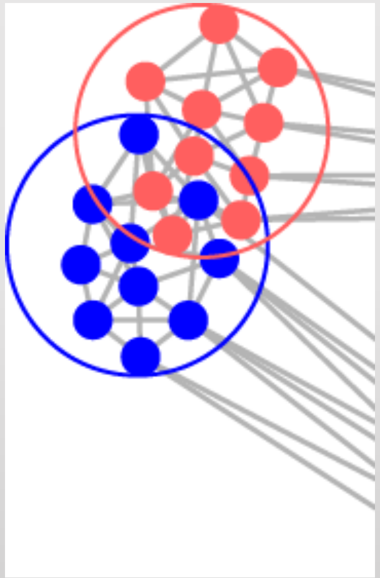
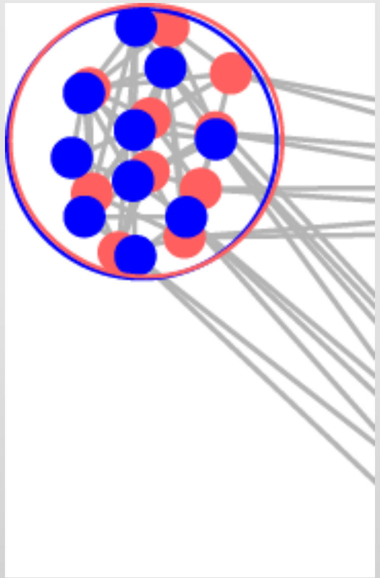
Node-node: increasing overlap

	Near-min	Some	Near-max
Layout			
Area-aware clutter	4.3	14.4	30.6

Node-node: increasing overlap

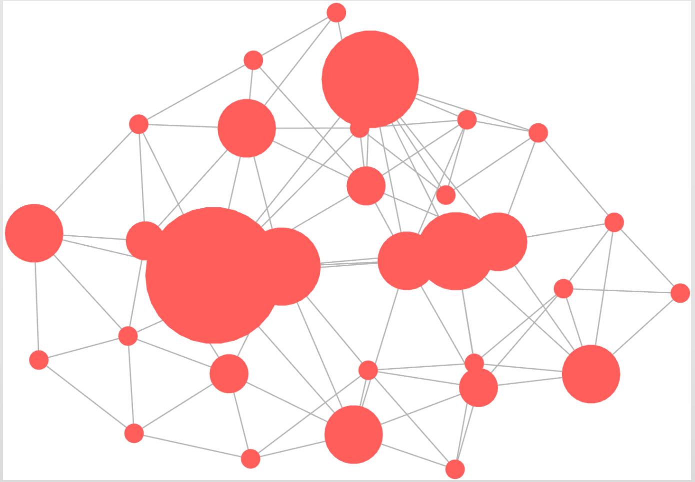
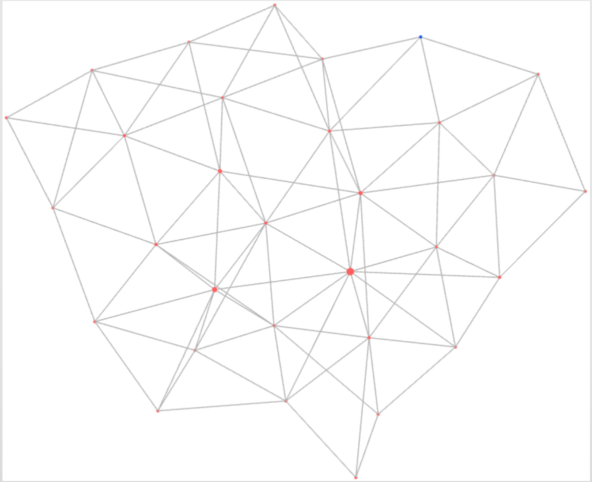
	Near-min	Some	Near-max
Layout			
Area-aware clutter	4.3	14.4	30.6
Count (baseline)	1	1	12

Node-node: increasing overlap

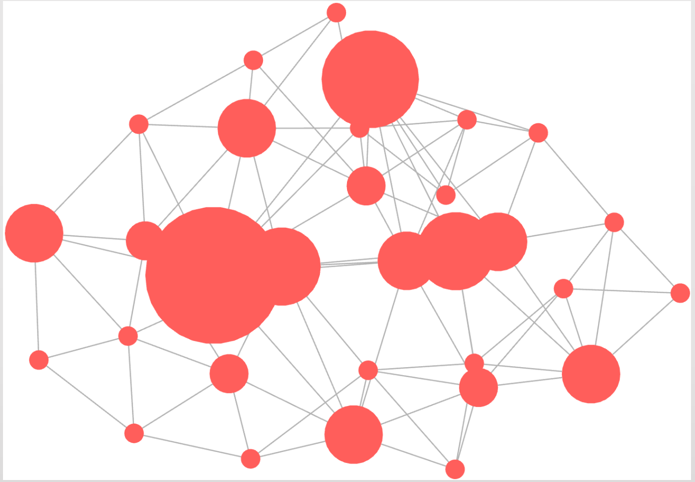
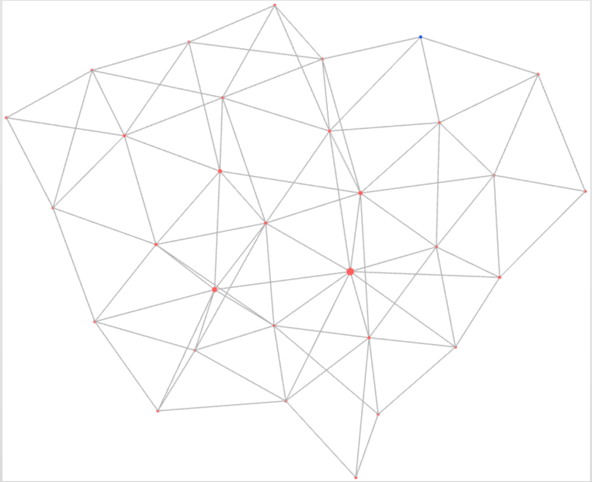
	Near-min	Some	Near-max
Layout			
Area-aware clutter	**4.3** + 0 = 4.3	**14.4** + 0 = 14.4	**22.1** + 8.6 = 30.6
Count (baseline)	**1**+ 0 = 1	**1** + 0 = 1	**1** + 11 = 12

Penalty for metanodes

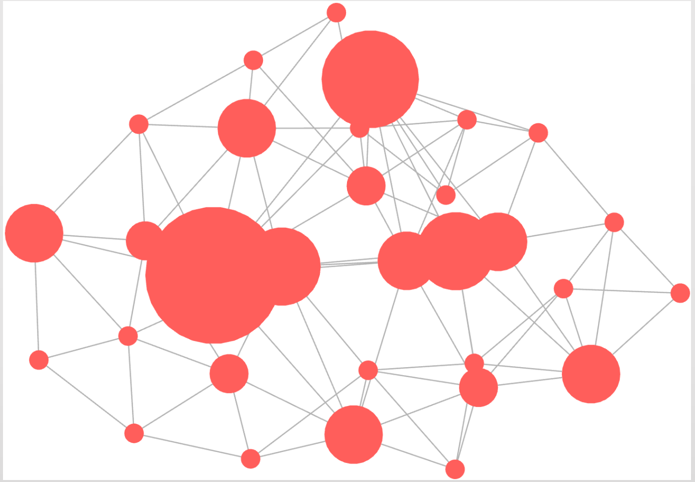
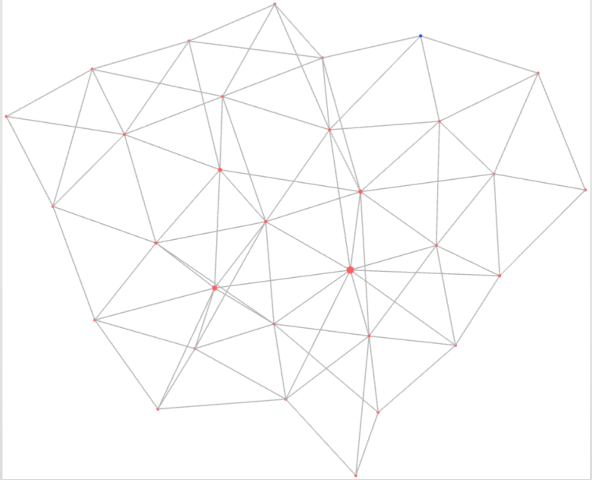
Node-node: variable node size layout

	Some clutter; less sprawl	No clutter; high sprawl
Layout		

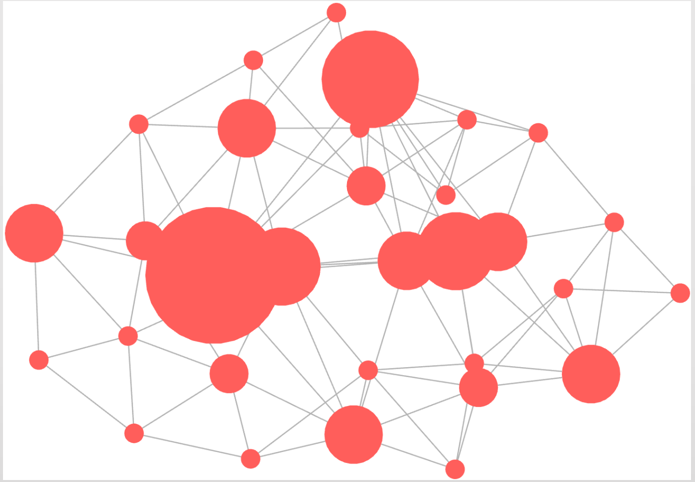
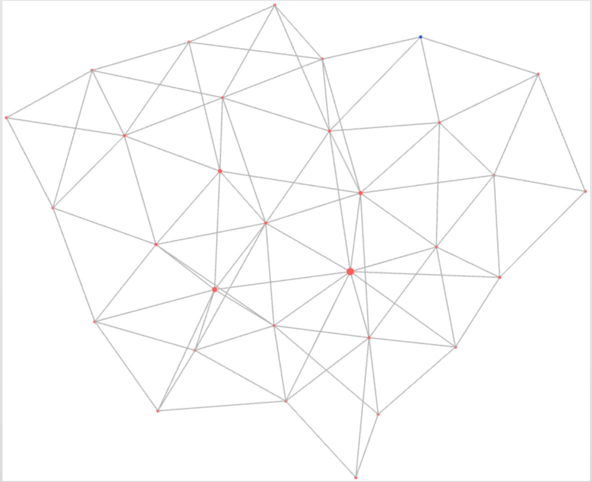
Node-node: variable node size layout

	Some clutter; less sprawl	No clutter; high sprawl
Layout		
Area-aware clutter	16.3	0
Count (baseline)	6	0

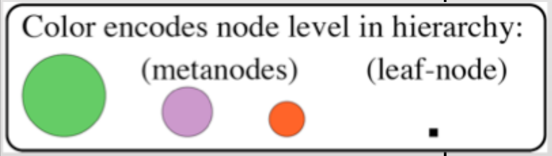
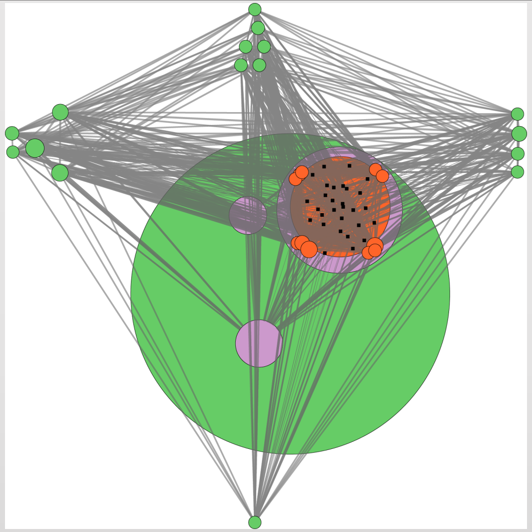
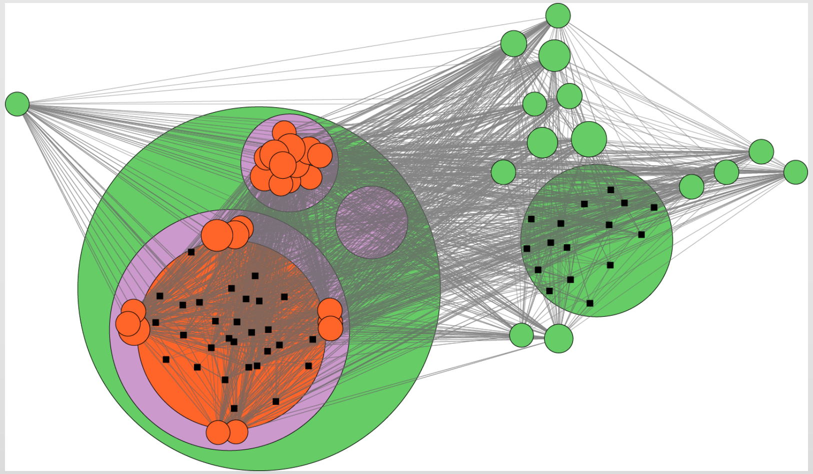
Node-node: variable node size layout

	Some clutter; less sprawl	No clutter; high sprawl
Layout		
Area-aware clutter	16.3	0
Sprawl	6.1	1605.3
Count (baseline)	6	0

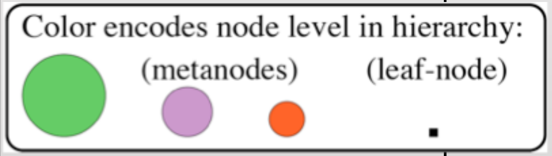
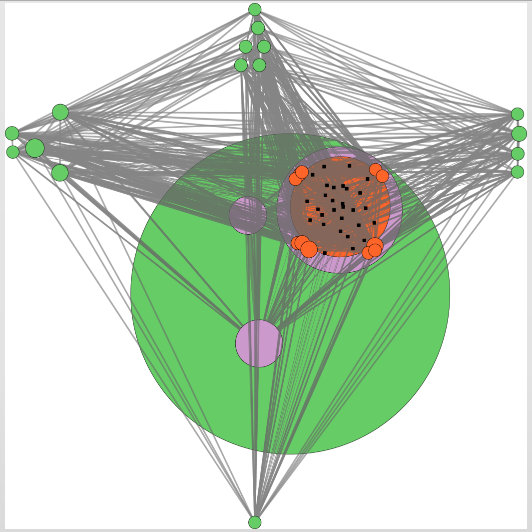
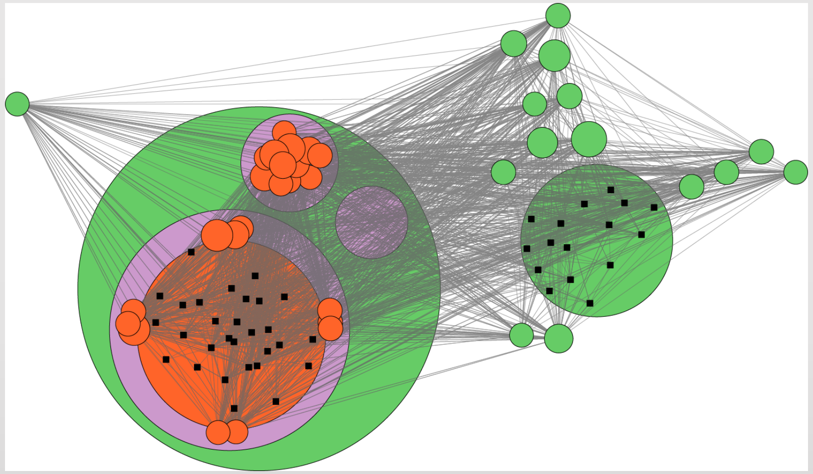
Node-node: variable node size layout

	Some clutter; low sprawl	No clutter; high sprawl
Layout		
Area-aware clutter	16.3	0
Sprawl	6.1	1605.3
Sprawlter	10.0	40.1
Count (baseline)	6	0

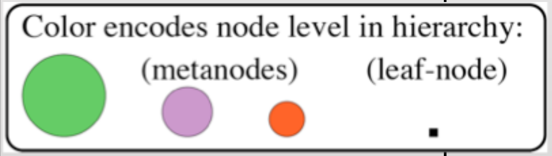
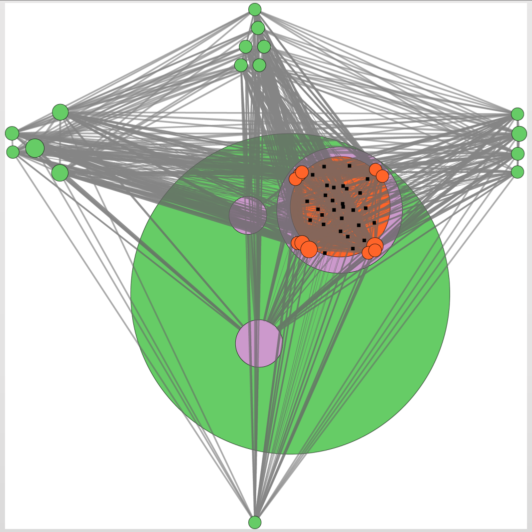
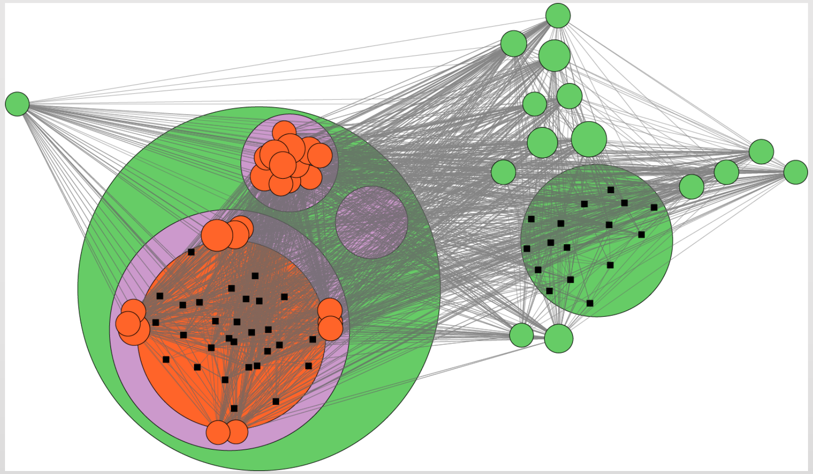
Node-node: real-world graph, multi-level layout

	Lower clutter; higher sprawl	Higher clutter; lower sprawl
Layout 		

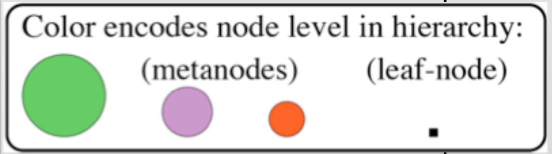
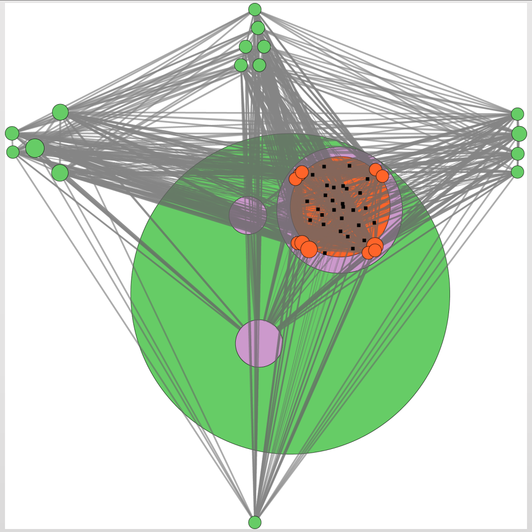
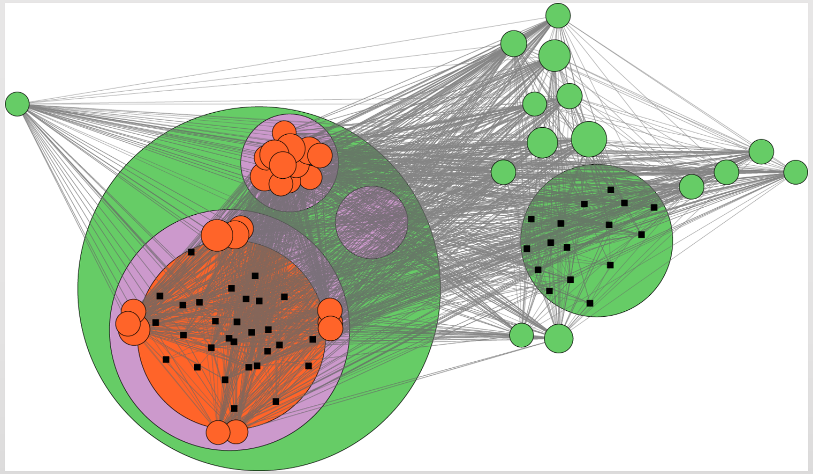
Node-node: real-world graph, multi-level layout

	Lower clutter; higher sprawl	Higher clutter; lower sprawl
Layout 		
Area-aware clutter	133.5 <	184.1
Count (baseline)	16 <	40

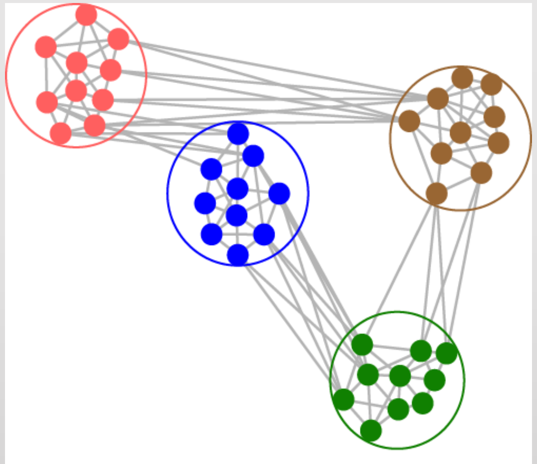
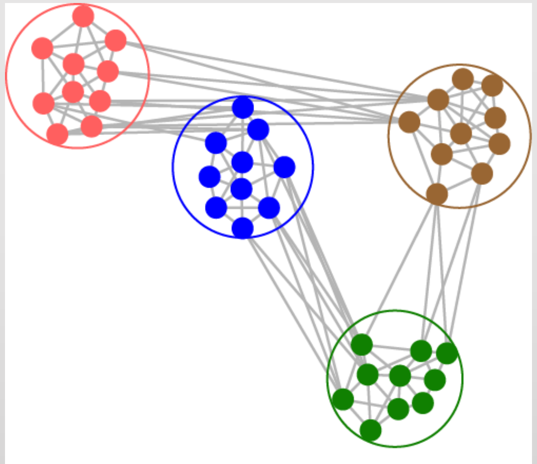
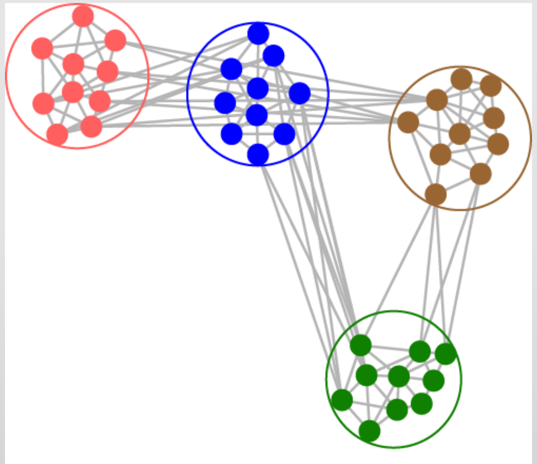
Node-node: real-world graph, multi-level layout

	Lower clutter; higher sprawl	Higher clutter; lower sprawl
Layout 		
Area-aware clutter	133.5	184.1
Sprawl	947.6 >	219.6
Count (baseline)	16	40

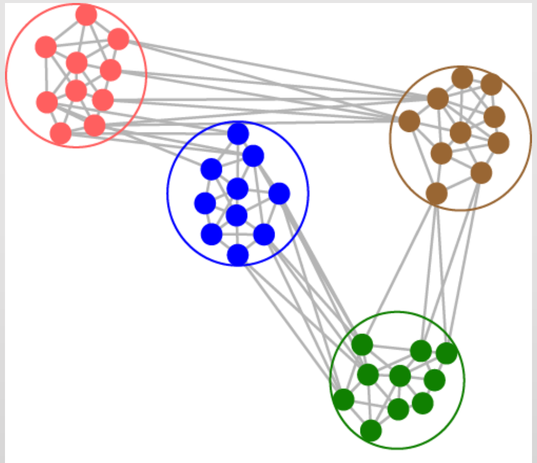
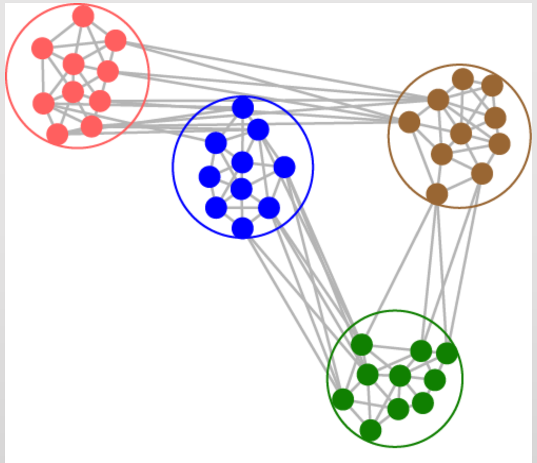
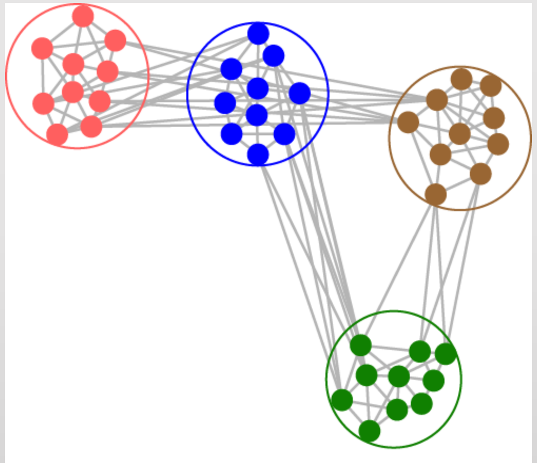
Node-node: real-world graph, multi-level layout

	Lower clutter; higher sprawl	Higher clutter; lower sprawl
Layout 		
Area-aware clutter	133.5	184.1
Sprawl	947.6	219.6
Sprawlter	355.7	201.1
Count (baseline)	16	40

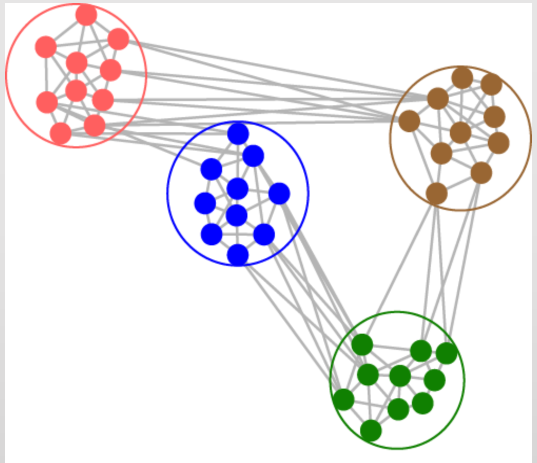
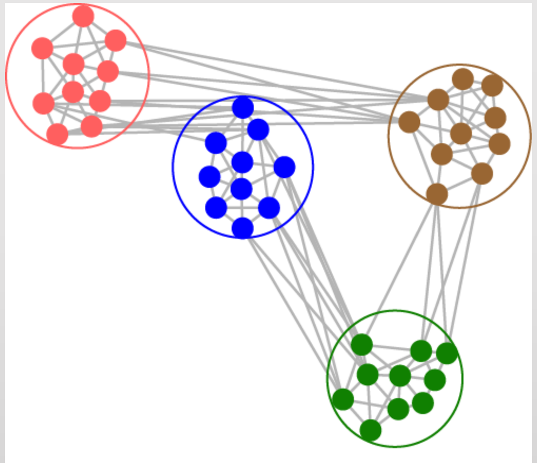
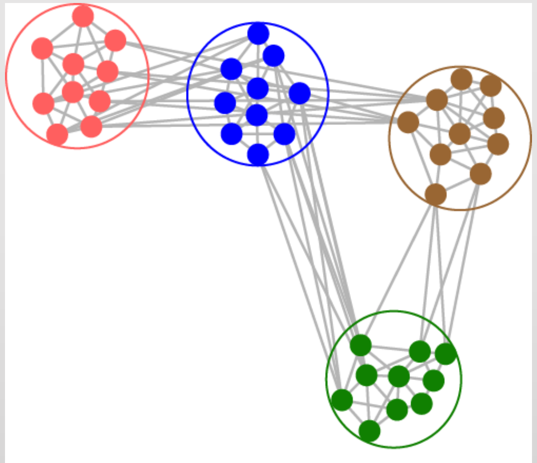
Node-edge: increasing overlap

	Near-min	Some	Near-max
Layout			

Node-edge: increasing overlap

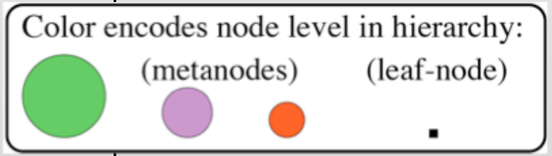
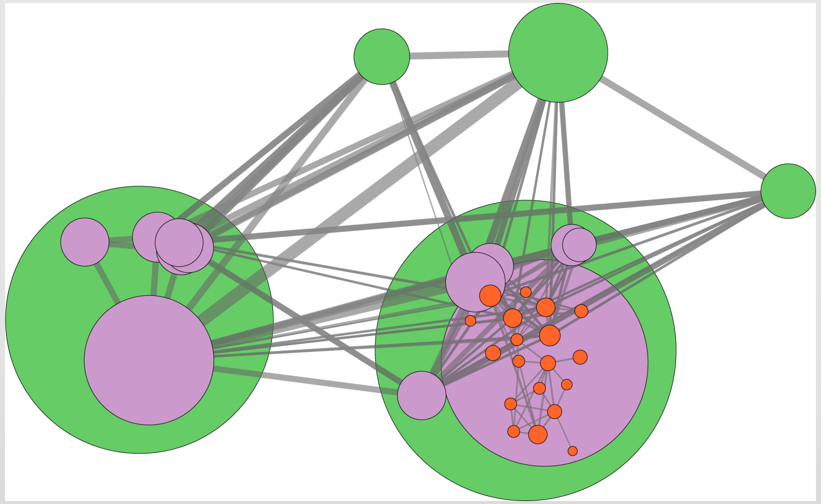
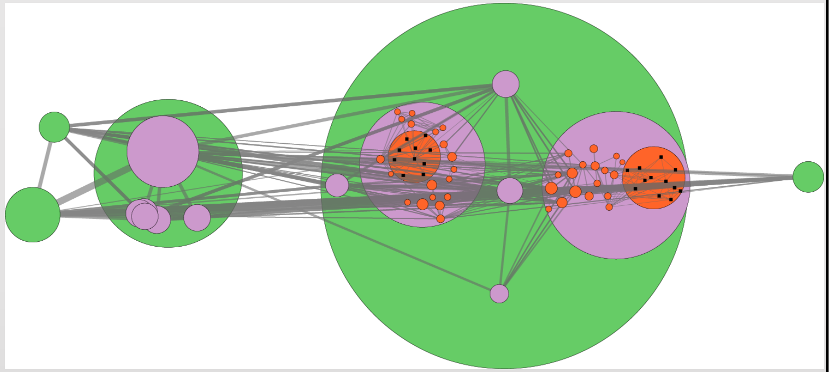
	Near-min	Some	Near-max
Layout			
Area-aware clutter	12.1	54.5	136.7
Count (baseline)	6	15	36

Node-edge: increasing overlap

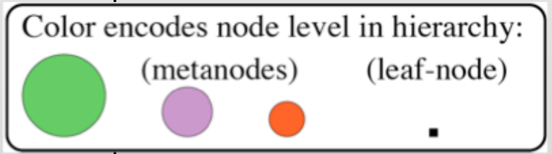
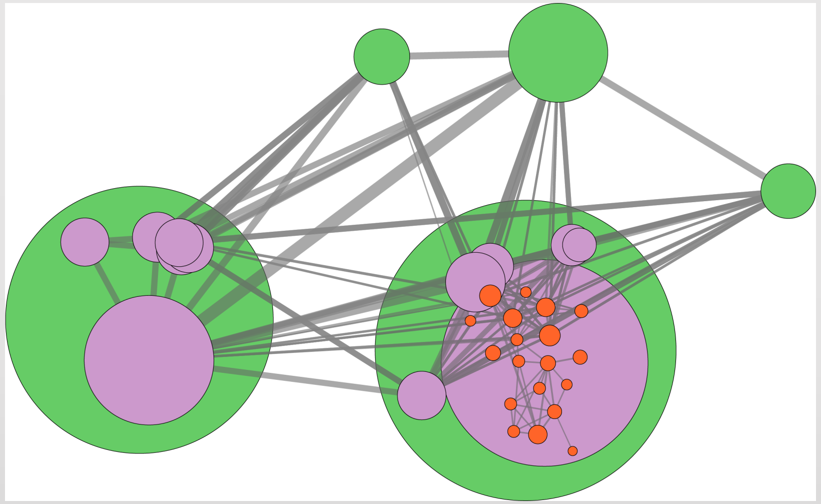
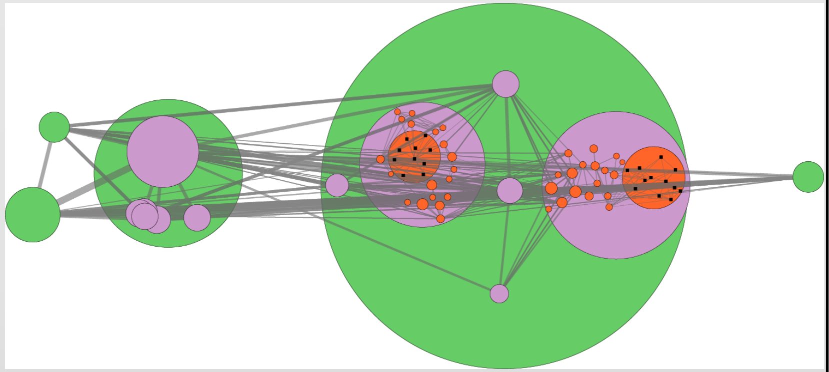
	Near-min	Some	Near-max
Layout			
Area-aware clutter	**5.4** + 6.7 = 12.1	**39.8** + 14.7 = 54.5	**97.5** + 39.2 = 136.7
Count (baseline)	**1** + 5 = 6	**5** + 10 = 15	**8** + 28 = 36

Penalty for metanodes

Edge-edge: real-world graph, multi-level layout

	Orthogonal angles	Glancing angles
Layout 		

Edge-edge: real-world graph, multi-level layout

	Orthogonal angles	Glancing angles
Layout 		
Area-aware clutter	703.7 <	5,936.9
Count (baseline)	808 <	4,154

Future work

- Incorporate sprawlter metric into layout algorithms
- Incorporate more families of readability metrics beyond clutter and sprawl
- Incorporate meta-edges with 2D area beyond 1D length

Conclusions

Propose **area-aware sprawlter** metric

- Account for geometric overlaps, beyond integer crossing counts
- Deal with multi-level layouts by design
- Handle the tradeoff between sprawl (geometric sparseness) and clutter

The Sprawlter Graph Readability Metric:
Combining Sprawl and Area-aware Clutter

Zipeng Liu, Takayuki Itoh, Jessica Q. Dawson, Tamara Munzner.
PacificVis 2020.
Invited to appear at Trans. Visualization and Computer Graphics 2020.
http://www.cs.ubc.ca/labs/imager/tr/2020/sprawlter/