networks, networks everywhere!

Eric J. Ma (MIT)

12 December 2016

slides are available online

- html notes: ericmjl.github.io/big-data-boston-2016
- slides: ericmjl.github.io/big-data-boston-2016/slides.pdf
- source: github.com/ericmjl/big-data-boston-2016

about myself

- doctoral candidate, MIT biological engineering
- self-taught pythonista
- using networks to problems in infectious disease ecology, evolution & biochemistry

outline

1. what are networks?
2. example 1: recommendation systems
3. example 2: panama papers
4. example 3: influenza ecology & evolution
5. example 4: neural networks on networks

what are networks

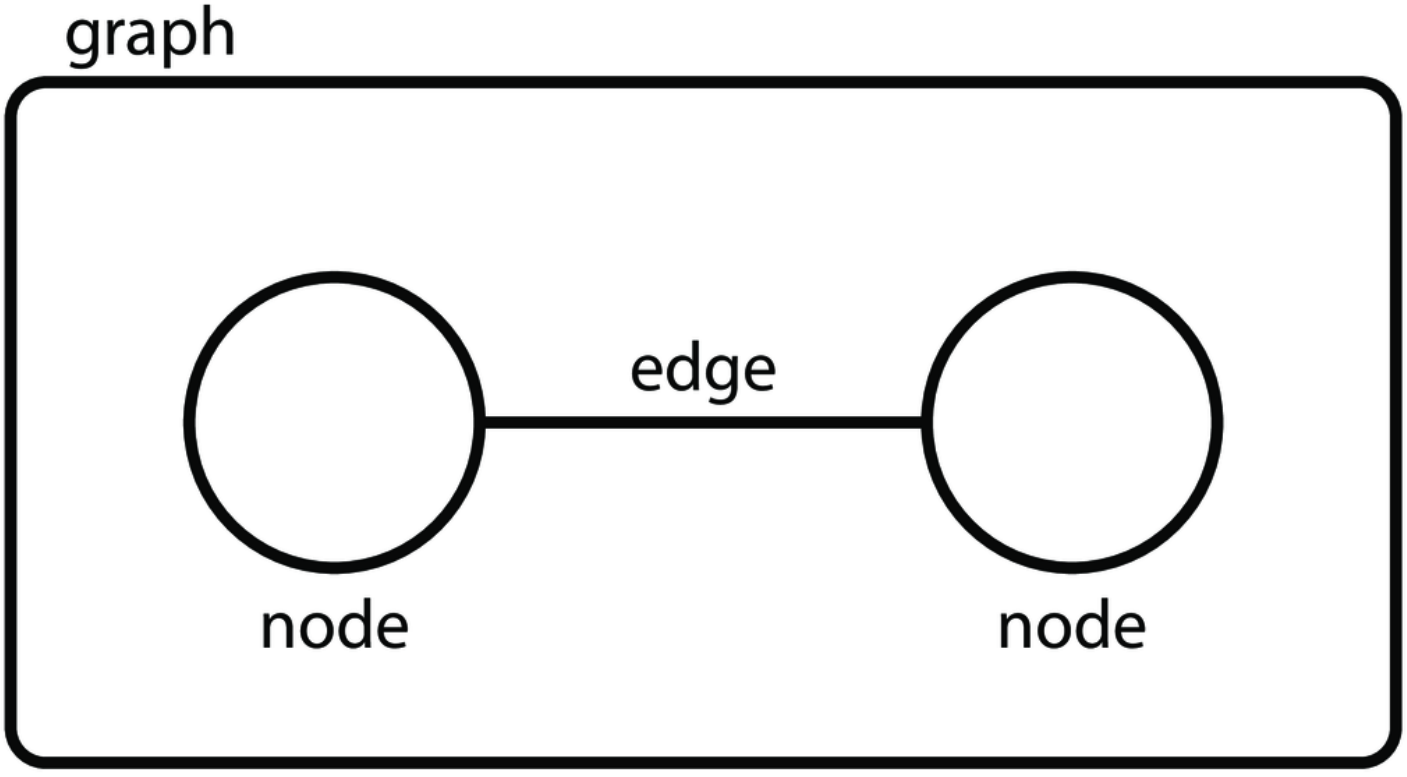

Figure 1: networks, a.k.a. **graphs**, are composed of **nodes** (circles) and **edges** (lines)

example 1: recommendation systems

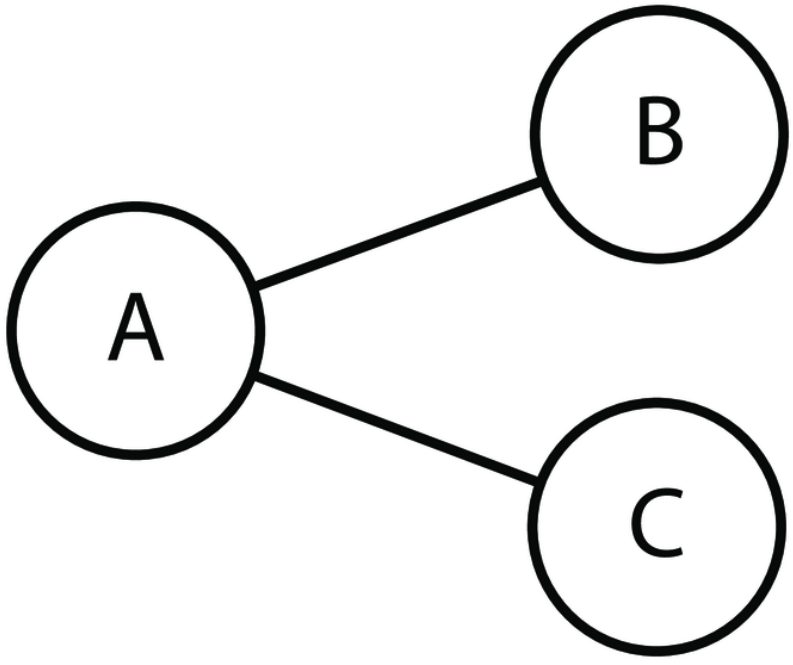

Figure 2: if `A` is connected to `B` and `C`, but `B` and `C` are not connected, then maybe they should be!

example 1: recommendation systems

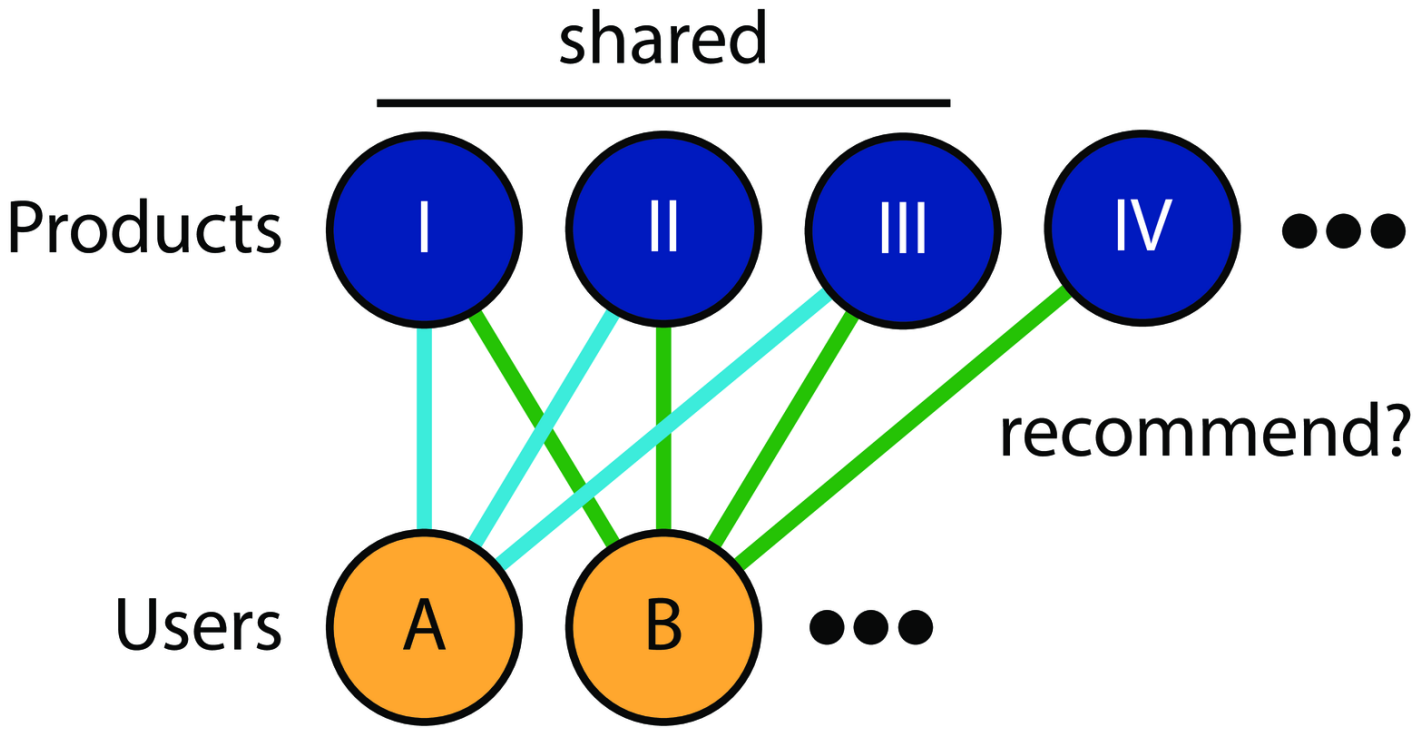

Figure 3: if A and B share overlapping interests, then maybe some of B's interests can be recommended to A.[2]

[2] Collaborative filtering

example 2: panama papers

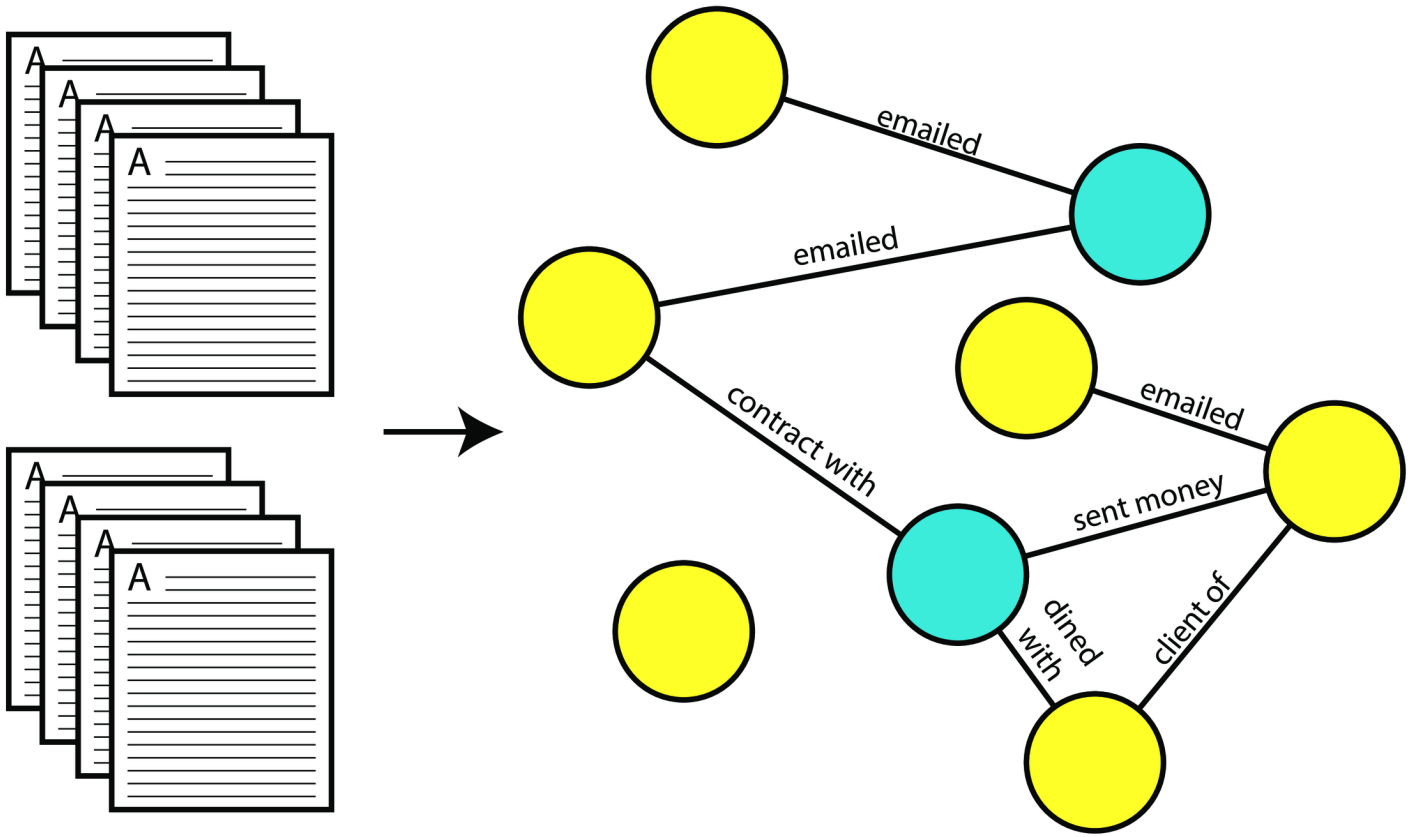

Figure 4: graph databases were used to show how the rich hide their money.[4]

[4] ICIJ and Neo4j unravel the panama papers.

example 3: influenza ecology and evolution

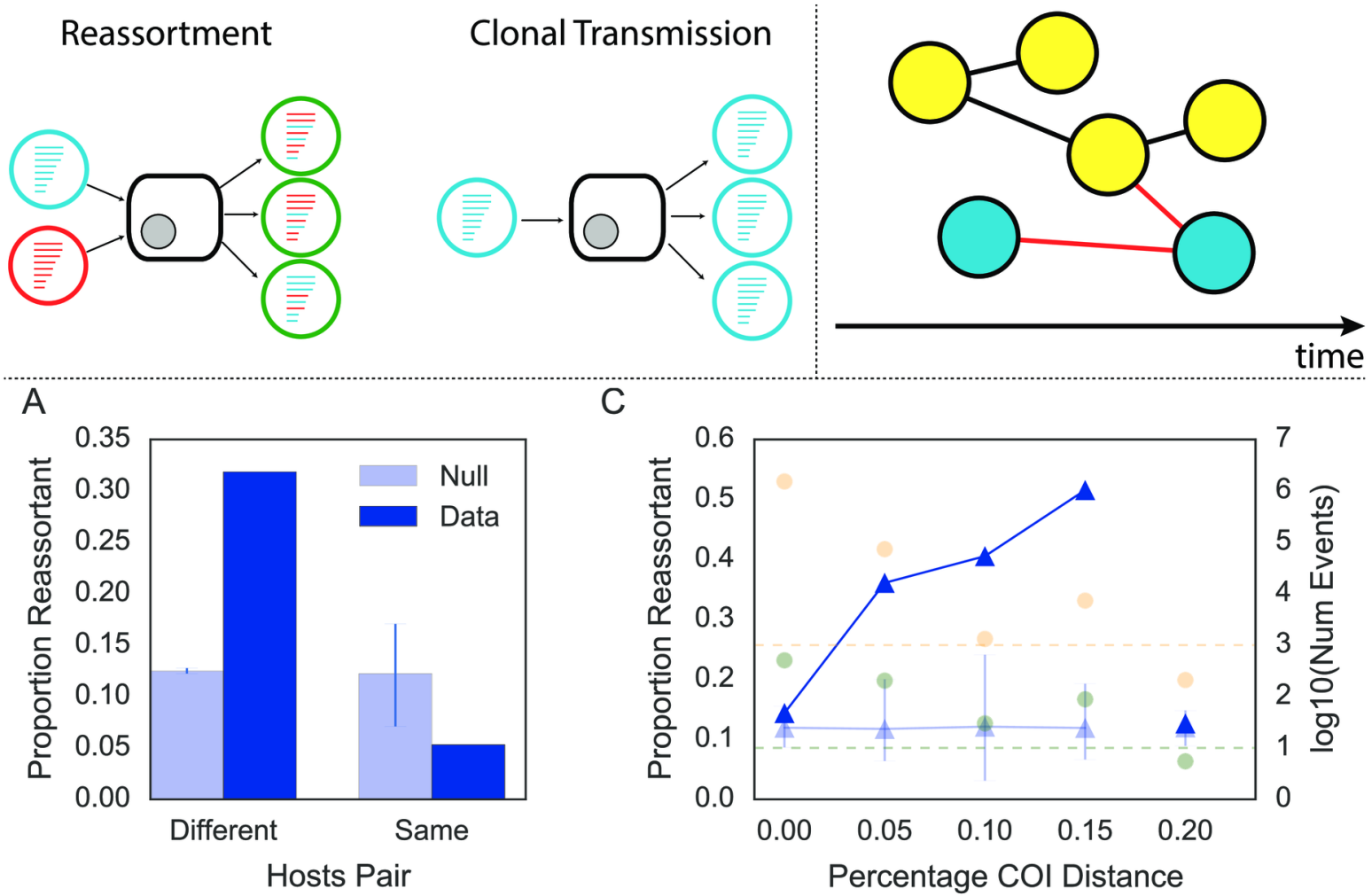

Figure 5: for influenza, gene shuffling probably helps in host switching.[5]

[5] Reticulate evolution is favoured in influenza niche switching.

example 4: neural networks on networks

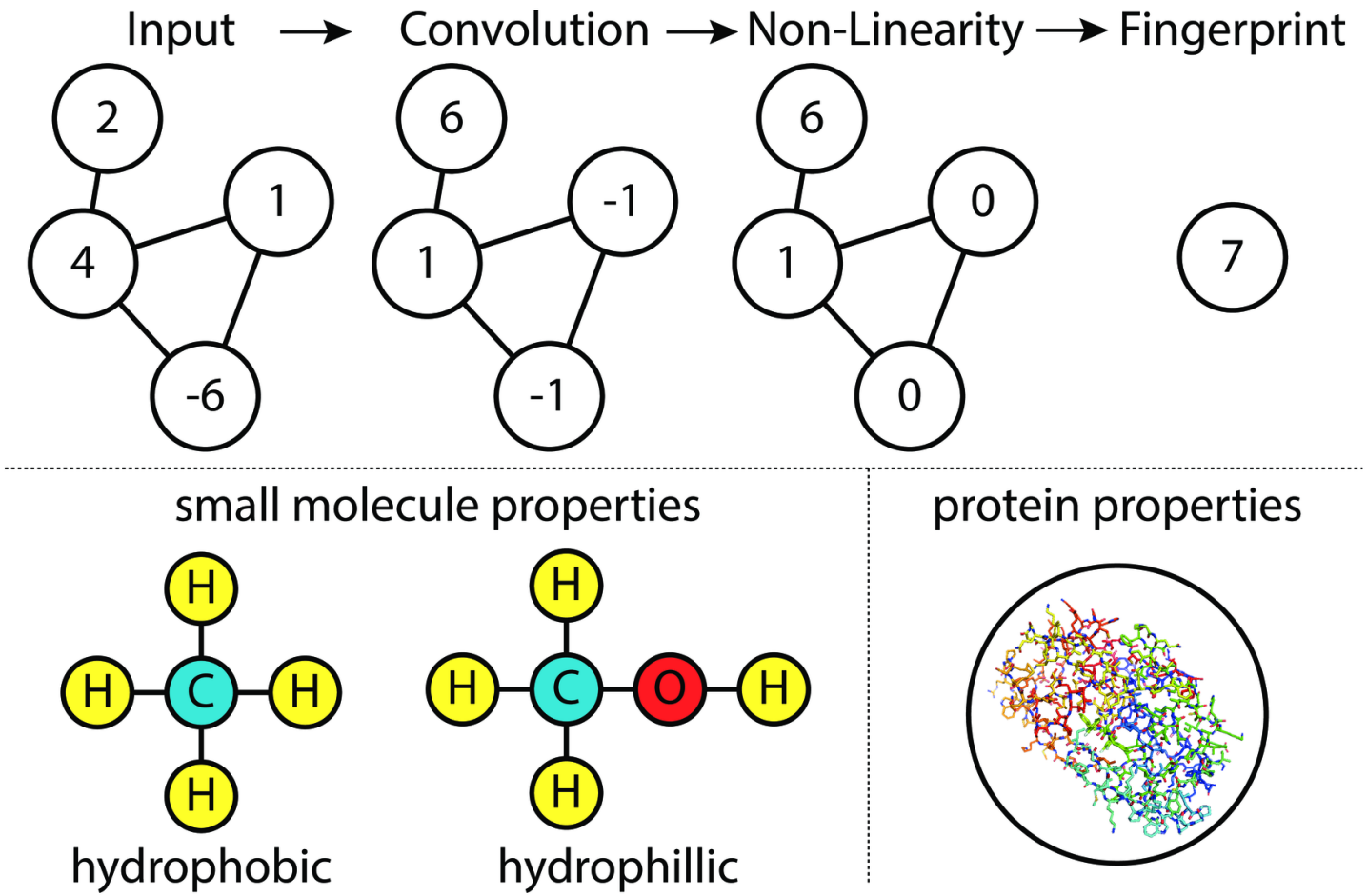

Figure 6: graph convolutions let us do machine learning on graph-structured data.[7]

[7] Convolutional Networks on Graphs for Learning Molecular Fingerprints

visualize networks rationally

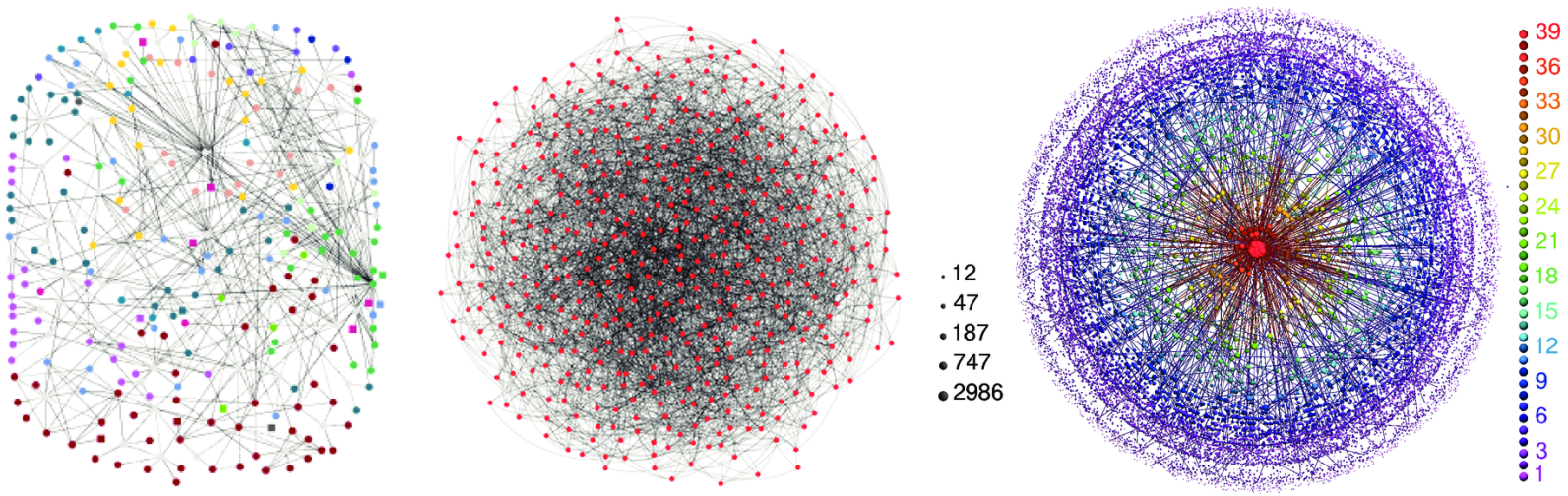

Figure 7: move away from the hairball!

visualize networks rationally

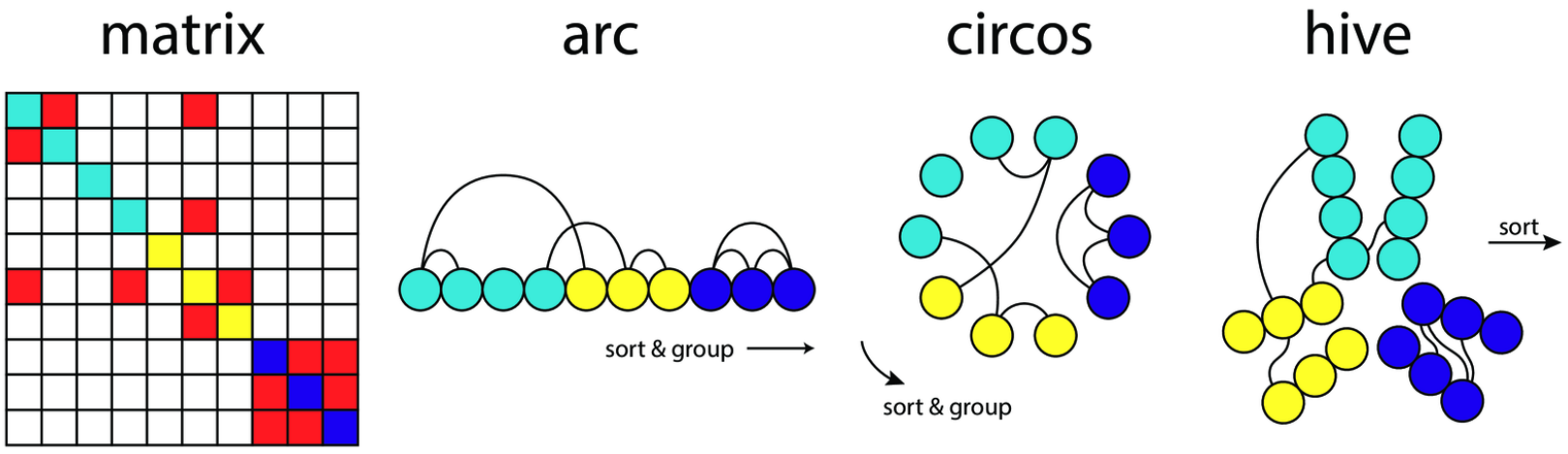

Figure 8: rational network visualizations prioritize placement of nodes

conclusions

- think relationally
- networks can be used creatively to solve all sorts of problems

keep in touch

- personal website: ericmjl.com
- linkedin: linkedin.com/in/ericmjl
- datacamp: network analysis course coming within the next few months!
- available for training your staff on data analysis and network science