Lecture 10
Networks 5

Andrew Sonta

CIVIL 534: Computational systems thinking for sustainable engineering

7 May 2025

Outline

- Percolation
- Network applications
  - Naturally-defined networks
  - Model-defined networks
  - Data-defined networks
- Networks and systems thinking

- Reading: Honti et al. 2019 (on Moodle)
- Networks Midterm next week
- Project milestone 2 due Friday
- Project final presentation on May 28

Final project deliverable

- Presentation on May 28
  - 15 minutes followed by Q+A
- What to cover:
  - What is your group's definition of sustainable development in the context of the world dynamics models?
    - Be explicit in how you construct your metric(s)
  - How did you design your automated exploration of different policies?
  - What did you find in analyzing the system dynamics model as a network?
  - Did your network analysis affect your policy development?
  - What parameters did you adjust in your final policy, and which had the biggest impact in moving your model toward your sustainable development target?
  - How does the model perform in the end?
  - What are the biggest strengths and weaknesses of both the model and your approach?

Final project presentation

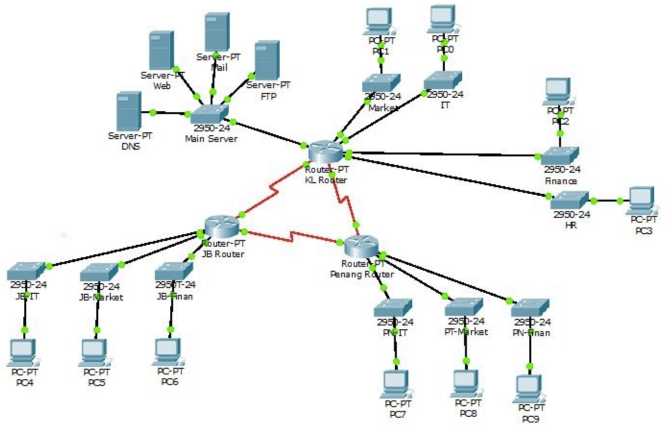
- Introduction to the project context (10 pts)
- Definition of sustainable development in the world dynamics models (20 pts)
- Network analysis of systems model (20 pts)
- Automated process for policy development (20 pts)
- Discussion of results and implications (20 pts)
- Conclusion, limitations, and outlook (10 pts)

Final project deliverable

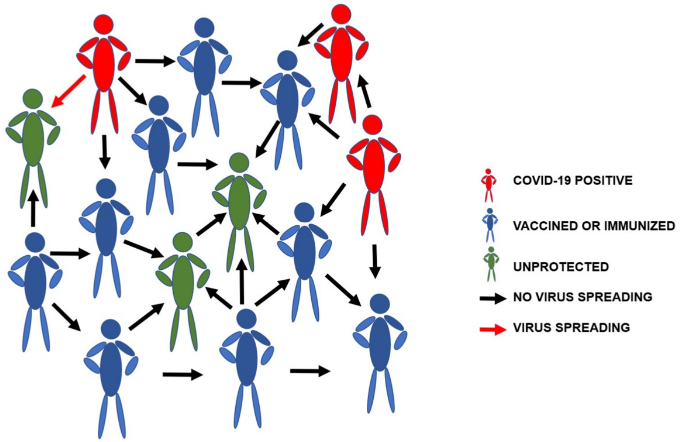
- You also need to submit a final version of your code for analysis
- Grading criteria:
  - Does it contain all the code you reference in your final presentation? (75%)
  - Is it well-documented (i.e. commented)? (25%)

Percolation

- The process of removing nodes from the network and the associated edges
- Helps us understand “what-if” scenarios
- Examples:



Internet



“Herd immunity”

Lots of urban systems!

Simulating percolation

- Key parameter: ϕ (occupational probability)
  - $\phi = 1 \rightarrow$ all nodes are present
  - $\phi = 0 \rightarrow$ all nodes have been removed
- Process of removing nodes
  - Uniform: all nodes have same probability of removal
  - Non-uniform: based on node properties (e.g., different forms of centrality)
  - The choice is normally driven by the domain

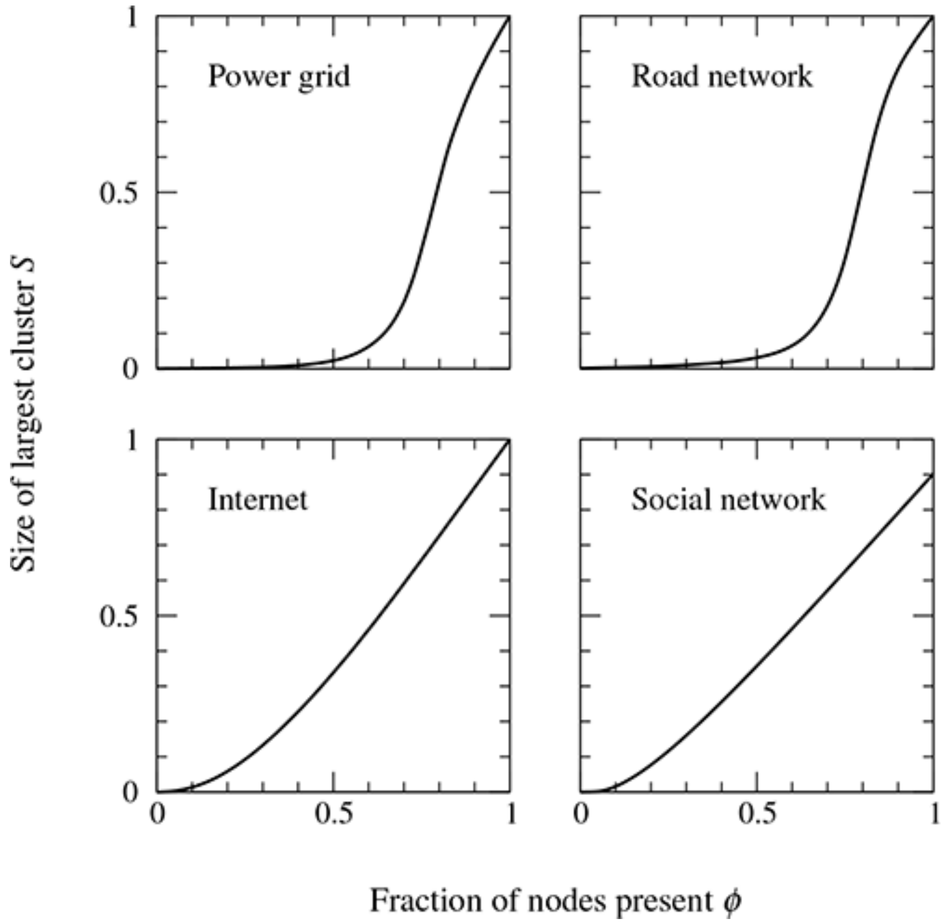
Percolation theory

- Newman goes into detail in Chapter 15
  - (And there can be a lot of mathematical detail)
- Percolation is an entire subfield of network science
- Offers an understanding of network resilience and robustness

Qualitative percolation analysis example

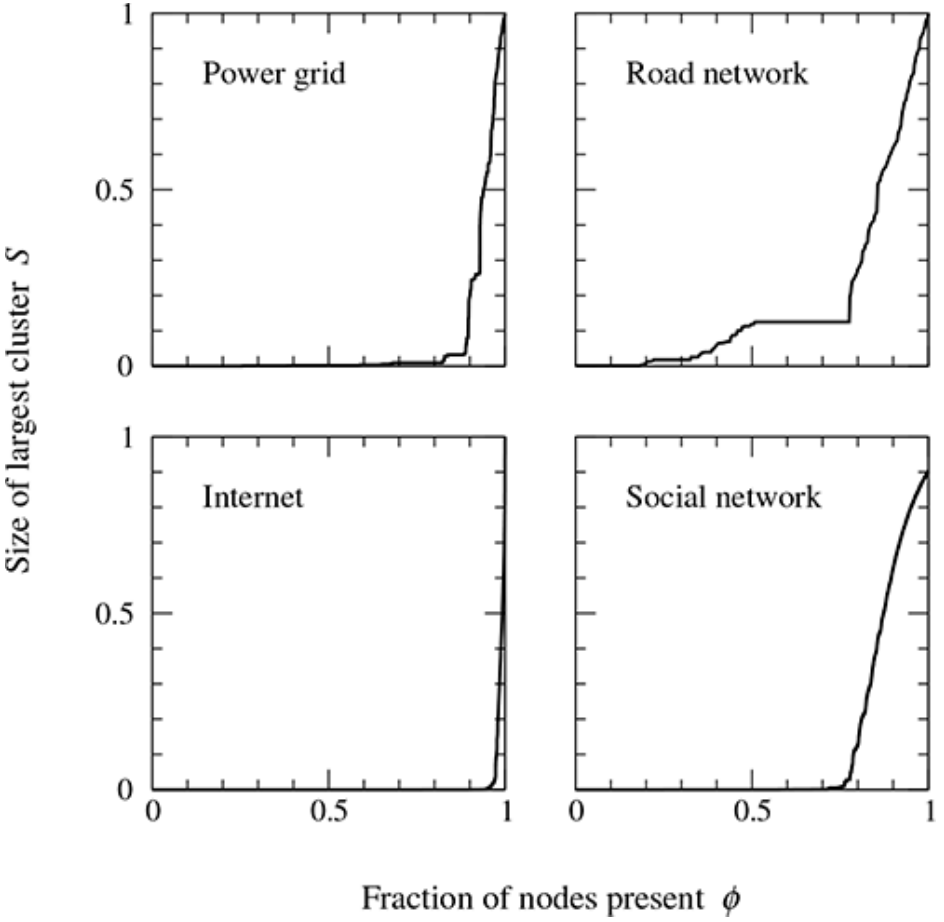
Uniform removal of edges

What does this signify?



Qualitative percolation analysis example

Targeted edge removal (target largest degree first)



General network analysis procedure

- Define network structure
- Analyze structure & extract insights
- Perturb structure to understand how changes impact performance
- Make design changes / recommendations for the system

Networks in urban systems

Naturally-defined network structure

Network is defined by the physical or natural structure of the data / system:

- Social network (people, relationships)
- Transportation network (roads, intersections) (destinations, paths)
- Water network (valves, pipes)
- Power grid (homes/transformers/etc., power lines)

Model-defined network structure

Network is defined by a model (models are imperfect):

- Theoretical between two entities in a city
  - How similar are two buildings?
  - How much do two buildings impact one another?
- Dynamic networks (processes or changes across edges/nodes)

Data-defined network structure

Network structure is learned from data

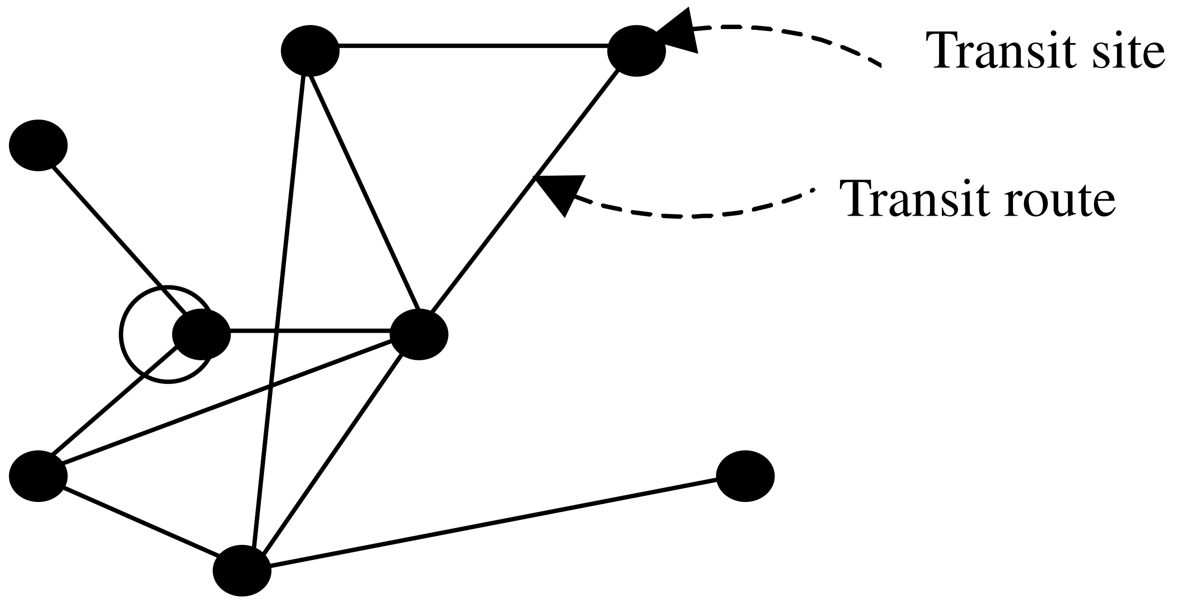
- Open problem in many domains
- Can help to understand interdependencies between infrastructure / systems / people
- Example: data on human behavior – learn social network structure?

Naturally-defined networks

Define network structure	✅ naturally defined
Analyze structure & extract insights	Structure: metrics from class Insights: e.g. node importance, edge importance What can we say about the system as currently designed?
Perturb structure to understand how changes impact performance	How does changing the capacity of edge ***x*** or node ***y*** affect performance?
Make design changes / recommendations for the system	Make decisions based on alternative analysis

Example: Beijing transit network

Model transit sites as nodes and transit routes as edges



Wu, J., Gao, Z., Sun, H., & Huang, H. (2004). Urban transit system as a scale-free network. *Modern Physics Letters B*, *18*(19n20), 1043-1049.

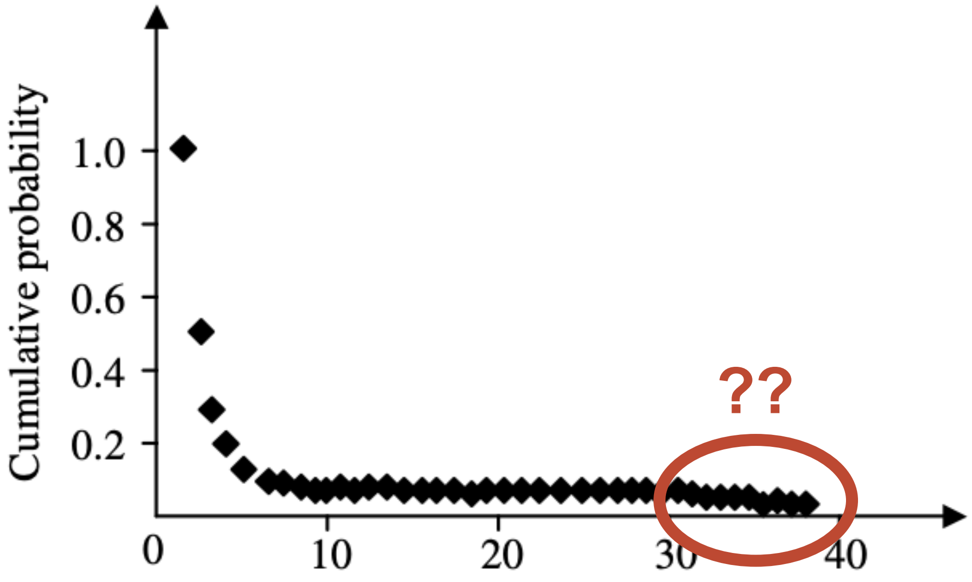
Example: Beijing transit network

Empirical analysis of Beijing urban transit network *(441 nodes, 776 edges)*

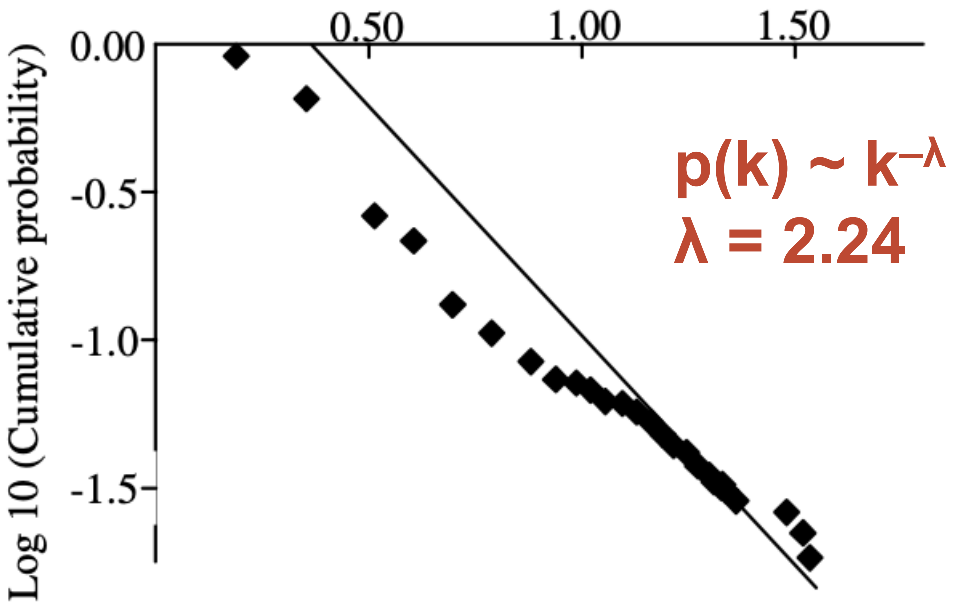
Wu, J., Gao, Z., Sun, H., & Huang, H. (2004). Urban transit system as a scale-free network. *Modern Physics Letters B*, *18*(19n20), 1043-1049.

Example: Beijing transit network

Demonstrates properties of a scale-free network



Degree



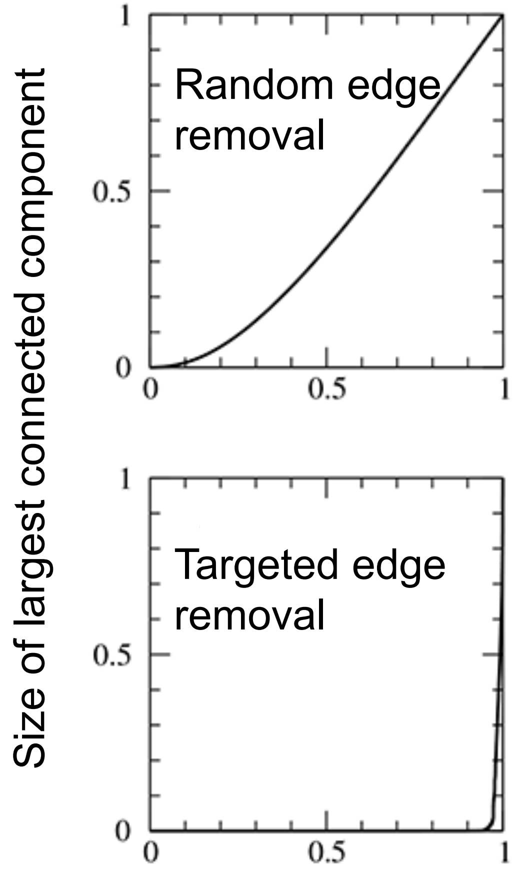
Log (Degree)

Example: Beijing transit network

Implications for urban transit:

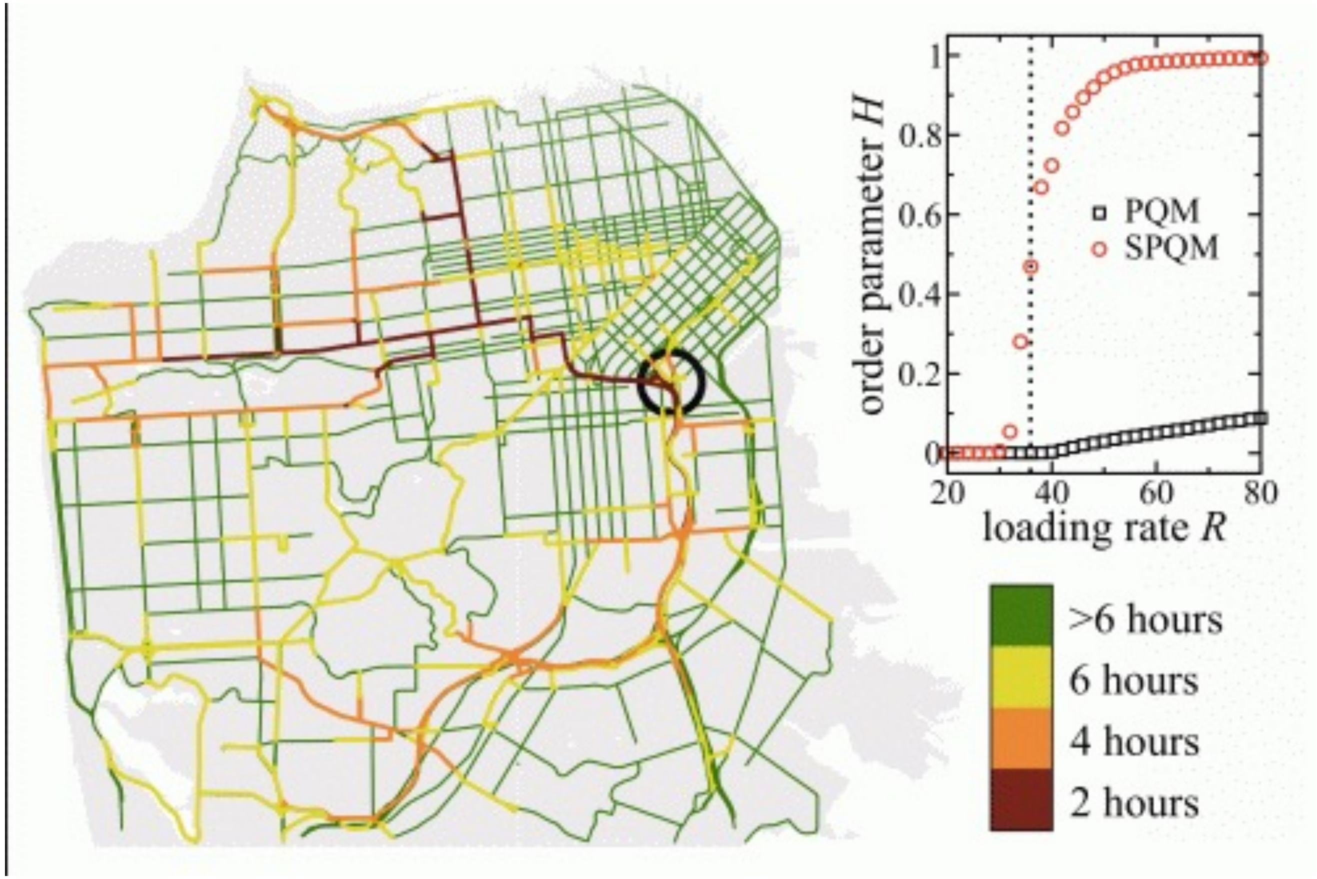
- *Effective in propagation* – people can move around across the entire network efficiently (hubs facilitate this)
- *Resistant to accidental (random) failures* – this is because of its nonhomogeneous topology (a few very important hubs)*;* non-hub breakdowns/accidents won't destroy connectivity
- *Vulnerable to targeted attacks* – breakdown at a hub can be devastating to the network (ex: July 11 rainstorm); where should you put resources for resiliency planning?

For scale-free networks:



Example: Traffic on a road network

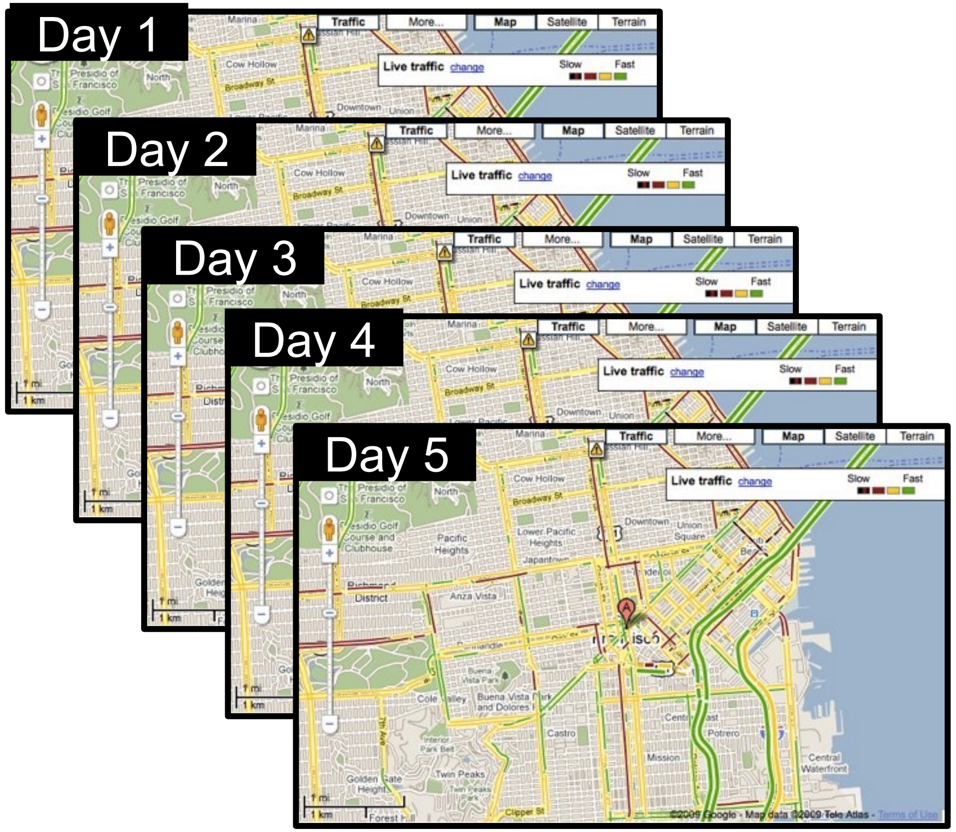
- Nodes: Intersections
- Edges: Roads
  - Weight: capacity or traffic (depending on analysis)
- Process:
  - Analyze structure
  - Extract insights
  - Perturb network
  - Make design recommendations



Serdar Çolak *et al* 2013 *New J. Phys.* **15** 113037

Example: Traffic on a road network

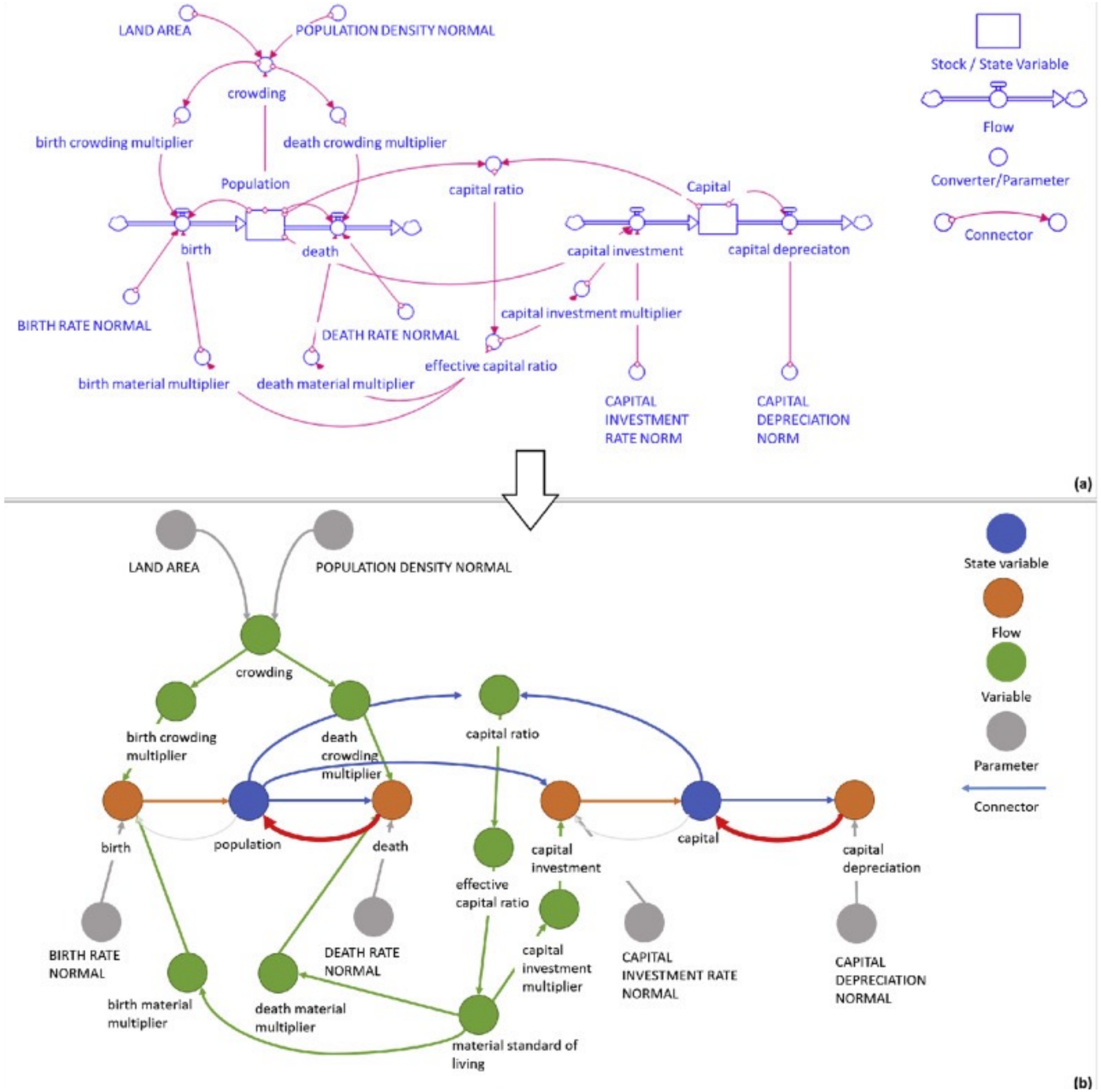
- Nodes: Intersections
- Edges: Roads
  - Weight: capacity or traffic (depending on analysis)
- Process:
  - Analyze structure
  - **Extract insights**
    - **What if networks is dynamic with time-series data? How do things change?**



Serdar Çolak *et al* 2013 *New J. Phys.* **15** 113037

Example: System dynamics model

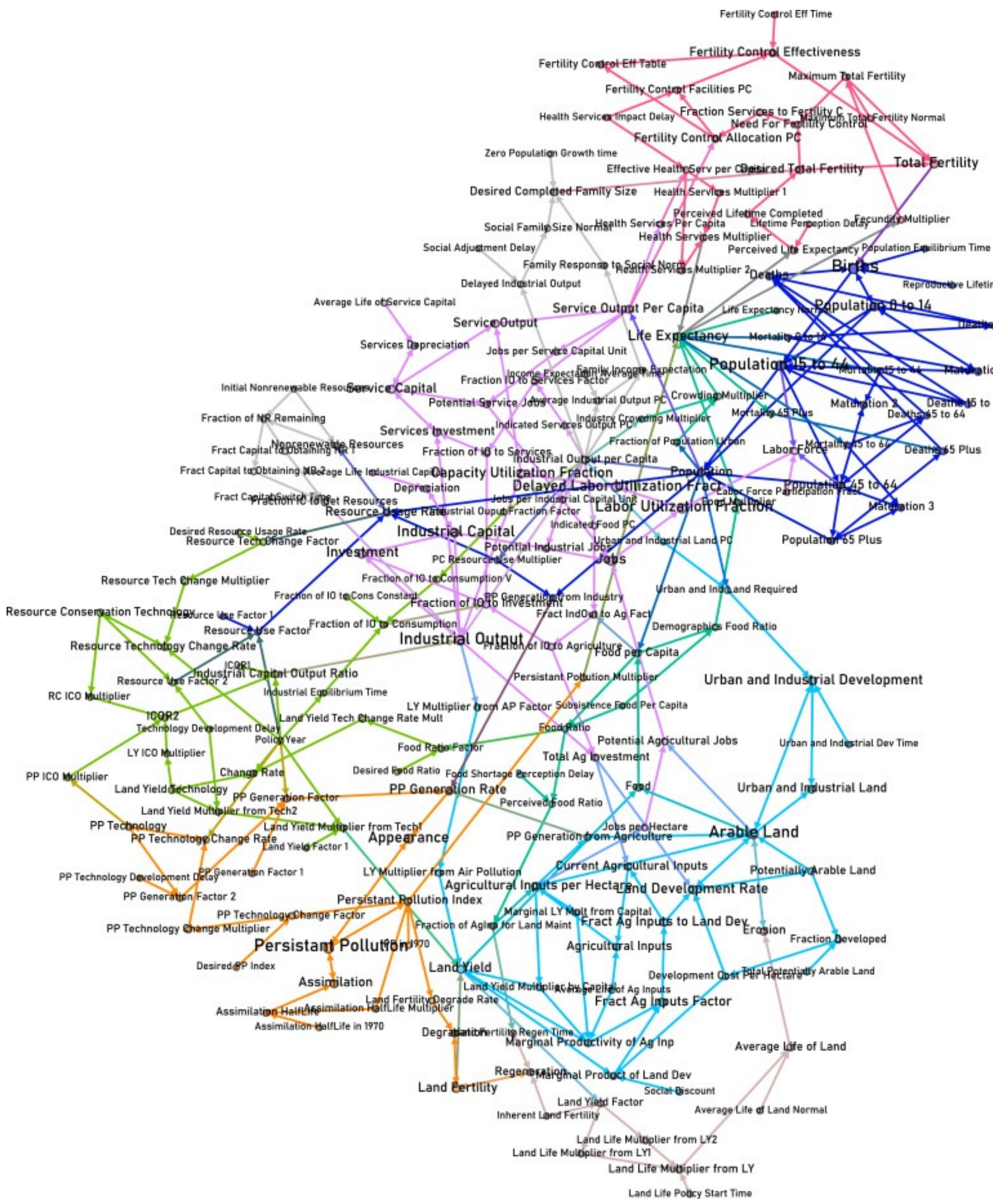
- All types of variables become nodes
    - Stocks
    - Flows
    - Parameters/Variables
- All influence arrows become edges
- All flow-stock pairs become edges



[Honti et al. (2019). Review and structural analysis of system dynamics models in sustainability science. *Journal of Cleaner Production*, *240*(118015)]

Example: System dynamics model

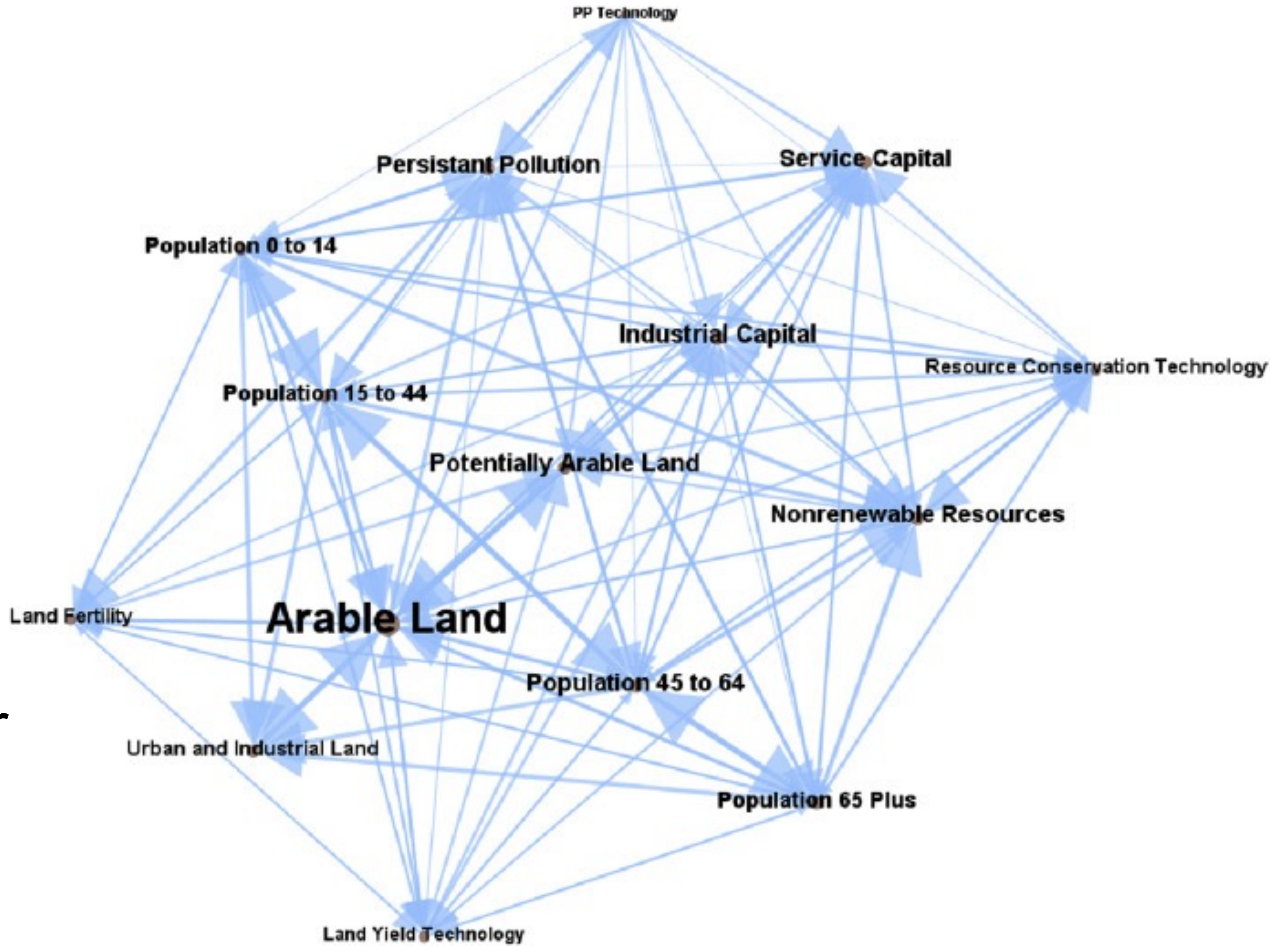
- Network diagram of world3 model



Honti et al. (2019). Review and structural analysis of system dynamics models in sustainability science. *Journal of Cleaner Production*, *240*(118015)

Example: System dynamics model

- Example where authors created a subgraph with only the stocks
- The size of the text corresponds to the eigenvector centrality
- How could this impact our understanding of the system?
- See reading on Moodle



Honti et al. (2019). Review and structural analysis of system dynamics models in sustainability science. *Journal of Cleaner Production*, *240*(118015)

Model-defined networks

- Define network structure
- Analyze structure & extract insights
- Perturb structure to understand how changes impact performance
- Make design changes / recommendations for the system

Derive relationships / network based on known model
Remember: all models are wrong (imperfect), but some are useful

- Useful to compare to ground truth network if known

Do patterns emerge that match the model?

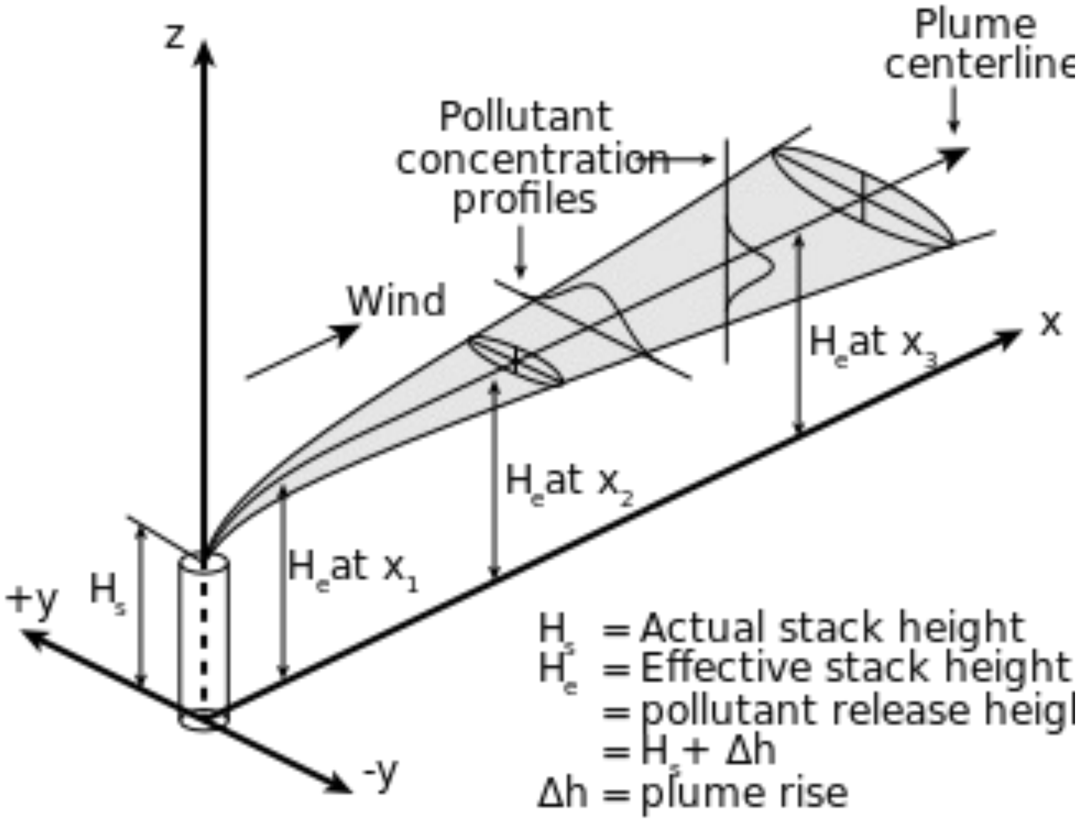
Example: Building pollution

Photo taken on Jan 8, 2014 in downtown Brooklyn, NY
by Rishee Jain, Stanford University

- Boilers produce more $PM_{2.5}$ than cars and trucks (City of NY, 2012)
- Local pollution: +2,300 deaths; +4,800 ER asthma visits
  (City of NY, 2013)
- **Goal: identify clusters of buildings for infrastructure re-design**

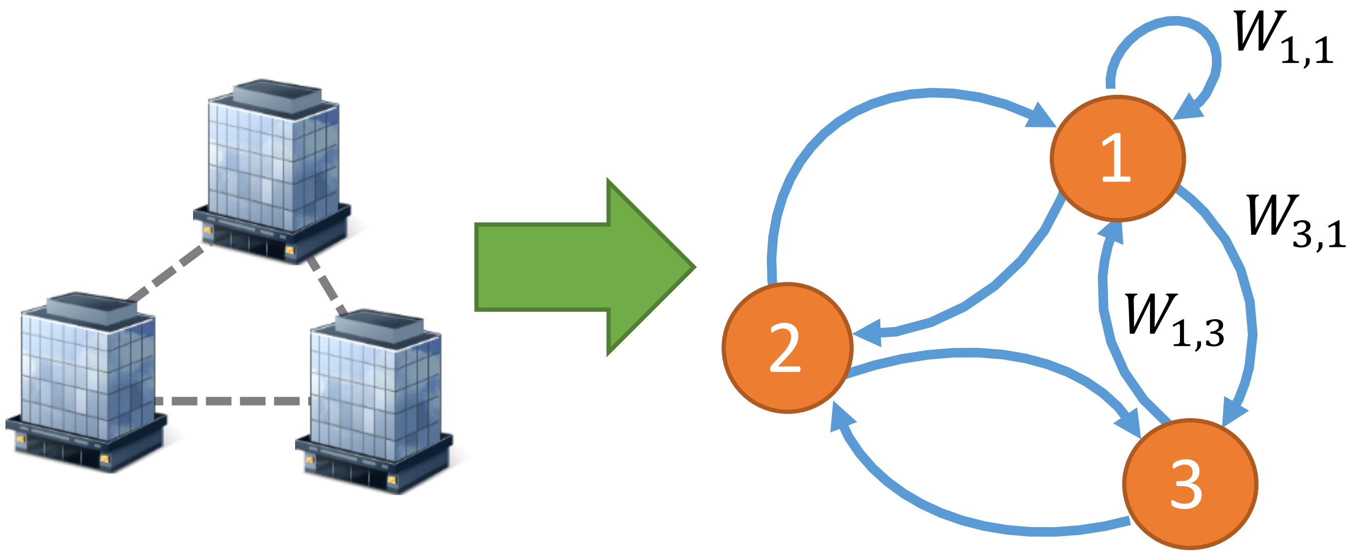
Example: Building pollution

Model to define network:
Gaussian Plume Model



Example: Building pollution

Nodes = buildings; Edges = pollution dispersion relationship



$$\mathbf{x_1} = \begin{bmatrix} EC_1 \\ \vdots \\ FLRS_1 \end{bmatrix}$$

$\mathbf{x_1}$ is the signal vector of building (node) 1 with EC = energy consumption, $FLRS$ = number of floors

R. K. Jain, J. M. F. Moura and C. E. Kontokosta, "Big Data + Big Cities: Graph Signals of Urban Air Pollution [Exploratory SP]," in *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 130-136, Sept. 2014,

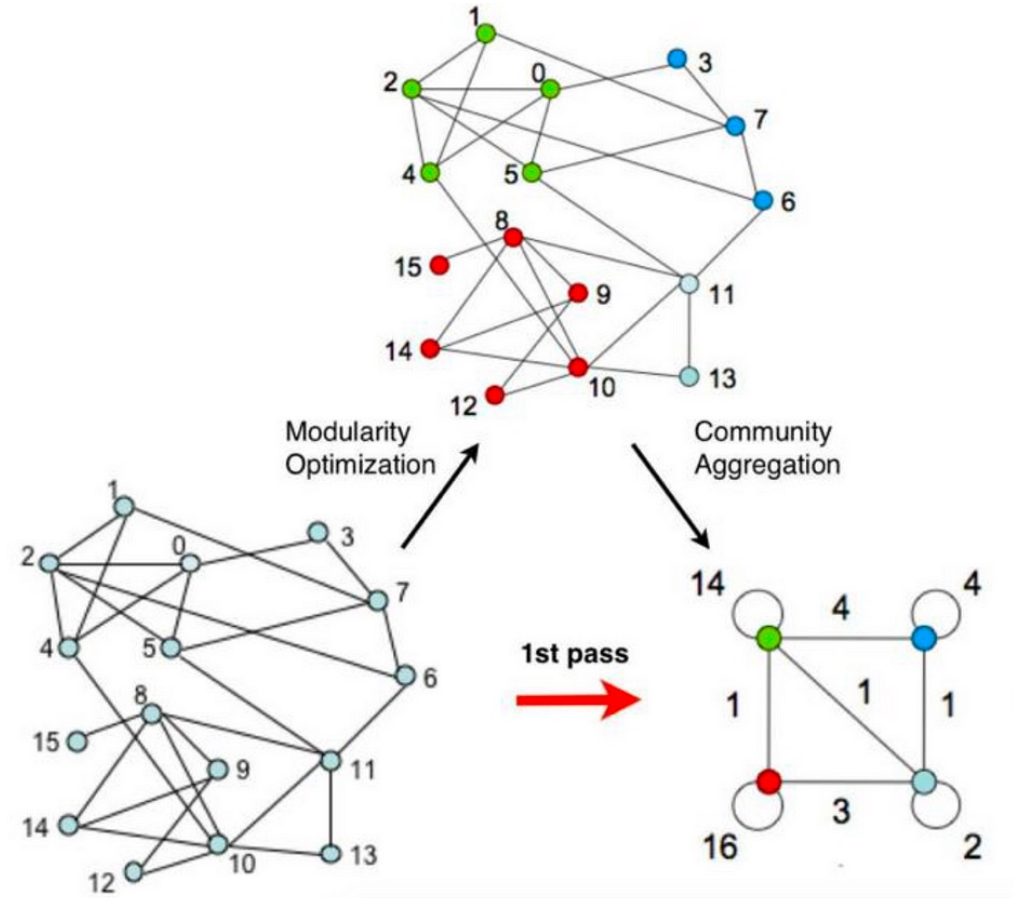
Example: Building pollution

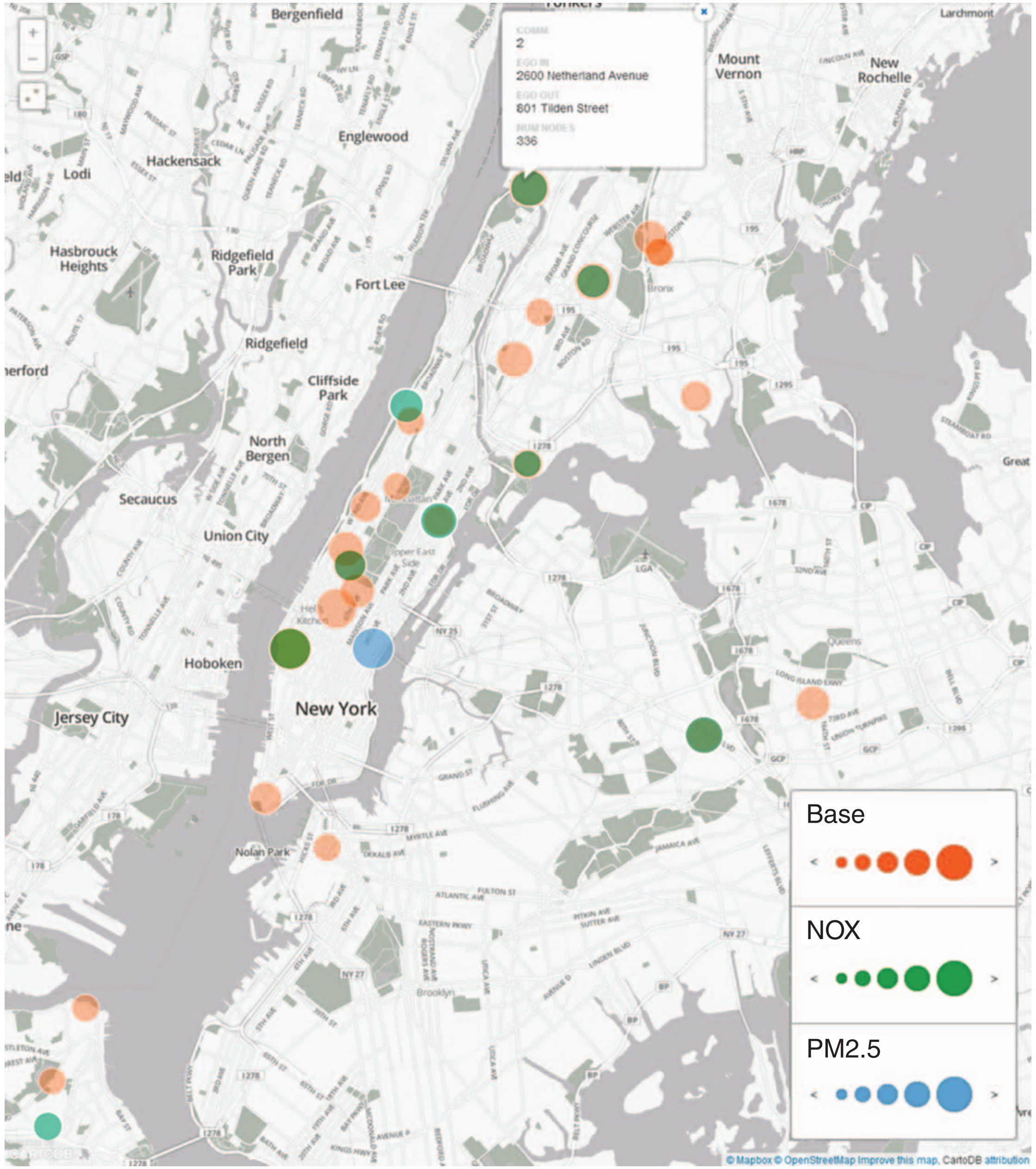
[TABLE 2] THE COMMON GRAPH METRICS APPLIED TO URBAN AIR POLLUTION.

GRAPH METRIC	FORMAL DEFINITION (DEFINITIONS ADAPTED FROM [11])	MATHEMATICAL FORMULATION	APPLICATION TO URBAN AIR POLLUTION
SELF-LOOP OF A VERTEX	THE VALUE OF AN EDGE CONNECTING A VERTEX TO ITSELF.	$W_{i,j}$ WHERE $i = j$	A MEASURE OF THE EMISSIONS OF BUILDING i.
IN-DEGREE OF A VERTEX	THE VALUE OF ALL IN-GOING EDGES CONNECTED TO A VERTEX ON A DIRECTED GRAPH.	$k_i^{\text{in}} = \sum_{j=1}^{N} W_{i,j}$	A MEASURE OF HOW MUCH NEIGHBORING BUILDINGS ARE CONTRIBUTING TO BUILDING i's AIR QUALITY.
OUT-DEGREE OF A VERTEX	THE VALUE OF ALL OUT-GOING EDGES CONNECTED TO A VERTEX ON A DIRECTED GRAPH.	$k_j^{\text{out}} = \sum_{i=1}^{N} W_{i,j}$	A MEASURE OF HOW MUCH BUILDING j IS CONTRIBUTING TO THE AIR QUALITY OF ITS NEIGHBORS.

R. K. Jain, J. M. F. Moura and C. E. Kontokosta, "Big Data + Big Cities: Graph Signals of Urban Air Pollution [Exploratory SP]," in *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 130-136, Sept. 2014,

Example: Building pollution Community detection



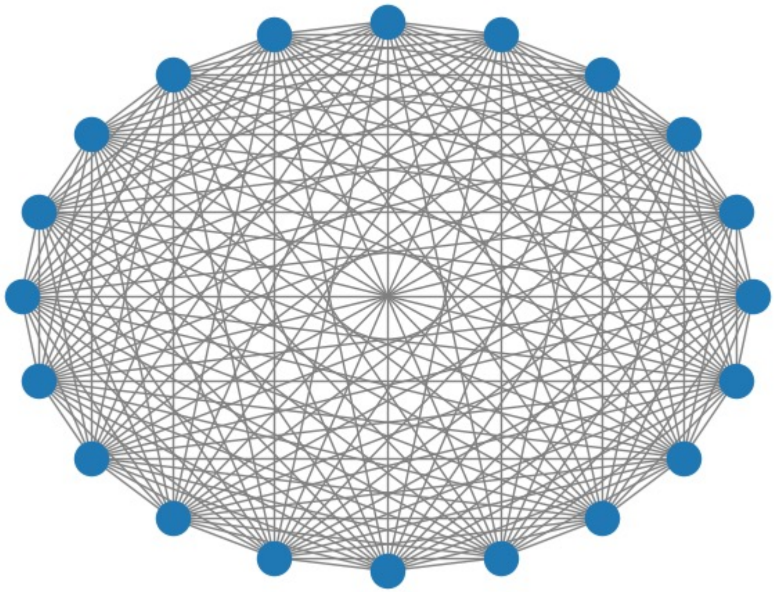


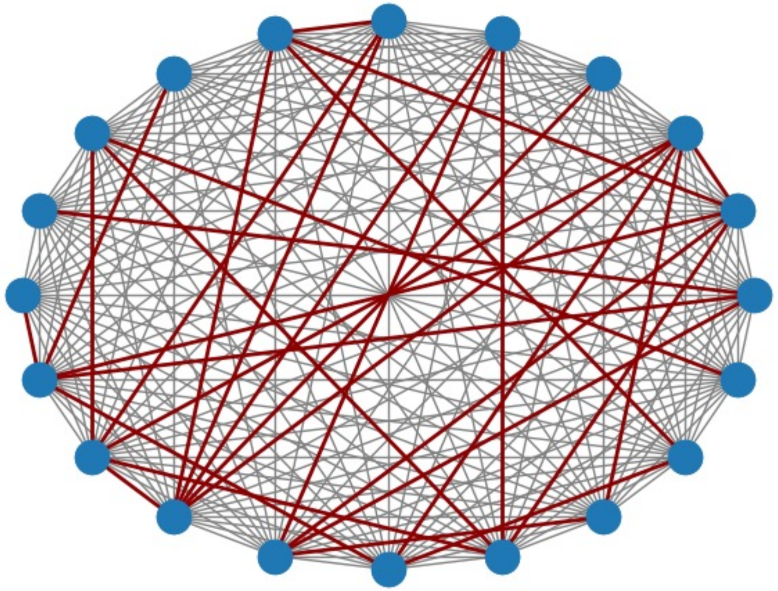
R. K. Jain, J. M. F. Moura and C. E. Kontokosta, "Big Data + Big Cities: Graph Signals of Urban Air Pollution [Exploratory SP]," in *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 130-136, Sept. 2014,

Network thresholding

- Some models / data give you a fully connected network
  - Example: network defined as the similarity in building energy use profiles
- We often threshold such graphs to identify edges deemed “significant”

$$e_{ij} = \begin{cases} e_{ij}, & e_{ij} \geq \alpha \\ 0, & e_{ij} < 0 \end{cases}$$





Data-defined network

- Define network structure
- Analyze structure & extract insights
- Perturb structure to understand how changes impact performance
- Make design changes / recommendations for the system

1. Feature extraction from data
2. Need to find graph structure that is the “best” fit
   - Difficult problem and domain specific

Example: Occupant network inference

- Example from our lab's research

Building systems

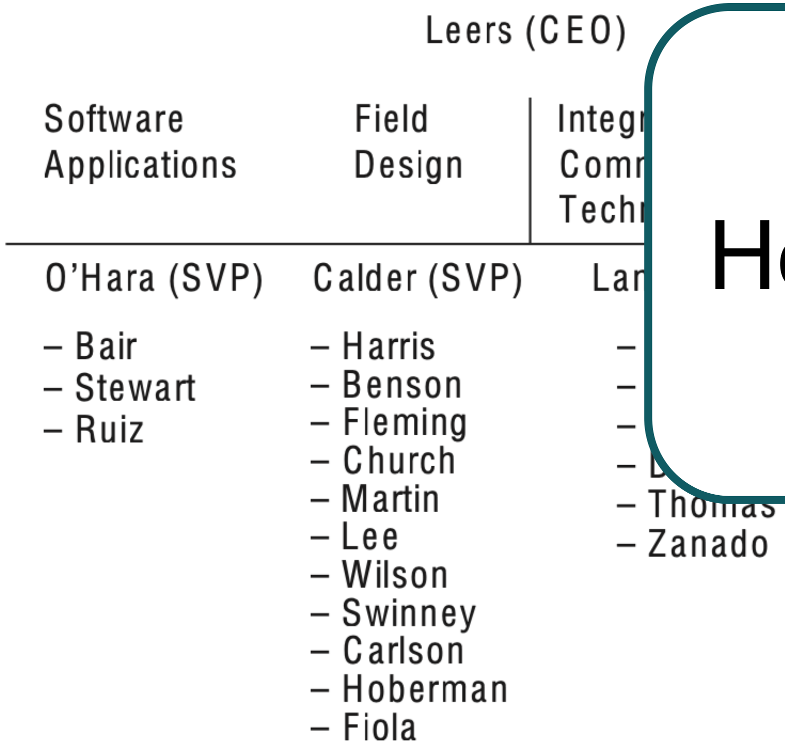
Data-driven design & management

Occupants

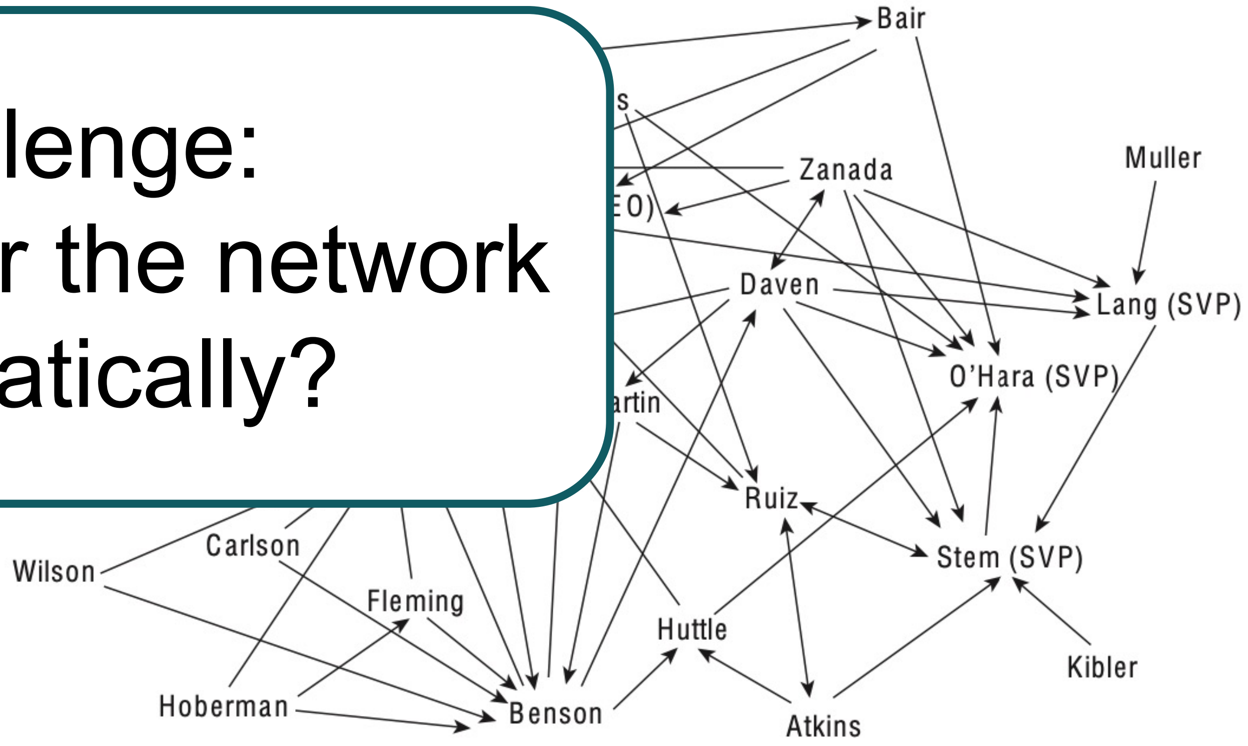
Co-optimization of organizations and building systems

Research gap: understanding networks

The formal chart shows who's on top



The advice network reveals the experts

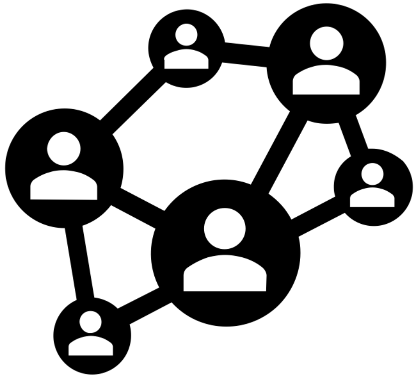


Challenge:
How to infer the network automatically?

Krackhardt and Hanson 1993

How to infer network automatically?

Learn graph $G = (\mathcal{V}, \mathbf{A})$
from time series



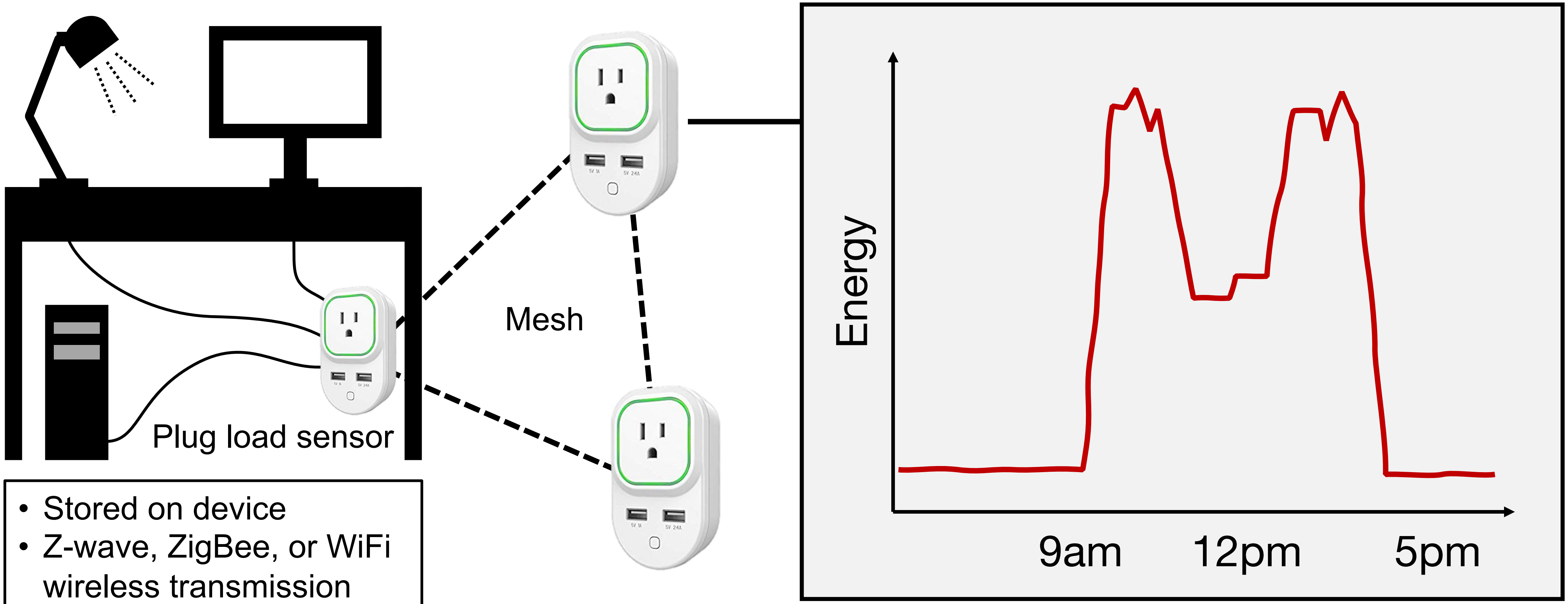
$\mathcal{V}$: set of nodes (occupants)
$\mathbf{A}$: weighted adjacency matrix

Develop or adapt method to infer network

Compare against ground truth

Hypothesis: novel method infers network structure similar to ground truth

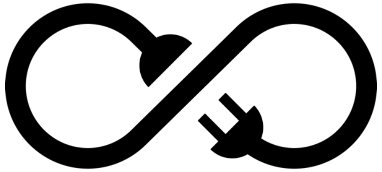
Sensor deployment details



- Stored on device
- Z-wave, ZigBee, or WiFi wireless transmission

Method inspiration: Gaussian Mixture Model

Heart rate



Power consumption

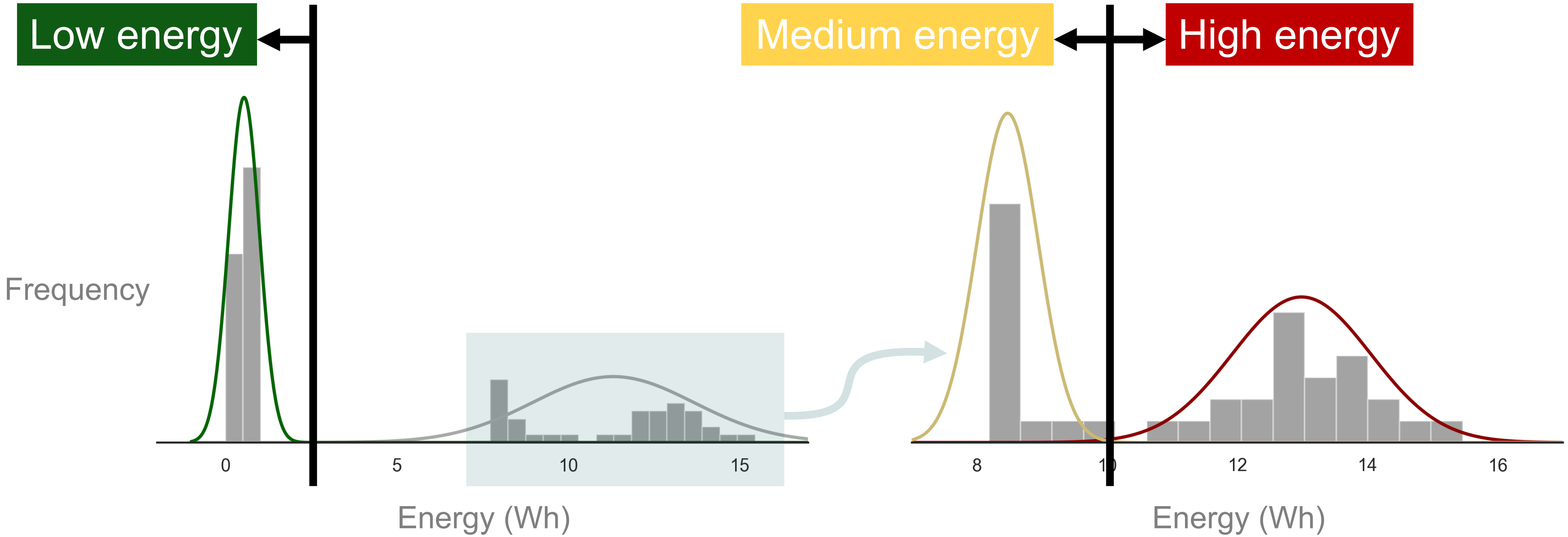
Variational Bayesian Gaussian Mixture Model clustering

Physiological state[1]

Working state

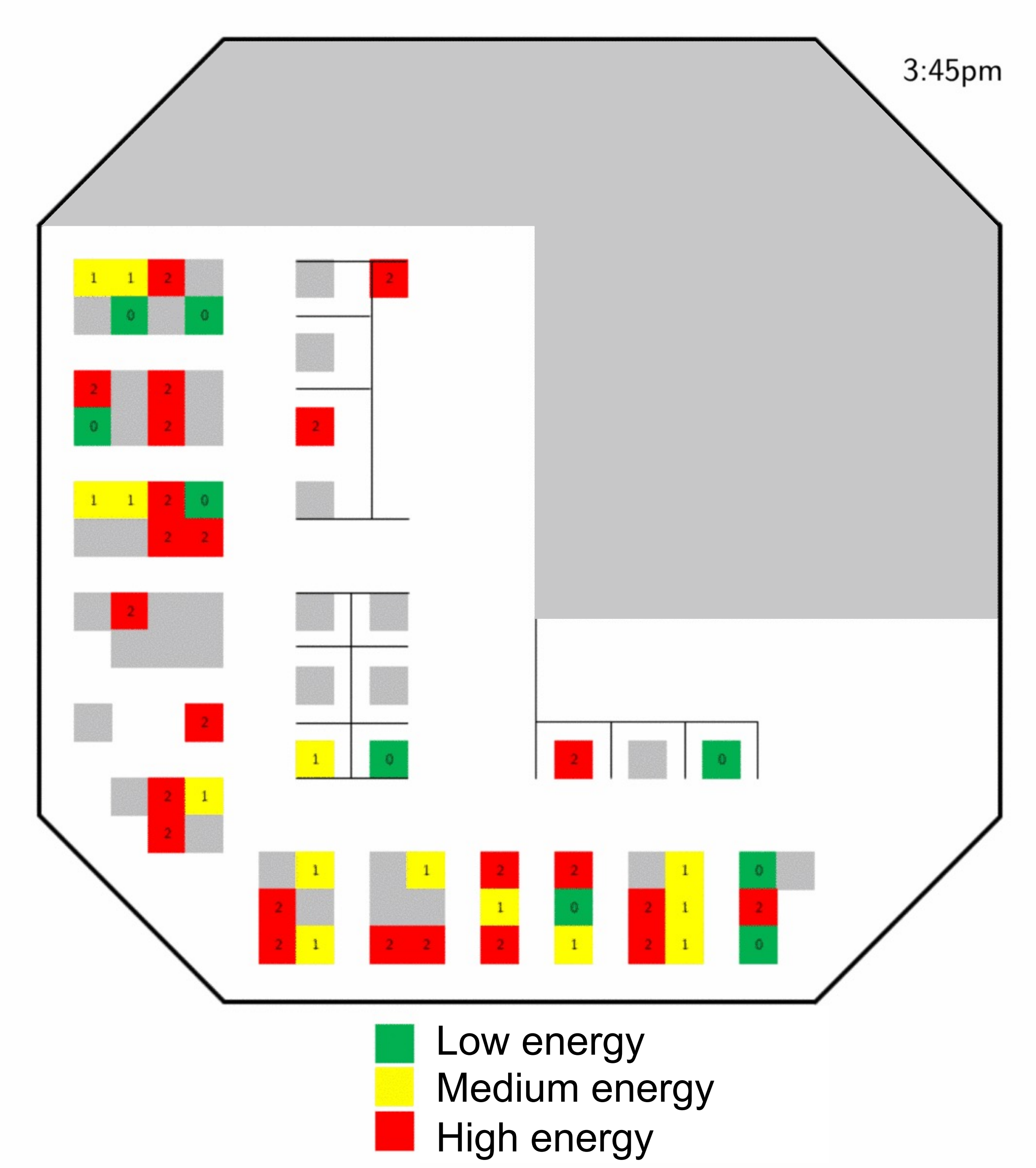
[1] Costa et al. 2012

Variational Bayesian Gaussian Mixture Model

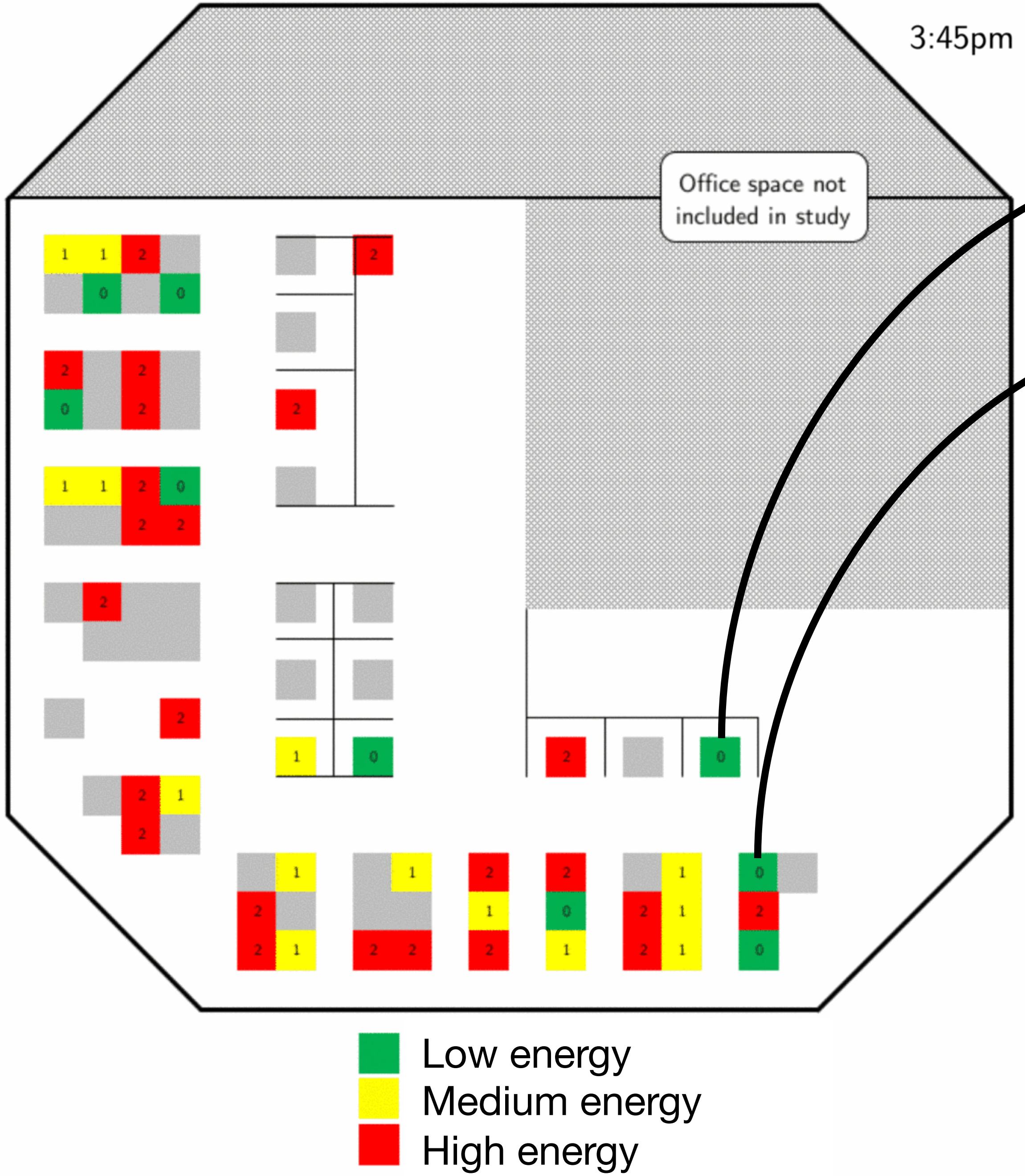


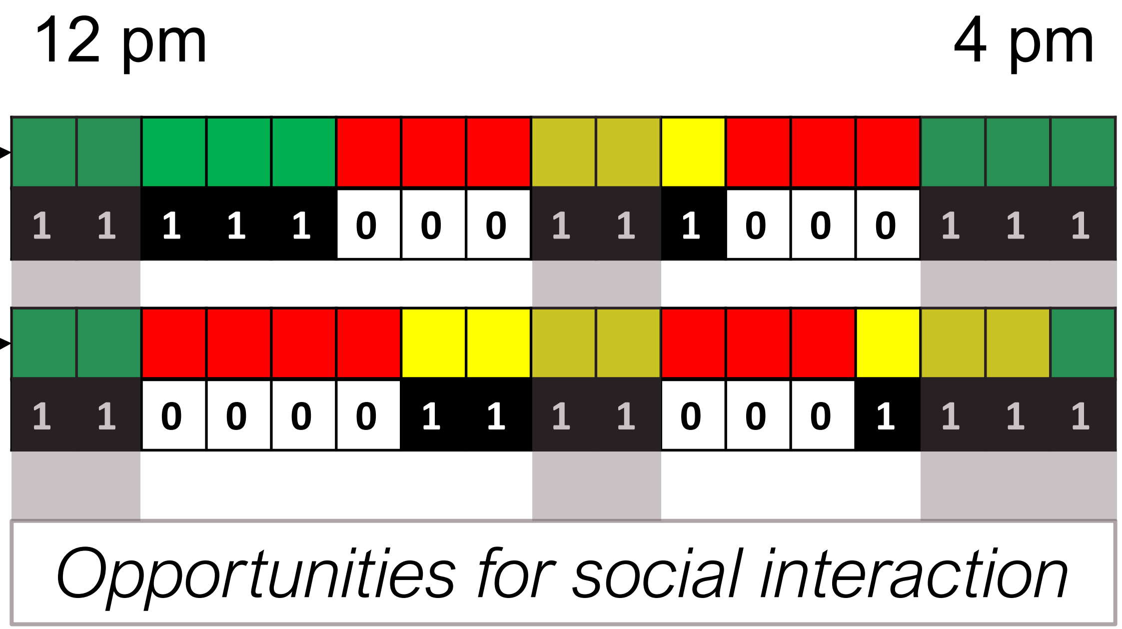
Distribution of 15-min energy consumption
(1 occupant over 1 day)

Activity States Visualized
(Real office in San Francisco)



Sonta, A. J., Simmons, P. E., & Jain, R. K. (2018). Understanding building occupant activities at scale: An integrated knowledge-based and data-driven approach. *Advanced Engineering Informatics*, 37, 1–13.





Fill adjacency matrix with Jaccard similarity:

$$\mathbf{A}_{1,2} = \frac{|\boldsymbol{V}_1 \cap \boldsymbol{V}_2|}{|\boldsymbol{V}_1 \cup \boldsymbol{V}_2|} = \frac{|\boldsymbol{V}_1 \cap \boldsymbol{V}_2|}{|\boldsymbol{V}_1| + |\boldsymbol{V}_2| - |\boldsymbol{V}_1 \cap \boldsymbol{V}_2|}$$

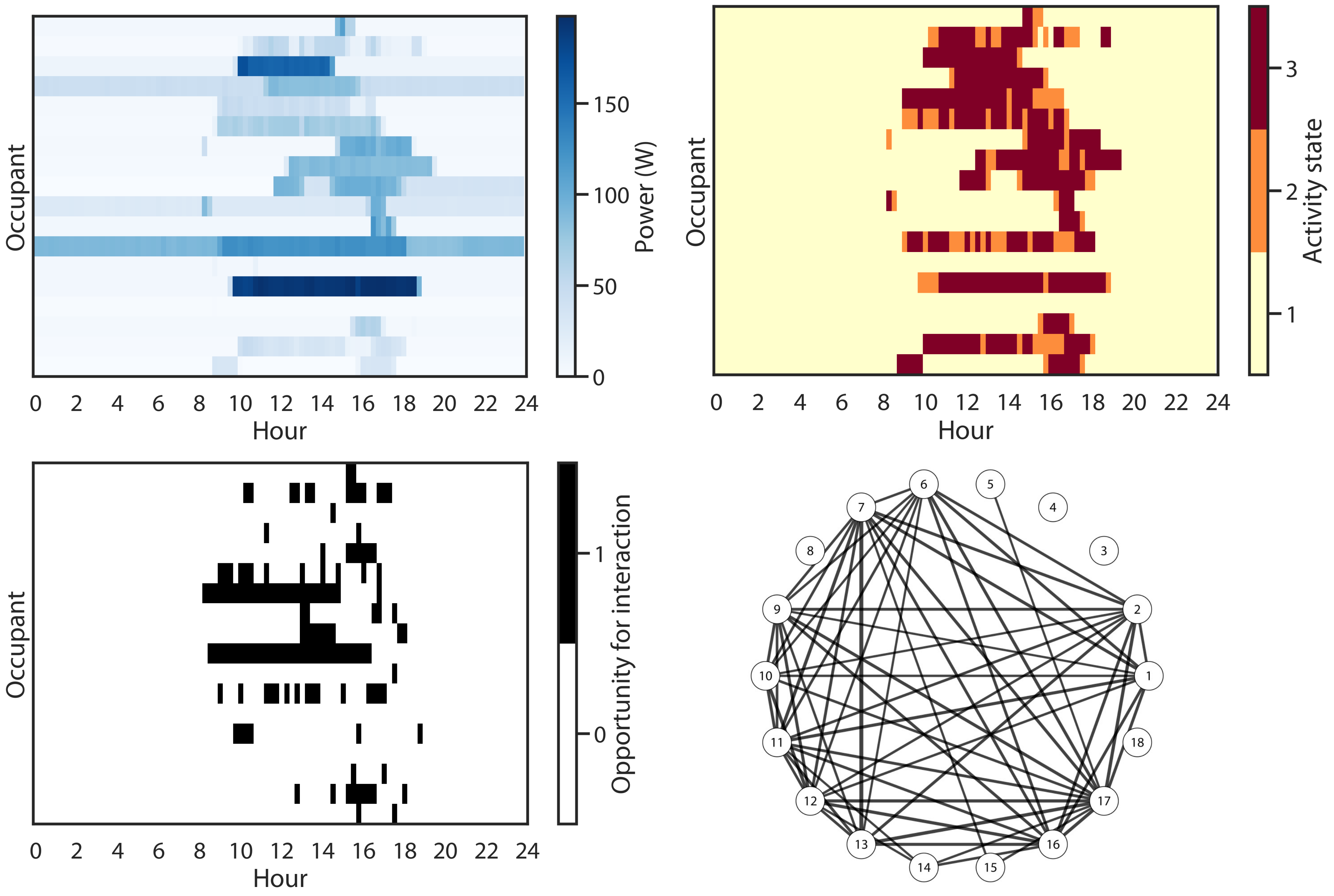
$$|\boldsymbol{V}| = \sum_i V^i$$

$$|\boldsymbol{V}_1 \cap V_2| = \sum_i \boldsymbol{V}_1^i \cdot \boldsymbol{V}_2^i$$

The Interaction Model

Example data from one day

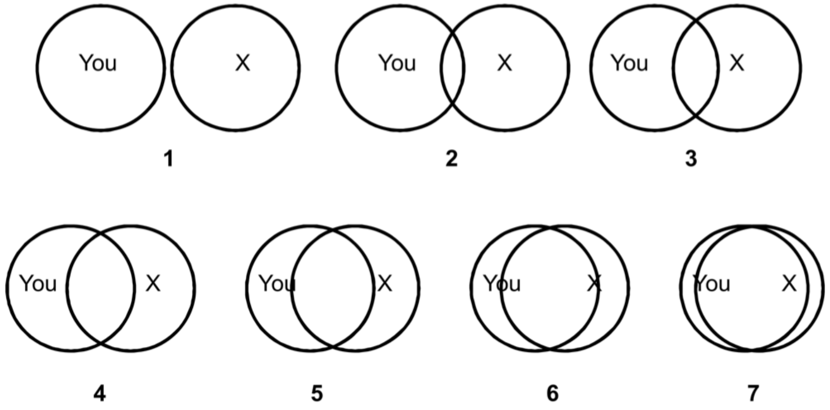
Office building in Berkeley



Survey to measure ground truth network

Social

"Inclusion of the other in the self" scale[1]



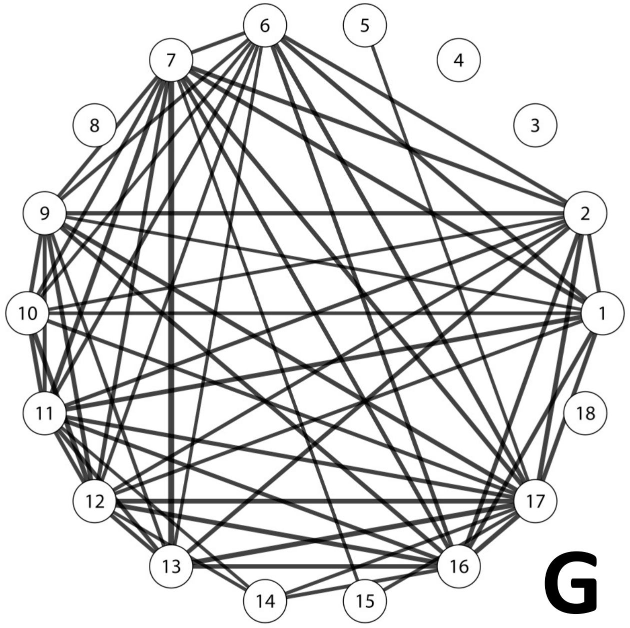
Organizational

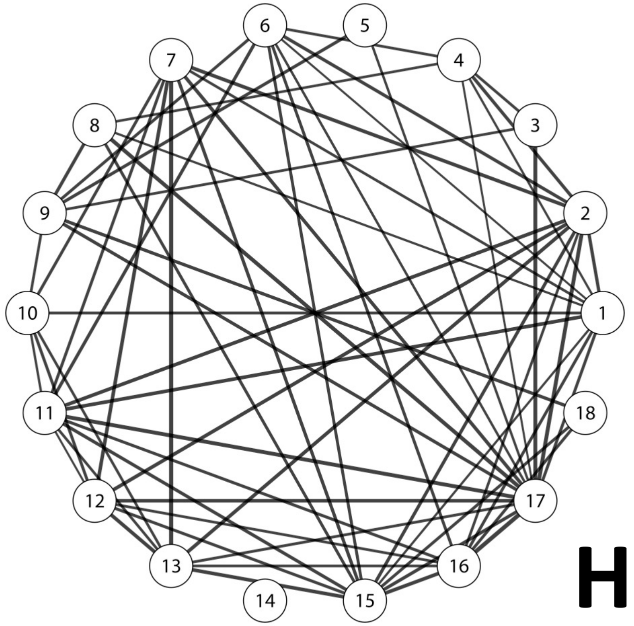
Survey questions measure 3 key attributes[2]

1. Communication (information sharing)
2. Advice (problem solving)
3. Trust (work-related support)

[1] Gachter et al. 2012 | [2] Krackhardt & Hanson 1993

Network comparison





Inferred from model
- Graphical Lasso
- Influence Model
- Interaction Model

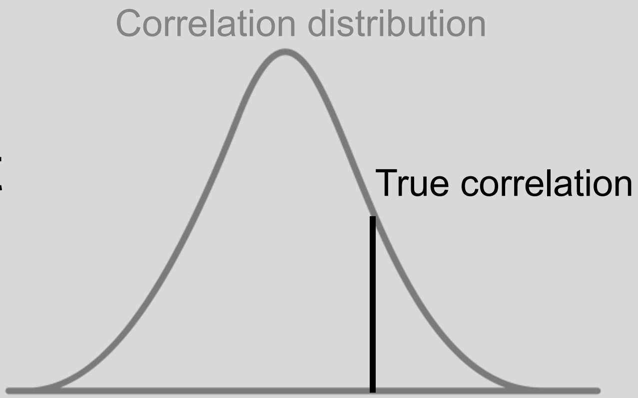
Ground truth from survey
- Organizational (communication, advice, trust)
- Social

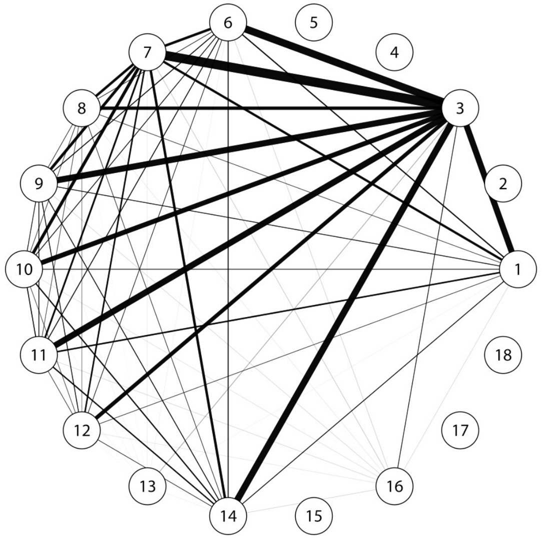
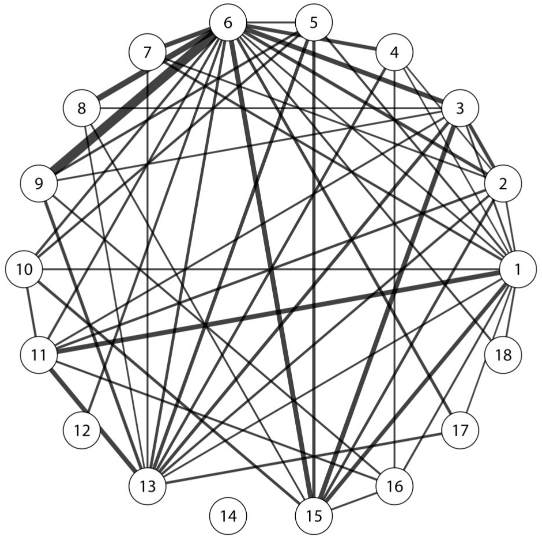
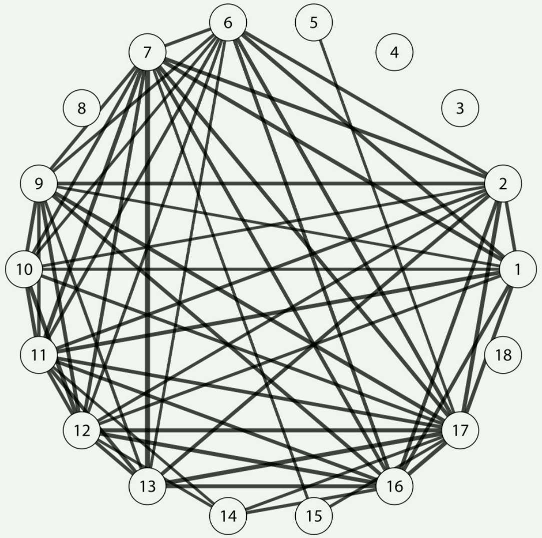
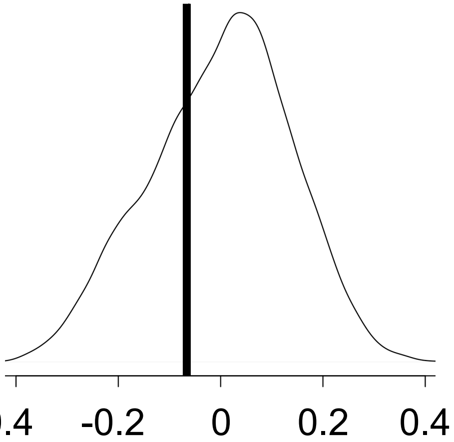
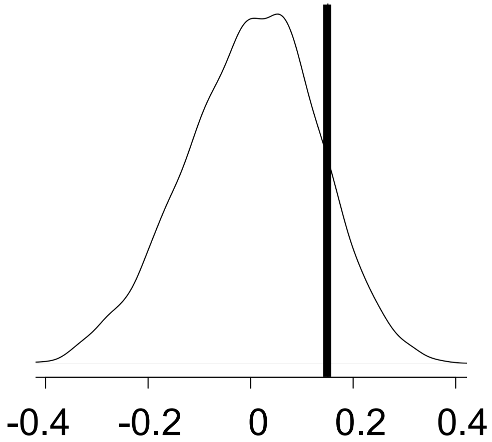
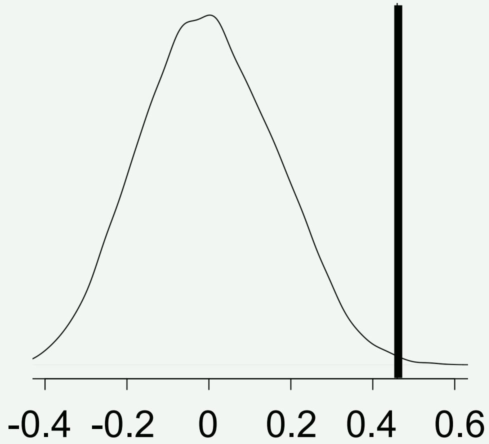
Graph correlation

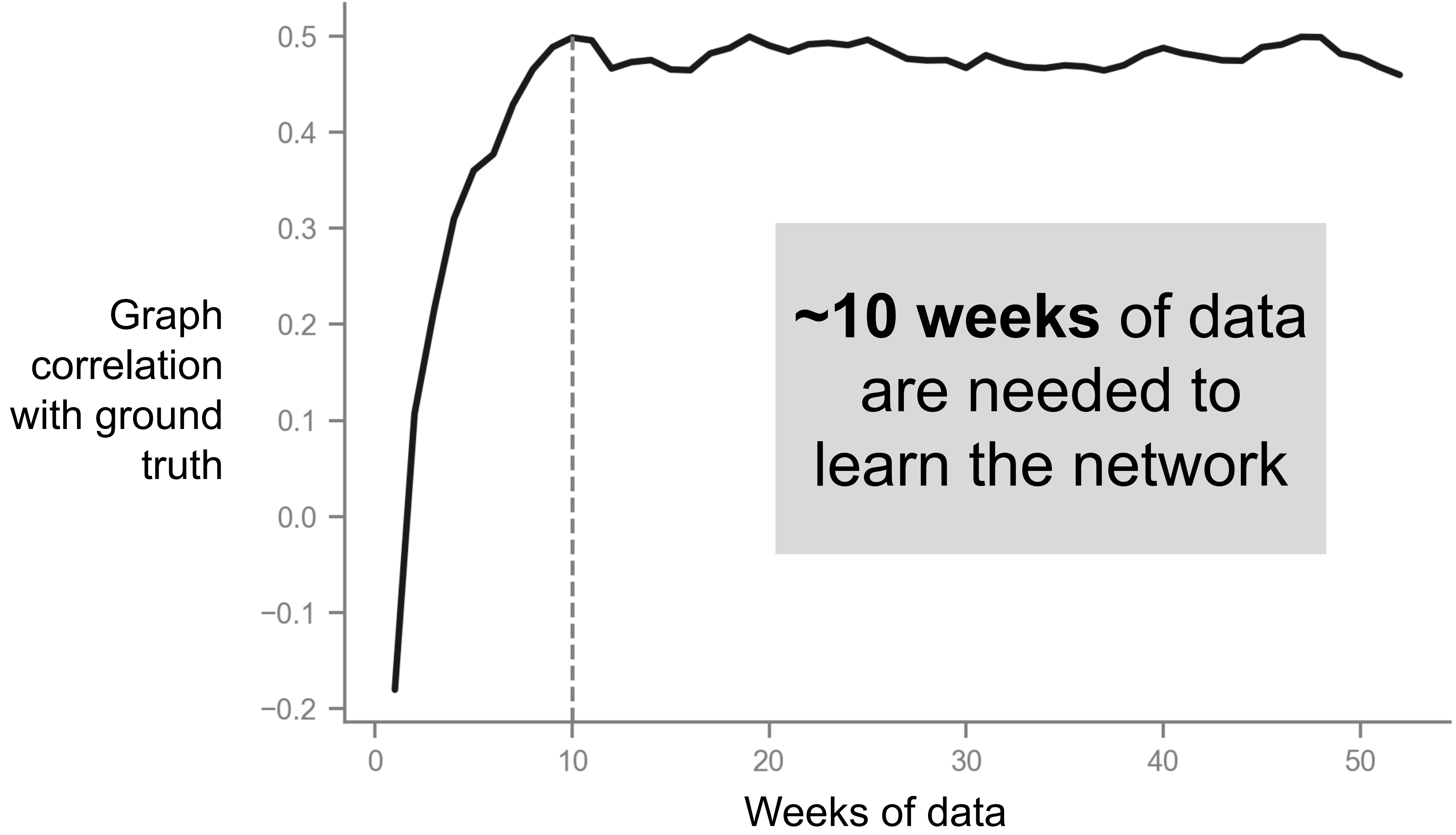
$$\operatorname{cor}(\mathbf{G},\mathbf{H}) = \frac{\operatorname{cov}(\mathbf{G},\mathbf{H})}{\sqrt{\operatorname{cov}(\mathbf{G},\mathbf{G}) \cdot \operatorname{cov}(\mathbf{H},\mathbf{H})}}$$

$$\operatorname{cov}(\mathbf{G},\mathbf{H}) = \frac{1}{|V|^2 - |V|} \sum_{i=j} \left(\mathbf{A}_{ij}^{\mathbf{G}} - \mu_{\mathbf{G}}\right)\left(\mathbf{A}_{ij}^{\mathbf{H}} - \mu_{\mathbf{H}}\right)$$

Quadratic Assignment Procedure (QAP)

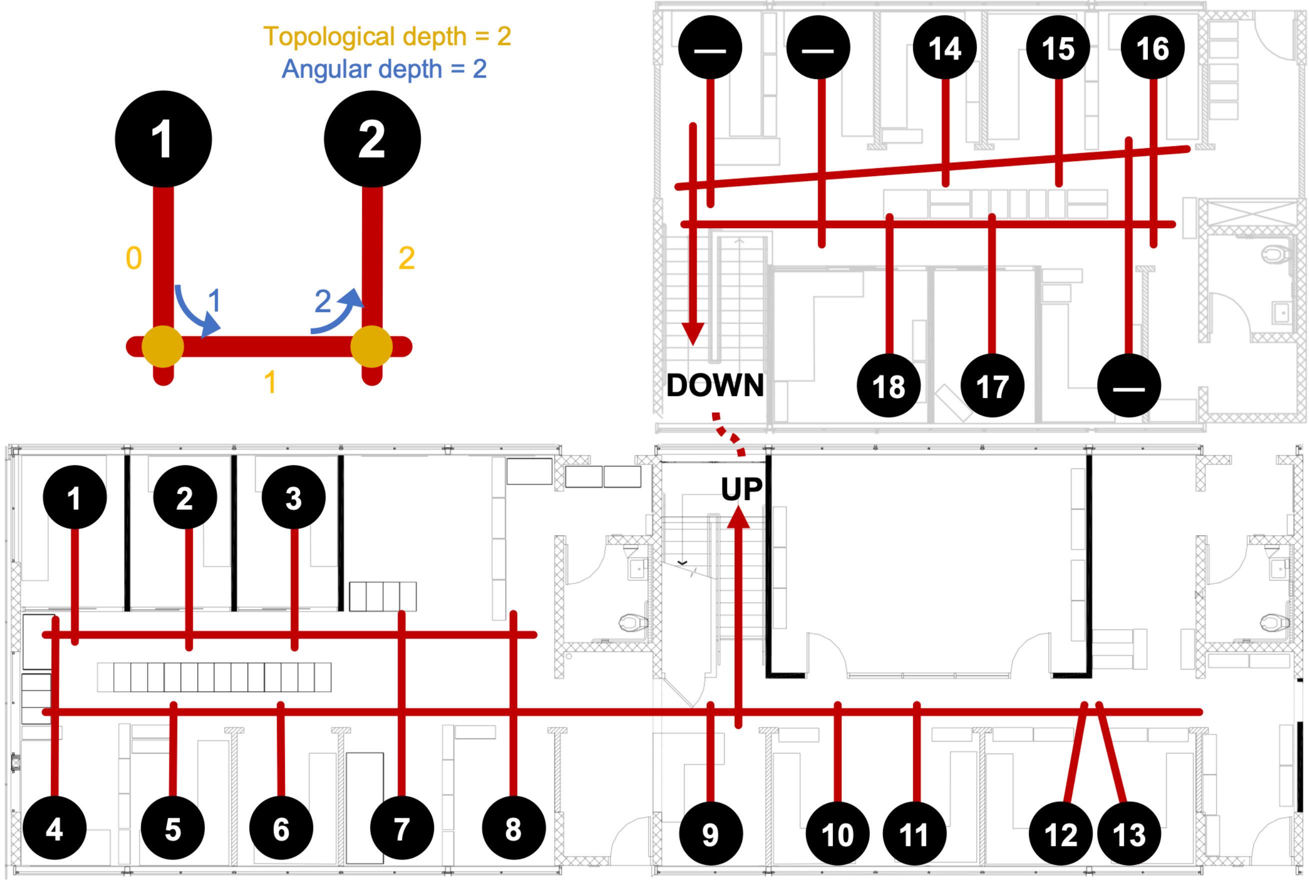


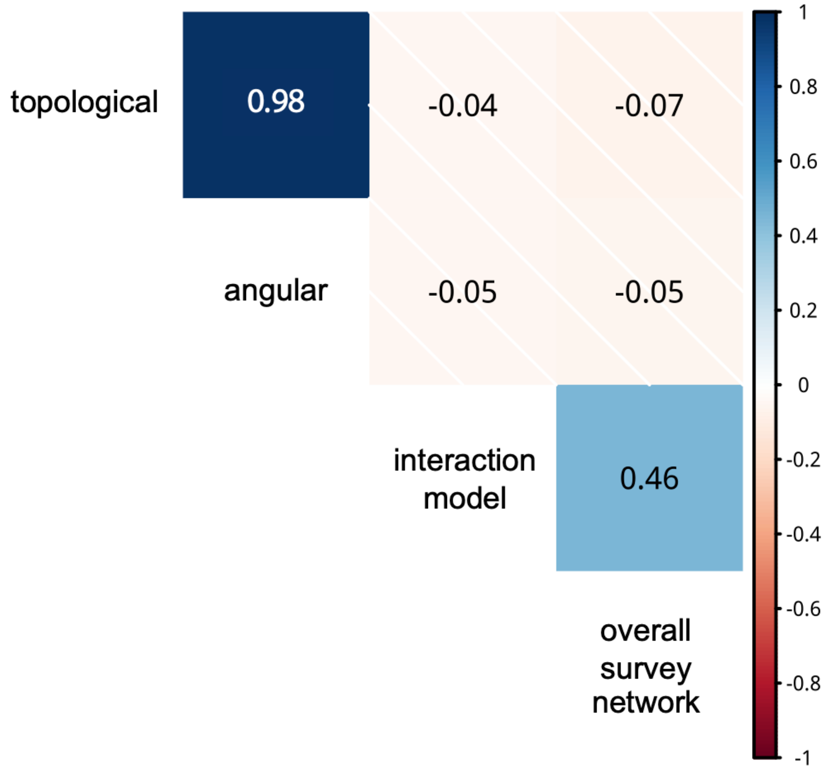
	Graphical Lasso	Influence Model	Interaction Model
Network visualization			
Graph correlation & estimated density from QAP test	**-0.06**, p = 0.66 	**0.15**, p = 0.25 	**0.46**, p = 0.002 



~10 weeks of data are needed to learn the network

How does social network compare with spatial network?





Exam 2 details

- Next class time – Wednesday 8th (9:15 to 11:00 am)
- Closed book and notes
- **You may (and should) bring 1 sheet of paper with notes**
- Content: based on materials in Lectures 6-10
- Format:
  - Multiple choice questions
  - Open-ended questions – application-oriented
    - If network calculations require matrix manipulation (e.g., eigendecomposition), this will be provided
  - No coding

Lecture 06

- Network notation
- Adjacency matrix
- Types of networks
- Degree
- Paths
- Independent paths / connectivity
- Cut sets
- Min-cut/max-flow

Lecture 07

- Components
- Graph Laplacian matrix
- Spectral partitioning; Fiedler method
- Centrality
  - Degree
  - Eigenvalue
  - Katz
  - PageRank
  - Closeness
  - Betweenness

Lecture 08

- Roles
  - Structural equivalence
  - Regular equivalence (intuition)
- Groups of nodes
  - Cliques
  - (k-)Cores
  - (k-)Components
- Clustering coefficient
- Homophily / assortative mixing
- Modularity and community detection
- Motifs

Lecture 09

- Random graph models
  - Erdős–Rényi
  - Configuration model (intuition)
  - Preferential attachment (Barabási–Albert)
- Network structure metrics
  - Degree distribution
  - Clustering coefficient
  - Size of largest connected component
  - Average shortest path length
- Network archetypes
  - Scale free
  - Small world

Lecture 10

- Percolation
- Natural, model, and data defined networks