Multipartite networks, projections, and filtering

What are bipartite graphs?

It’s when we have two disjoint sets of nodes, **U** and **V**

Edges can only be between **U** and **V**, never within

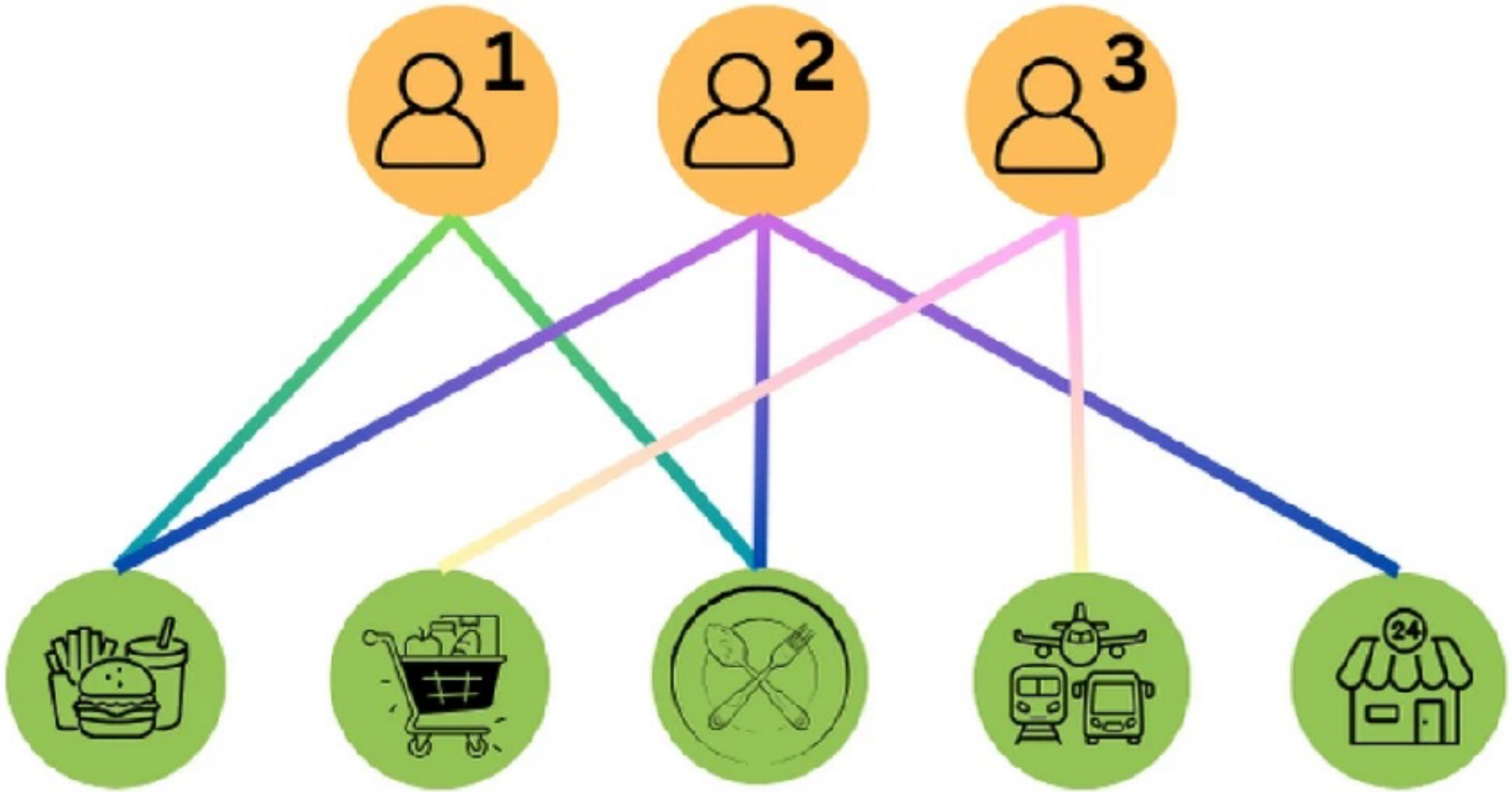
Adjacency matrix is rectangular, not square

Why are bipartite graphs?

Many financial/economic systems naturally connect two (or more*) types of entities

Modelling these networks as bipartite preserves all the information

People

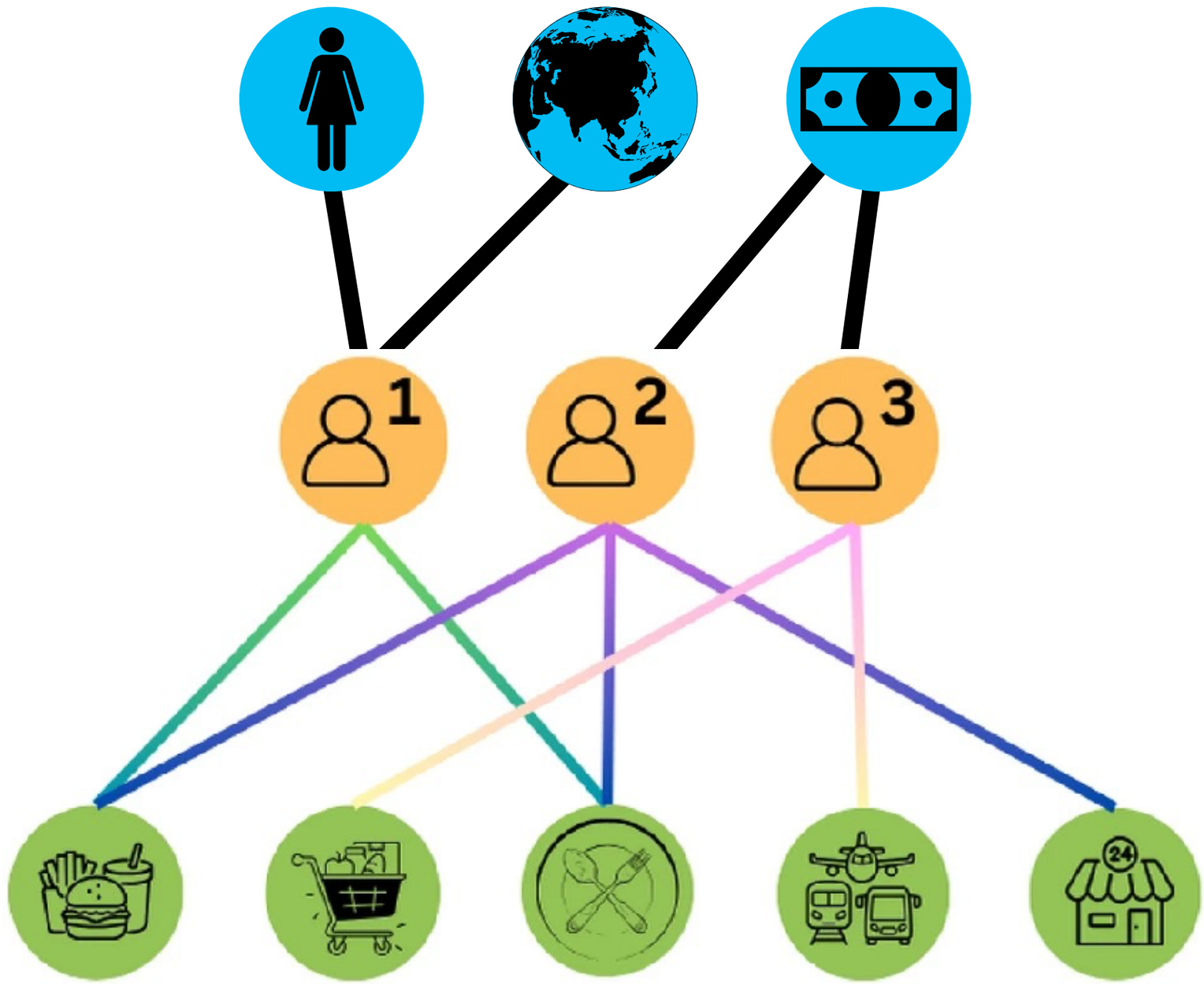


Merchants

What are multipartite graphs?

It’s when we have multiple disjoint sets of nodes, for instance **U, V, W**

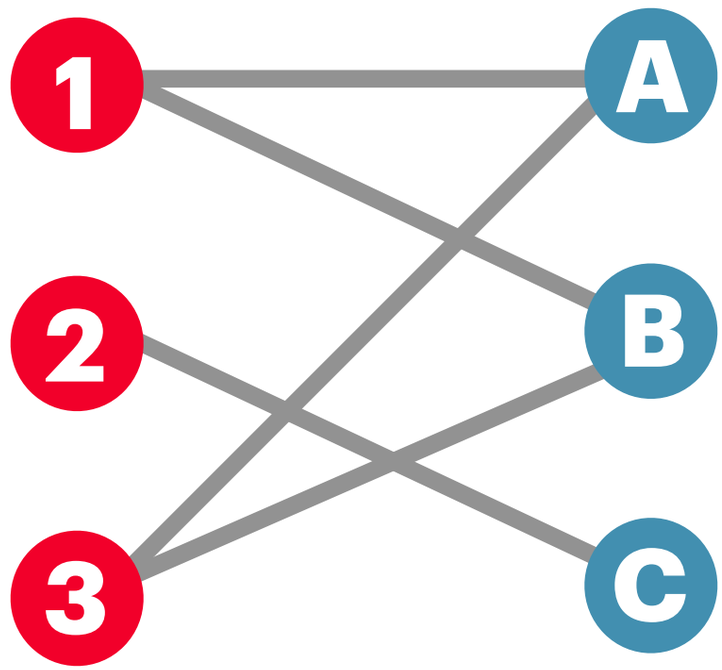
Edges cannot be between sets. Other sets have arbitrary connection rules, based on the system


1
2
3
24

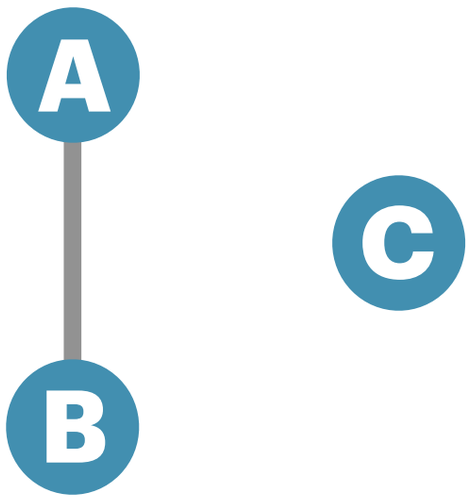
Projections

To use the metrics we learned so far, we need to transform a bipartite network into a “normal” network

This is called projection

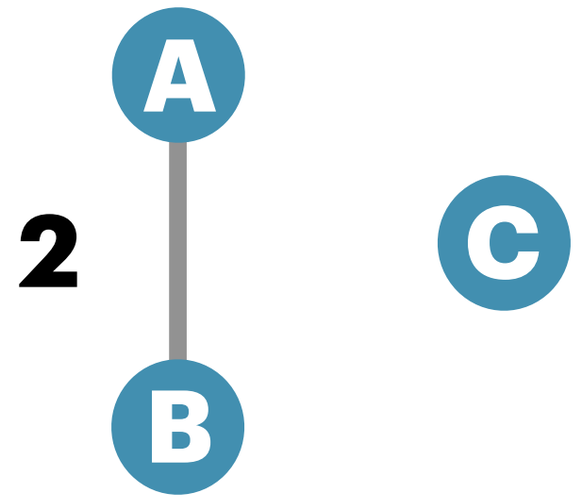

1
A
2
B
3
C

Unweighted projections




A
B
C
1
2
3

Weighted projections



Issues

Nodes with high degrees bias the counts

We can have a high count even if there isn’t much overlap

Similarity measures

We need something that addresses these issues

Standard correlation/similarity measures can be used!

Similarity measures

$$Jaccard(u, v) = \frac{|u \cap v|}{|u \cup v|}$$

Similarity measures

$$Salton(u,v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{\sqrt{|\Gamma(u)|\,|\Gamma(v)|}}$$

Resource diffusion/allocation algorithms

HeatS/ProbS

They favour “important” and “niche” nodes, over high-degree nodes

You can find code here:
https://www.michelecoscia.com/?page_id=734

Exercise

Recommender systems

Think of some concrete examples where you think recommender systems are used

Filtering/backboning

What? Remove edges from the network

When? Networks are too dense

Why? Remove noise

Thresholding

Simplest method

Many issues: favours some nodes - introduces bias

Spanning trees

Great if you need only the essential network

Too few edges otherwise

Advanced methods

Disparity filter

Noise corrected

High salience skeleton

Issues

They all have some free parameters

You need experience

Require further analysis

Issues

They all have some free parameters

You need experience

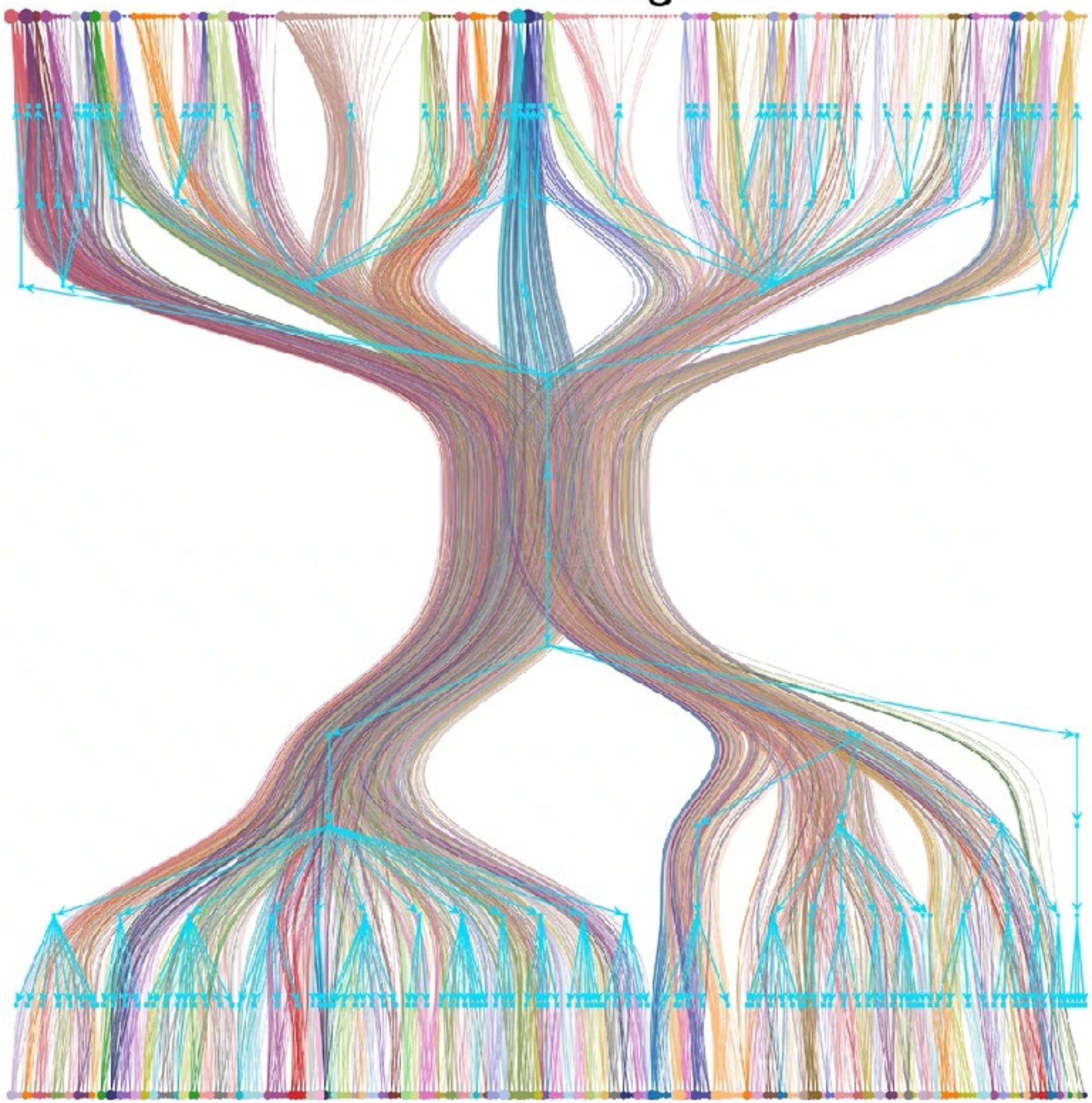
Require further analysis

Is filtering the solution to all of our problems?

Community detection

Bipartite SBM

Merchant Categories



Customers

Examples

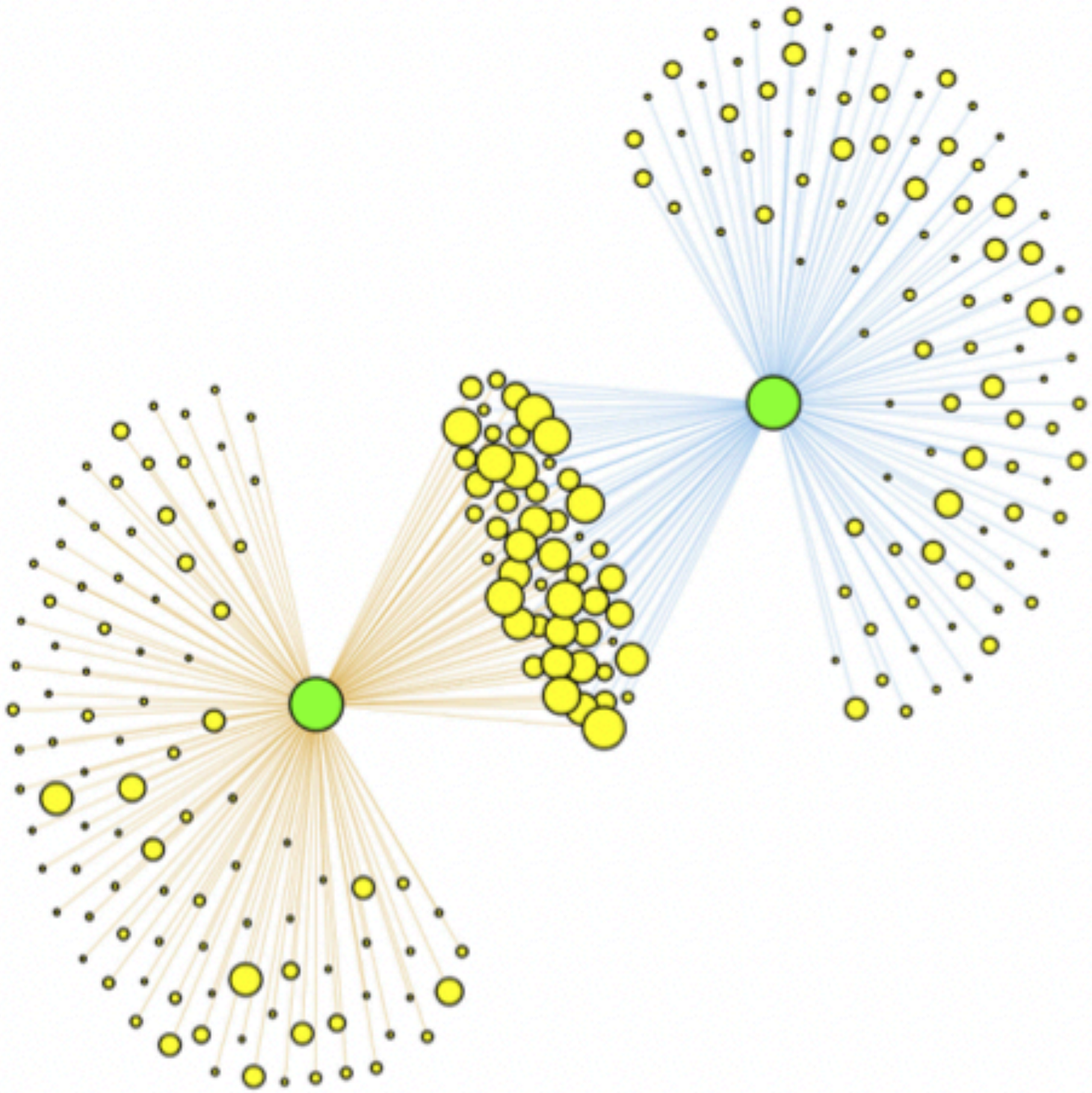
Network of institutions and assets held

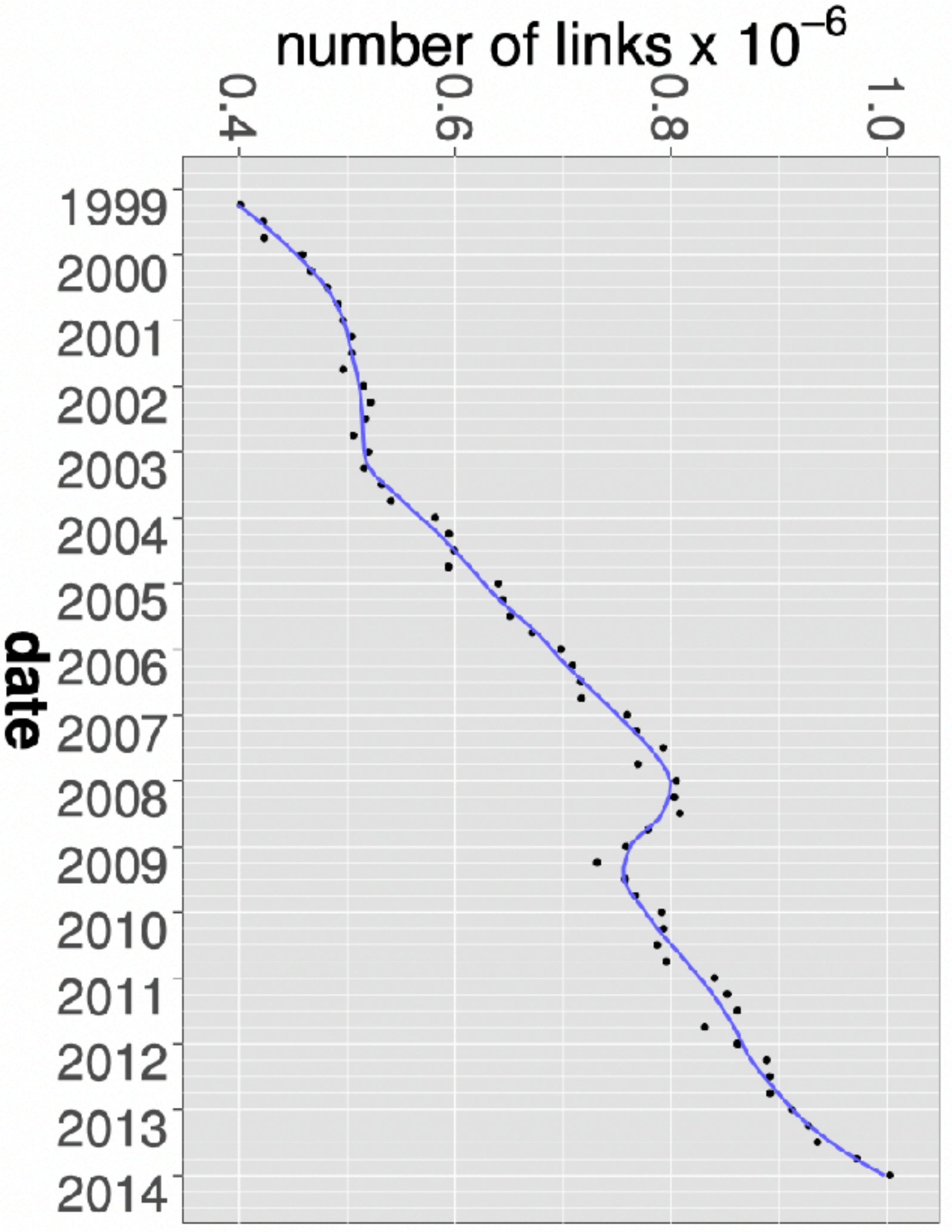
Gualdi et al (2016) Statistically validated network of portfolio overlaps and systemic risk

Institutional **asset holdings**

Period: **1999 through 2013**

Projection (one-mode) of institutions: they draw links between pairs of institutions if they share holdings. Only statistically validated overlaps count (so it's **not just raw overlap**).




number of links x 10^{-6}
0.4
0.6
0.8
1.0
date
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Bank-firm credit network

Marotta et al (2014) Bank-firm credit network in Japan: An analysis of a bipartite network

Bank-firm credit relationships in Japan

Period: **1980–2011**

Links are **loans and mortgages**

They perform community detection without projecting the network

There find **persistence** in communities: certain bank–firm clusters remain stable over years

Changes are gradual rather than abrupt.

There is **persistence** in communities: certain bank–firm clusters remain stable over years

Changes are gradual rather than abrupt.

Sometimes you don’t need to project

Startup networks

Esposito et al (2022) Venture capital investments through the lens of network and functional data analysis

80000+ nodes

Period **2000-2020**

Multiple **countries**

They **project** the network to perform centrality analysis

They find that early investment from **central investors** matters

Bridge investors matter less