Networks: An Introduction

Social Networks (SNs)

Social networks consider social actors not just as isolated individuals but as linked together by relationships.

Economic Networks (ENs)

Economic networks are a subset of social networks.

An economy can be depicted as a network or graph that links economic actors with one another in a flow of exchange. But so what? Why might it matter that an economy (or market or firm) is a network?

- [zuckerman-2003-jel]

Basic Issue: Can network analysis help explain observations of economic behaviour and activity?

Three Concepts of ENs

- concentrated exchange
- primordial affiliation
- structures of mutual orientation.

Some Questions

An EN can be described as “a set of nodes and a pattern of ties among such nodes”

- What is a node?
- How is the boundary of the set of nodes defined?
- What is a tie?
- What is a pattern?
- When is a network an EN, and not just a social network (SN)?

What is a Node?

There are two main types of node in an EN:

- individual human beings
- organisations (human collectivities)

Other Node Interpretations in the Literature

- country
- industry
- innovation
- product

Interpreting Nodes

How do nodes relate to actors: is the link explicit and justified?

- may use aggregation --- a node captures common effects or common causes
- may proxy for unobserved actors

Delimiting the Set of Nodes

The "boundary specification problem":

- are the rules for including or excluding nodes sensible?
- do the rules generate data that are not artefacts of those rules?

Two broad approaches to boundary specification (Laumann et al, 1983):

- "nominalist" approach --- the set comes from a priori criteria
- "realist" approach --- actors (nodes) are included if they are judged relevant by the actors themselves

Example: an analysis of competitors in an industry:

nominalist use an ANZSIC code at a particular fineness (number of digits)

realist look for evidence among the actors on who's in and out
Erickson's "snowball" survey sampling approach

Problems with the Realist approach

Plusses:

- include relevant nodes, and exclude irrelevant nodes,

Minusses:

- but this is likely to be highly contingent to time and place
- and moreover the data collection method might skew the boundaries.
  - who is around to be asked?
  - who will respond to the survey?

Graph Theory: Some Vocabulary

We use graph theory to analyze social networks.

- **node** also called vertex or agent
- **edge** also called tie or link
- **Vertex Set** $V(G) = \{1, \ldots, N\}$
- **Edge Set** $E(G)$ is a set of ordered pairs of vertices

adjacency matrix: $A_{ij} = 1$ iff $(i, j) \in E(G)$

Directed and Undirected Graphs

undirected graph symmetric relationship between nodes (agents) (e.g., friendship)

directed graph (digraph) possibly asymmetric relationship between every two actors for twomode network
buyers and seller, borrowers and lenders, etc

Weighted Graph

add weights to the links

- strength of friendship;
- cost of transaction;
- etc

Simple graph

- no loops
- at most one edge between each pair of nodes.
- unweighted edges

The following are usually synonmous

- “simple undirected graph”.
- “simple graph”.
- “graph” (with no modifiers)

Occasionally we talk about simple directed graphs.

Complete Graphs

A complete graph has an edge between every pair of distinct nodes.
A complete (undirected) graph therefore has $n(n-1)/2$ edges.
If we the vertexes of a regular polygon as our nodes, then the drawing of the complete graph is known as a "mystic rose".

```
layout-circle turtles
ask turtles [create-links-with other turtles]
```

clique the nodes of any complete subgraph (so there is an edge between any two nodes of the clique)

Connected Graphs

walk sequence of adjacent edges $((v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k))$

path a walk with no repeated nodes

cycle a path plus an edge from the final node to the first node

connected graph every two nodes in the network are connected by some path in the network.

component maximally connected subgraph (i.e., you cannot connect its nodes to any additional nodes)

Connected Digraphs

connected directed graph underlying undirected graph is connected (i.e., ignoring the directions of edges).

strongly connected directed graph each node can reach every other node by a directed path

Example: a directed graph that is connected but not strongly connected
adjacency list: (1,2), (1,3), (2,3)

Random Graphs

Simplest random network:

1. create agents (creating the nodes),
2. link each with a random other (creating the edges), until we have the desired number of edges

See: Random Network Example (NetLogo Models Library)

Random Networks (Classic ER Network)

Seminal reference: Erdös and Rényi (1959). Result is often called a “Poisson” network.

Consider each possible link between n agents, and create it with a fixed probability p.

This is sometimes called a $G(n,p)$ network.

The expected degree of a node is then $p(n-1)$. Correspondingly, the expected number of edges is then $p(n-1)n/2$.

Very simple, except for one tricky part: make sure each link is considered only one.

See: Random Network Example (NetLogo Models Library)

Resource: `https://en.wikipedia.org/wiki/Erd%C5%91s%E2%80%93R%C3%A9nyi_model`

Classic ER Network, Ordered Agents

If agents have a natural order, consider all the agents in any order, but only create links between an agent and those who come later in the natural order
E.g., from Random Network Example in the NetLogo Models Library:

```
ask turtles [
  create-links-with turtles with [
    self > myself and random-float 1.0 < p-link
  ]
]
```

Classic ER Network, Using Lists to Impose Order

Traverse a list of the agents, and always consider only links between then agent and agents later in the list

E.g.,

```
let _tlist [self] of turtles
while [1 < length _tlist] [
  let _t first _tlist
  set _tlist butfirst _tlist
  ask turtle-set _tlist with [random-float 1 < p-link] [
    create-link-with _t
  ]
]
```

Code sprint: Campus Friendships

50 students arrive on a dorm floor.

For any pair, and equal probability (say, 0.2) of a friendship link.

Code Sprint: Campus Infection

Infect a random student.

Each “friend” then becomes infected, and so on.

How far does the infection spread?

Characterizing Networks

We can characterize networks in many ways. Here we focus on the following:

- degree distribution
- diameter
- path length distribution
- clustering coefficient distribution

Vertex Degree

$k(v) =$ number of vertices directly connected to vertex v

A digraph vertex may have in-degree different from out-degree

In a simple graph, if there are n vertexes, each vertex has a maximum degree of $n - 1$.

In a complete graph, every vertex has this degree.

Mean Degree

mean vertex degree: sum all the vertex degrees, and divide by number of vertices

Equivalently, divide twice the number of edges by the number of nodes.

$$< k >= \frac{\sum_{v \in V} k(v)}{\# V} = \frac{2 \# E}{\# V}$$

Q: What is the mean degree in the Campus Infection model?

Degree Distribution

$P(k) = n_k/n$ the fraction of notes with degree $= k$

In a complete graph, $P(k) = 0$ for every k except for $k = n - 1$

Degree as Centrality

A node's degree is an indication of its "centrality" in a network.

In a social network, a highly connected node may have:

- more influence
- more access to information
- more risk of infection

Degree Distribution of ER Random Graph

$P(k) = C(n-1,k)p^k(1-p)^{n-1-k}$
This is just a binomial distribution.
If we hold *np* constant while letting $n \to \infty$, we get a limiting Poisson distribution.

Path-Distance

Define $d(v1, v2)$ to be the shortest path length from $v1$ to $v2$.
Problem: what if there is not such path?
Divergent solutions:

- diameter is ∞
- report some arbitrary value
- raise error

NetLogo Example (`nw` extension):

```
ask turtle 0 [show nw:distance-to turtle 1]
```

(Returns `false` if there is no such path.

Path-Distance Distribution for an Undirected Connected Graph

Mean path length:

$$< d >= \sum_{i<j} d(i,j)/C(\#V,2)$$

Diameter

diameter of a connected graph the longest distance between any two vertexes.

Problem: what if the graph is not connected?

Divergent solutions:

- diameter is ∞
- diameter is diameter of largest component

Mean Path Length

The mean number of edges for the shortest paths for all possible pairs of network nodes.

Average path length: add up all the $d(v_i, v_j)$ for $i > j$, and divide by the number of paths $n(n-1)/2$.

Problem: what if the graph is not connected?

Divergent solutions are used.

- raise an error (e.g., networkx)
- return some arbitrary value (e.g., NetLogo's `nw` extension returns `false`)
- set $d(v_i, v_j) = 0$ if there is no path from v_i to v_j.
- use the average path length of the largest component

Implication: small mean path length implies quick spread (of disease, information, innovation, etc.)

Clustering Coefficients

- overall clustering coefficient
- individual clustering coefficients
- average clustering coefficient

Individual Clustering Coeffient

This is for undirected graphs.

The individual clustering coeffient for node i is

- the number of triangles including node i
- divided by the number of triples centered at node i

Finding the Number of Triples

```
to-report n-triples
  let _n (count my-links)
  report (count my-links - 1) * count my-links / 2
end
```

Finding the Number of Triangles

```
to-report n-triangles
  let _ct 0
  let _nbrs link-neighbors
  ask _nbrs [
    ask link-neighbors with [member? self _nbrs] [
      set _ct (_ct + 1)
    ]
  ]
  report (_ct / 2)
end
```

Finding the Individual Clustering Coefficient

```
to-report clustering-coefficient
  report n-triangles / n-triples
end
```

Average Clustering Coefficient

```
to-report average-clustering-coefficient
  let _n (count turtles)
  let _sum sum [clustering-coefficient] of turtles
  report (_sum / _n)
end
```

Global (Overall) Clustering Coefficient

This is for undirected graphs.

Question: How likely is it that an agents friends are themselves friends.

triplet a central node, with 2 branch nodes

open triplet a triplet where the branch nodes are not connected

closed triplet a triplet where the branch nodes are connected

triangle a set of 3 nodes that are completely connected (so one triangle implies 3 closed triplets); a 3-vertex clique

global clustering coefficent: #(closed triplets)/#(triplets) = 3 × #(triangles)/#(triplets)

Network Topologies

There are three common network topologies:

- random
- scale-free
- small-world

Scale-Free Networks

Idea: any subnetwork shares key properties of the global network.
Classic references: Barabási & Albert (1999) Barabási (2002)
Result: a few nodes are very highly linked (“hub-and-spoke” networks).

Preferential Attachment

Probability of a new arrival linking to an agent depends on the current degree of the agent.

E.g., start with one linked pair of nodes, and keep adding nodes as follows until you have as many as you want.

- pick a link at random,
- pick one of the nodes of this link,
- link this node to a newly created node

```
let _link one-of links
let _partner one-of [both-ends] of _link
crt 1 [create-link-with _partner]
```

Two further boundary issues:

one mode or two: single type of actor, or two types (buyer **and** seller)

death and birth e.g., firms are born, firms merge, firms exit

OneMode or TwoMode Networks?

SN analysis is most often concerned with "onemode" networks,
That is, for instance, every actor (node) might both potentially `send` and `receive` a tie of interest to every other actor, with no a priori agent-typing.
In contrast, ENs are usually "twomode" networks: two types of nodes ex ante (either `sender` or `receiver` nodes, but no nodes that do both)
E.g., directors and boards they sit on.
So markets are *interfaces* between, say, buyers and sellers.

- the EN modeller must generally delimit two sets of nodes: *buyers* and *sellers*

But: sometimes no clearcut distinction in an EN.
E.g. traders on eBay can both buy and sell (onemode), but most specialise.
Nonetheless, the pattern of specialisation might be significant.

The Meaning of the Absence of a Tie?

Q: How to interpret the absence of a tie (of a specific kind) between two nodes, when absence

- either indicates impossibility? or
- or indicates actors' choice not to tie? e.g. sexual contacts?

Moreover, what do interfirm relations mean?

- another form of network tie?

using reliance on networks to predict the location of the firm's boundaries, or

- something different from market integration? qualitative increase in commitment; need a theory of the firm to analyse ENs.

Ties? Patterns? EN v. SN?

. What manner of orientation among the nodes is meaningful such that it consitutes a structure that has causal implications for outcomes of interest?
A network is an EN if it has effects on future events that are considered economic.
An EN is of interest if it cannot be fully reduced to the constrained choices made by actors, that is, a complete account requires attention to the EN and its patterns.
"manner of orientation" among nodes

- tie definition: infinite possibilities of what the ties model;
- tie pattern: central to the analysis of ENs.

Pattern is Central to Analysis of ENs

Two meanings of “network”:

- the pattern of ties among a set of nodes, or
- a high degree of pattern in the ties among nodes, with a specific theoretical or empirical meaning.

e.g. network vs market

e.g. network vs organisation

Economic Network (EN) any collection of actors ($N > 2$) that pursue repeated, enduring exchange relationships with one another and, at the same time, lack a legitimate organisational authority to arbitrate and resolve disputes that might arise during the exchange. (selforganisation)

Four Approaches to ENs

1. Ties as market exchanges. The EN is more patterned or concentrated than expected from market models.
2. ENs are economic interactions shaped by ascribed or primordial relationships Do network structures (or commitments) have causal impact? But “primordial” blurs the distinction between ENs and SNs.
3. ENs as structures of mutual orientation (I and II are subsets of III)
4. Increasingly, the structures of interfirm orientation designed to be orthogonal to market exchange. To meet needs unmet by market exchange (e.g. coopetition)

Do ENs have Causal Implications?

Not just a pretty picture?

Does the structure of the EN have causal implications for the actors of interest?

Two issues (Reagans et al. 2003):

1. unobserved heterogeneity, and
2. reverse causality

To argue that particular network position confers advantage, it's necessary to show, first, that any observed association between position and success does not reflect underlying differences in actor *type* or, second, that expectations of success did not determine the observed network pattern.

Some Network Metrics (from Tesfatsion/Zhou)

Clustering Coefficient measures how well connected my neighbouring vertices

C(v) = Actual Total possible

where Actual = number of connections among my neighbouring vertices.

Distance L ij = shortest path length between vertices i and j

Characteristic Path Length of Graph G = L(G) = the average of L ij for all i , j in graph G, *math:‘i neq j*.

Properties of Graph Models

Path length L(G) Clustering C(v) Regular Longest Largest SWN Short Large Random Shorter Small ScaleFree Shortest Small Small World Networks (SWN) are resiliant against random failures of vertices (nodes), but highly vulnerable to deliberate attacks on hubs (vertices of high degree k (v)).

Limitations (from Girvan)

We need to think more about the interplay between network topology and dynamics.

Can we find "universality classes" with respect to topology and dynamics?

Can we determine which topological features are most important to different types of dynamics?

1. Are networks a fad?
2. What are we doing wrong in the field of complex networks?
3. Where do we go from here?

Scientific “Fads” from statistical physics

Their claims have often been overasserted:

- Selforganized criticality SOC---a supposed explanation for why we see power laws in so many natural systems when power laws in physics are only seen at critical points.
- Econophysics --- the application of statistical physics toward the understanding of market patterns.
- Fractals --- selfsimilar patterns observed in a variety of natural systems --- snowflakes, river networks, forest fires?
- Spin glasses (or spin systems generally) applied to neural networks, gene regulation, economy, opinion formation, war, . . .
- The edge of chaos, EOC
- And networks?

Are Networks a Fad?

Yes, but . . .

Networks can still be useful, if . . .

- you have knowledge of the system that you are studying.
- you have a problem that naturally calls for a networkbased approach.

Complex networks should not be the answer in search of the problem.

Al Wilhite remarks ...

There seem to be two types of economic network studies.

- theoretical *games on networks* and *network evolution* studies that use highly abstract networks embedded with some sort of decision making criterion.
- empirical papers that focus on a single network using reams of data to map out that actual network's structure.

But the Wilhite & Fong (2009) lies somewhere in between.
The theory part of the paper used abstract networks to simplify potential relationships which give suggestions as to what might happen. These relationships are then tested using data from actual organizations, but we do not map out the 400 networks of these organizations---that task is impossible.

References

[zuckerman-2003-jel] Zuckerman, Ezra W. 2003. On Networks and Markets by Rauch and Casella, eds.. *Journal of Economic Literature* 41, 545-565.

[1] E. W. Zuckerman, On Networks and Markets by Rauch and Castella, eds., Journal of Economic Literature, 41: 545-565, June 2003.
[2] Leigh Tesfatsion/Zhou,
http://www.econ.iastate.edu/classes/econ308/tesfatsion/NetworkNotes.ModifiedZhou.pdf
[3] Michelle Girvan, lecture 4 in The structure and dynamics of complex networks,
http://www.itp.ac.cn/csss05pic/download/Girvan/Girvan%204.ppt
[4] Al Wilhite and Eric A. Fong, Agentbased models and hypothesis testing: an example of innovation and organizational networks.
http://www.agsm.edu.au/bobm/KER/FinalDrafts/AWKER.pdf
[5] Barabási, A.-L., and R. Albert, 2002, Rev. Mod. Phys. 74, 47.