Economics of Information Networks

Stephen Turnbull

Division of Policy and Planning Sciences

Lecture 4: December 10, 2020

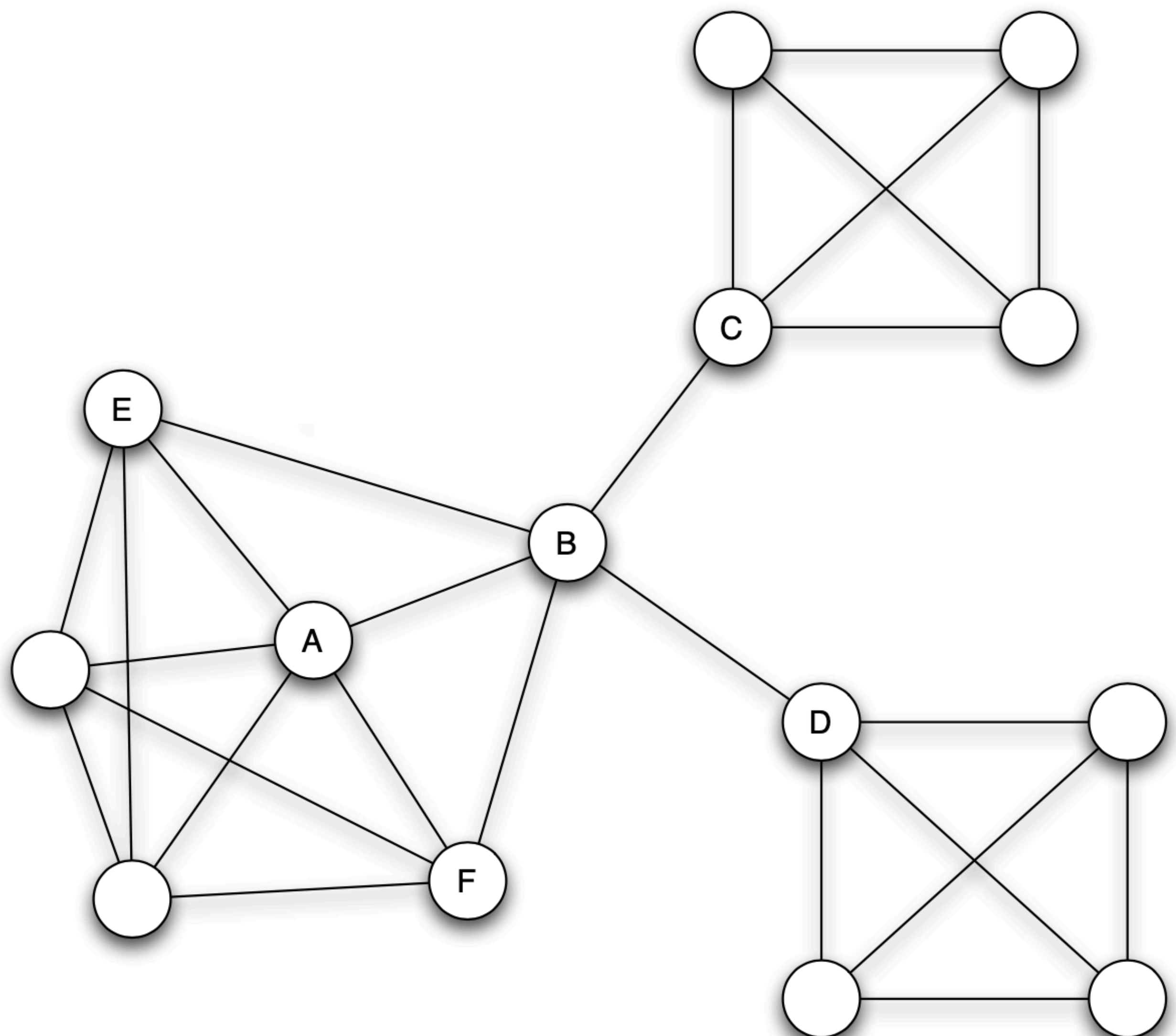
Abstract

We continue discussion of the "modern" economics of networks, which considers the effect of structure of networks on economic (including social) behavior.

Structure and behavior

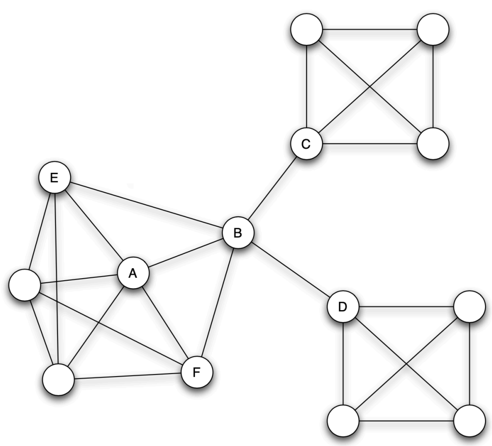
- Network structure induces constraints on behavior.
  - At a granular scale, in a city you can go from anywhere to anywhere by automobile.
  - By contrast you can only go from station to station on a train; you can't even get off between stations as you pass by.
- In this sense, economics (which studies how people deal with constraints) can be considered to "include" social behavior.

Symmetry and asymmetry of nodes



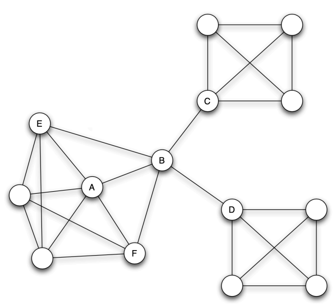
Symmetry and asymmetry of nodes

- Note the symmetry between E and F, and between the two unnamed nodes in that group.
  - E and F are connected to all other group nodes except each other.
  - The unnamed nodes are connected to all other group nodes except B.
- E and an unnamed node are different in a minor way (E is not connected to its twin F, while the unnamed nodes have a common unconnected group member), and in a more important way: E-C is *length 2*, while from an unnamed node to C is *length 3*.



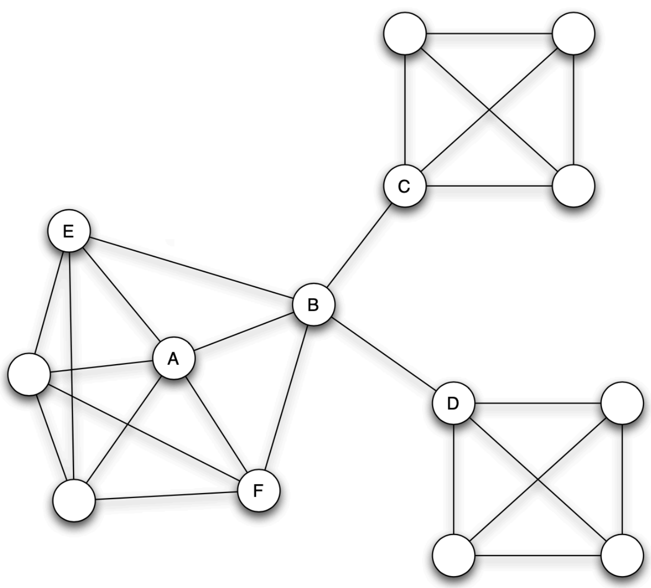
Symmetry and strength of gatekeepers

- One obvious feature of the network is the gatekeeper position of B is stronger than C and D, though they are also gatekeepers.
  - You might think that it matters whether there are *more* nodes in one group than the other, or more *valuable* nodes in one group than the other.
  - But as usual in networks, we focus on the value of the *link* between B and C. Since they are equally able to *deny access* to the link, they are equally strong as gatekeepers for that link.
- B's greater strength is due to its gatekeeping for two subnetworks, while C and D control only one each.
- In fact, B is also a gatekeeper for C's access to D.



Centrality and neighborhood overlap

- Another is the centrality of A *in* its group.
- An important aspect of social (including economic) networks is *centrality*. This applies to both nodes and edges. Simple binary notions are those of nodes that are *pivots* and *gatekeepers*, and edges that are (local) *bridges*.
- Earlier we discussed *neighborhood overlap*, which is a quantitative generalization of bridge. Connected nodes with low neighborhood overlap may function like gatekeepers for each other.



Centrality and embeddedness

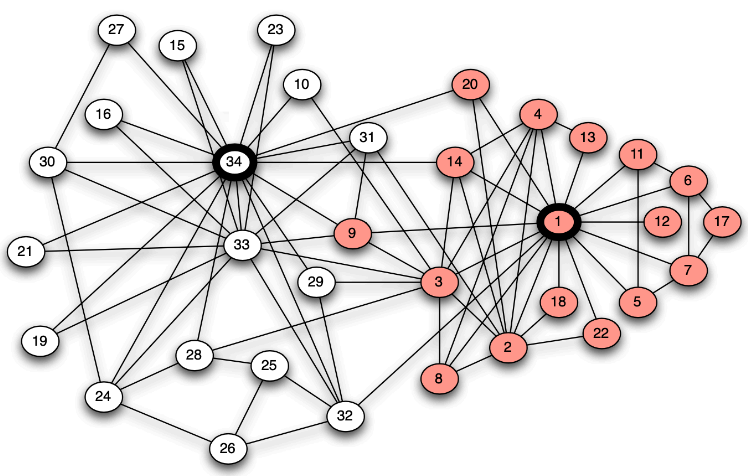
- We define an absolute version of *neighborhood overlap*, called *embeddedness*. The *embeddedness of an edge* is the number of common neighbors of the edge's endpoints, $|N(A) \cap N(B)|$.
  - An edge with embeddedness zero is a local bridge.
- Kleinberg and Easley write:

  > [W]hat stands out about A is the way in which all of his edges have significant embeddedness. A long line of research in sociology has argued that if two individuals are connected by an embedded edge, then this makes it easier for them to trust one another, and to have confidence in the integrity of the transactions (social, economic, or otherwise) that take place between them [117, 118, 193, 194, 395].

- Note that neighborhood overlap is interpreted as *power* while *embeddedness* is interpreted as *trust*.

Graph partitioning

- Links are defined by friendship (each end declares the other to be a friend).
- The two nodes with heavy borders have special roles. Node 1 is the instructor (who is authorized to test candidates and bestow rank), and Node 34 is the club president.
- The club split into two clubs while under study The node colors indicate who joined which club.
- Could this split have been predicted from the friendship structure?



Betweenness, flow, and partitioning

- Betweenness can be defined in terms of an abstract flow. Imagine that one unit of fluid flows between each pair of nodes, say A and B. This flow is equally divided among all shortest paths between A and B.
- The *betweenness* of an edge E is computed as follows:
  1. Find all shortest paths between all pairs of nodes.
  2. For each pair of nodes, determine the amount of flow through E. (A pair with no path through E will contribute 0.)
  3. Sum flows through E over all node pairs to get *betweenness of* E.

Calculating betweenness

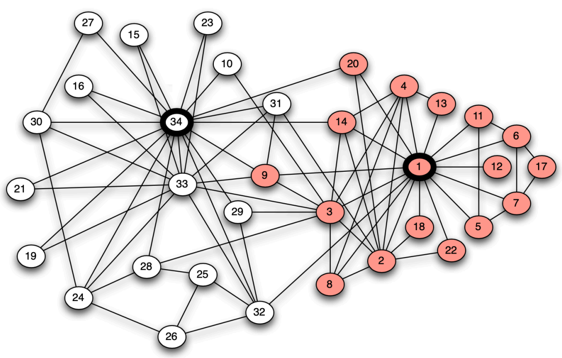
- This is a typical example of how algorithms often build on other algorithms.
- To find *all* the shortest paths from A to *all* the other nodes, perform the breadth-first search algorithm, labeling each edge with its distance from A when it is traversed.
- When each node is reached, check if this is a shortest path. If not, ignore that edge. If so, make a copy of the node reached, and attach the edge to it.
- This new graph is a *tree*, so it is easy to list all the paths. Each node in this tree corresponds to a path, and the path connects A to the node. Group by node labels.

Girvan-Newman partitioning

1. Calculate betweenness for all edges.
2. Remove all edges with highest betweenness.
3. If the graph becomes disconnected, this is the n-th step partition.
4. If there are any edges left, go to 1.
5. Stop.

Test of Girvan-Newman algorithm

- The actual division of the clubs members is given in the image.
- Girvan-Newman predicts almost the same division, except that as suggested by its position in the image, Node #9 is predicted to join the component led by #34, but instead joined the component of #1.
- With other information, we can guess why the algorithm failed. #1 is teacher qualified to evaluate a member in a test for *shodan*, *i.e.*, the "first-degree black belt." At the time of the split, #9 was preparing for the test in a few weeks. #34's group had no qualified teacher.



Algorithms *vs.* intuition

- Don't be mistaken: the "right way" is algorithms *and* intuition.
- A skilled consultant could have *predicted* the outcome. This is the bread-and-butter of the Shako-trained consultant: use a scientific method (computer, data, and algorithm in this case) to get a "good" answer, then apply deep knowledge and human intuition to get a better one.
- It's not obvious, but besides checking whether an *incorrect* prediction is "explained" by other factors, we should be careful to see if *correct* predictions are "explained" by other factors (often called "missing variables").
- For example, suppose we discovered that all of the students in the #34 faction live in one dorm, and all of the students in the #1 faction live in another that is far away. Then it might be that the location of the dorms determines *both* the friendship patterns and the clubs' memberships.