Social Network Analytics, Empirical Exercise #2
Due on Monday, October 29, at 12:00pm

Group structures, clustering, and core-periphery analysis in venture capital co-investment networks

Loading venture capital data

For this exercise, we will analyze groups and clusters in the networks of venture capital firms. Venture capital firms often co-invest with each other on startup venture as part of syndicate investment teams. The ties formed by co-investing together influence the strategy and performance of the venture capital firms and the entrepreneurs they fund, and certain positions in this network are more beneficial than others.

Information about venture capital firms and their investments is contained in several files:

Files generating network data

The two files "Funding_events_7.14.csv" and "Funding_events_7.14_page2.csv" contain information on venture capital investment events, from June 1981 until July 2014.

- In these files, each row represents an investment round made by one or more venture capital firms into a startup company

- The venture capital firms are listed in the column "Investors"

- The startup company is listed in the column "Portfolio Company Name"

- Consider a relationship tie to exist between venture capital firms when the firms invest together in the same round of a portfolio company—the firms show up in the same row together

- Consider firms as tied together if they invest together at least once—ignore multiple instances of a relationship

- Allow relationships to persist over time, so that the network in July 2014 is comprised of all cumulative ties that have occurred up to this point

Files generating venture capital performance data

The file "Venture_capital_firm_outcomes.csv" contains information about the performance of venture capital firms in each year they are active investors.

- The number of successful investments a venture capital firm has in any year is listed in the column "successful_investments"—successful investments represent acquisitions, IPOs, and other events in which the startup generates cash for the venture capital firm

- A venture capital firm goes out of business when it cannot raise a new fund from a limited partner to make new investments—firms that go out of business have a 1 in their most recent observation in the column "out_of_business", and otherwise this column is filled by 0s

Exercises

1. First, perform the Kevin Bacon Hollywood Actor exercise on the venture capital firm network.

   (A) Which firm is the center of the venture capital firm network as of July 2014? Consider the most central firm to be the firm with the largest closeness centrality, as in the Hollywood Actor example.

   (B) Next, compute the average shortest path length between all firms in the July 2014 network and verify that the firm with the highest closeness centrality also has the lowest average path distance. You can consider nodes that are unreachable to be separated by a number of steps equal to the total number of the firms in the network.

(C) What is the average shortest path length for all firms? Why is this number so high?

2. Next, we will look at the development of the network over time. Allow the network to be updated monthly for each month t in the data, adding the new ties that occur through investments in the current month to be added to the existing network of ties that have occurred in previous months.

   In Class Session 3, a figure plotted the average "coreness" of the venture capital network over time. Coreness represents the highest-degree k-core to which a firm belongs. When a node is a member of a k-core with a high degree, its surrounding ties are very dense. When many nodes are members of k-cores with high degrees, this suggests that there may exist dense clusters within the network.

   (A) Construct a figure similar to Class Session 3's, plotting the average "coreness" of firms in the network over time. This can be computed using the igraph function `coreness`. On the x-axis should be the age of the network in months. On the y-axis should be the highest-degree k-core each venture capital firm belongs to, averaged over all firms in the network up to that month.

   (B) Construct a plot similar to (A), but now allow ties to "decay." Remove ties from the network if they are not renewed within 10 years. Does the figure appear different than before? What does this suggest about the nature of relationships in the co-investment network?

3. Continuing to look at the development of the venture capital over time, do you think that the recent network now exhibits more of a core-periphery structure or a structure made up of distinctly clustered components? Provide at least two pieces of descriptive evidence for your conclusion.

4. Last, we will analyze whether being in the core of the network helps venture capital firms' and the entrepreneurs they work with to perform better. You may use whichever statistical approach you wish to determine the direction and strength of the relationship between network position and a venture capital firm's performance.

   (A) Is a venture capital firm being at the center of the network related to having more more successful investments in a given year, in terms of its closeness centrality? Is this true for other kinds of centrality as well?

   (B) Is a venture capital firm being at the center of the network related to being less likely to go out of business, in terms of its closeness centrality? Is this true for other kinds of centrality as well?