Mathematical Challenge December 2018

Network centralities for financial services

References

[1] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32:245–251, 07 2010.

[2] Thiago Silva and Liang Zhao. *Machine Learning in Complex Networks*. 01 2016.

[3] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

Description

Introduction

With network analysis we are able to investigate the structure of networks. Network analysis is often applied in the context of social networks, but can also be applied for use cases in the financial sector. The nodes in the network can represent companies, and links can be business relations between these entities. Examples of such business relations are 'individuals at these two companies have held meetings' or 'the money flow between these companies'. The question at hand defines the type of nodes and the type of links between these nodes.
Networks of companies in different forms can provide useful information that may not arise from focusing on individual entities only. For example, the network of banks and the money flow between banks can provide insight in the stability and resilience of the financial sector as a whole. In this challenge we explore a certain kind of network analysis: centralities, and investigate how to apply centralities in the financial sector.

Centralities

Centrality measures help to identify the 'most important' nodes in a network. Several definitions of importances of nodes have been developed over time. Degree centrality is by far the most know centrality and is defined as the number of edges a node has a connection to, [3] defines degree centrality as follows

$$C_D(n_i) = \sum_j x_{ij}$$

where n_i is node i and x_{ij} is the presence or absence of an edge between nodes n_i and n_j. Closeness centrality is the average length of the shortest paths between the node and all other nodes in the network, closeness centrality is defined as

$$C_C(n_i) = \left[\sum_j d(n_i, nj)\right]^{-1}$$

where $d(n_i, nj)$ is the distance between nodes n_i and n_j. Betweenness centrality is the number of times a node acts as a bridge on the shortest path between other nodes, betweenness centrality is defined as

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

where n_i denotes node i, and g_{jk} the number of shortest paths linking nodes j and k and $g_{jk}(n_i)$ the number of connected shortest paths that contain node i.

Centralities in Weighted Networks

The above mentioned centrality measures are all defined on unweighted networks, that means the weight of the edges are defined as one. In weighted networks, the edges have a value, an example of a weight can be an amount of contact between two nodes. Centralities can be defined with a weighted component, [1] shows $C_D^w(n_i) = \sum_j w_{ij}$ is often used as a measure for weighted degree. However, they argue that this only takes into consideration a node's total level of involvement in the network, and not the number of nodes it's connected to. Therefore, they propose the following measure

$$C_D^{w\alpha}(n_i) = C_D(n_i) \times \left(\frac{C_D^w(n_i)}{C_D(n_i)}\right)^\alpha$$

where α is a positive tuning parameter. If α is set between 0 and 1, higher degrees are in favor, if it's set above 1 a low degree is favorable. Centralities closeness and betweenness can be more easily generalized to a weighted version as they take into account path lengths.

Communicability and Vitality

Communicability is a measure for how easily two nodes can communicate. That is, how many options of paths exist between two nodes. [2] defines communicability as

$$C_{com}(n_i, n_j) = \frac{1}{s!}P_{ij} + \sum_{k>s} \frac{1}{k!}\left(A^k\right)_{ij}$$

where P_{ij} are the shortest paths between n_i and n_j and s is the length of these shortest paths. The term A^k_{ij} is the element (i, j) of the k-th power of adjacency matrix A, this matrix gives the number of paths of length k between n_i and n_j. For each length of path we count the number of paths and divide it by its factorial length, and sum over all different path lengths. If for a path of length 4, 2 alternatives exist, the measure will be increased with $\frac{2}{4!} = \frac{1}{12}$. With more path alternatives, the measure increases while on the other hand longer distances get penalized non-linearly due to the factorial term. Ultimately, a higher measure of communicability means more ways of communication between two nodes.
Flow betweenness vitality [2] is defined by

$$C_{BV}(n_i) = \sum_{\substack{n_j, n_k \in V \\ n_i \neq n_j, n_i \neq n_k}} \frac{f_{n_j,n_k}(n_i)}{f_{n_j,n_k}}$$

where $f_{n_j,n_k}(n_i)$ is the flow that must go through n_i and is determined as follows $f_{n_j,n_k} = f_{n_j,n_k} - \tilde{f}_{n_j,n_k}$ with $\tilde{f}_{n_j,n_k}$ the maximal flow between n_j and n_k when removing n_i from the network. The user can define an appropriate function f, for example the communicability function.

Feedback Centrality

Feedback centralities are centralities that are built on the concept of importance of one node influencing the importance of neighboring nodes. Eigenvector centrality is an example of such feedback centrality. The centrality of a node n_i is proportional to the sum of the centralities of the neighboring nodes, [3] defines eigenvector centrality as

$$\lambda x_i = \sum_{j \in V : j \to i} x_i = \sum_{j \in V} A_{ji} x_j = (A^T x)_i$$

That means that x_i is the i-th component of the transposed eigenvector of adjacency matrix A and eigenvalue λ, which is how it obtained its name Eigenvector centrality.

Questions

- **Q1:** *Describe the difference between betweenness centrality and flow betweenness vitality.*
- **Q2:** *Describe an application in financial sector, and explain which centrality measure would be appropriate and why.*
- **Q3:** *Describe the pros and cons of the different centrality measures for your application.*

We look forward to your opinions and insights.

Best Quant Regards,

swissQuant Group Leadership Team