

**Beyond Social Networks:  
Advanced Uses of Gephi in Humanities Research  
Session 2 Handout**

**Brian Tsz Ho Wong**  
Ph.D. Candidate, Asian Studies  
CDCS Training Fellow  
University of Edinburgh

Today we will use Gephi to reconstruct and analyse the ownership networks in the contemporary Hong Kong stock market. On 17 May 2017, David Webb, a retired investment banker and activist investor, published an article on his website. The article, entitled ‘The Enigma Network: 50 stocks not to own’ (<https://webb-site.com/articles/enigma.asp>). Webb argued that investors should not invest in 50 HK-listed companies, which he said had cross-owned each other’s shares. This allowed the ‘big bosses’ behind the ‘Enigma Network’ to easily manipulate the prices of these stocks, and ultimately damage investors’ interests. Webb’s article had shocked the market, and the Securities and Futures Commission of Hong Kong began to investigate Webb’s allegations. As a result, several executives of the companies in Webb’s networks were arrested and prosecuted, although most were eventually acquitted.

**Activity 1: How were the networks built?**

Webb’s website provided comprehensive data on the ‘Enigma Network’. Many of the 50 companies in the networks are holding companies, meaning that they own almost all the shares in their subsidiaries, some of which were also listed companies.

There is an elegant visualisation of the ‘Enigma Network’ in Webb’s article. Webb probably drew it in Microsoft Word or PowerPoint or similar software. In other words, he drew it by hand. Based on the data on Webb’s website, discuss with your group whether the information on the 50 companies and their subsidiaries is sufficient for us to reconstruct the ‘Enigma Network’ in Gephi? After the discussion, we will test it in Gephi.

### **1.1. Is it the ‘Enigma Network’?**

1. Go to the BeyondSocialNetworks repository on github.
2. Download The Enigma Network.xlsx and open Gephi. Do not open the spreadsheets in Excel, if you have opened them, close them before opening Gephi.
3. Click on ‘New Project’ and then go to ‘Data Laboratory’.
4. Click on ‘Import Spreadsheet’, then select The Enigma Network.xlsx and press ‘open’.
5. Import the ‘Nodes table’, then go to ‘Import spreadsheet’ and import the ‘Edges table’.
6. Select ‘Integer’ for the ‘Weight’ column.
7. Select ‘Directed’ for the ‘Chart type’ and then select ‘Append to existing workplace’.
8. Go to ‘Overview’ and select ‘Yifan Hu’ for the layout.

Discuss the results with your group and try to identify what information we need to complete the network.

### **Activity 2: Visualising the ‘Enigma Network’ in Gephi**

#### **2.1. Import the second dataset into Gephi**

1. Go to the BeyondSocialNetworks repository on github.
2. Download The Enigma Network2.xlsx and open Gephi. Do not open the spreadsheets in Excel, if you have opened them, close them before opening Gephi.
3. Click on ‘New Project’ and then go to ‘Data Laboratory’.
4. Click on ‘Import Spreadsheet’, then select The Enigma Network2.xlsx and press ‘open’.
5. Import the ‘Nodes table’, then go to ‘Import spreadsheet’ and import the ‘Edges table’.
6. Select ‘BigDecimal’ for the ‘Weight’ column.
7. Select ‘Directed’ for the ‘Chart type’ and then select ‘Append to existing workplace’.
8. Go to ‘Overview’ and select ‘Yifan Hu’ for the layout.
9. Change the ‘Optimal Distance’ to 300.

## 2.2. Measuring the networks

1. Go to 'Statistics'.
2. Run 'Average degree', 'Avg. Weighted Degree', 'Network Diameter', 'Modularity', and 'Eigenvector Centrality'.
3. Go to 'Data Laboratory', and you will find statistics for several types of centrality measures, including degree (centrality), weighted degree, betweenness centrality, and eigenvector centrality.

## 2.3. Interpreting the networks

1. **Degree centrality:** number of edges linked to each node (popularity, well connected)
2. **Weighted degree centrality:** Number of edges connected to each node, taking into account the weight of the edges (popularity, well connected)
3. **Betweenness centrality:** the number of times a node is on the shortest path between other nodes (bridges, gatekeepers)
4. **Eigenvector centrality:** similar to degree centrality, but takes into account the centrality of the nodes it's connected to

How would you interpret the results? Discuss with your group.

## Activity 3: Composing financial datasets for Gephi

Discuss with your group:

- The differences between The Enigma Network2.xlsx and the spreadsheets you used for Gephi before.
- What do we need to include in the spreadsheets for financial networks?
- What functions can we use in Excel to make the spreadsheet for Gephi more efficient?