# Complex Network

## Analysis of Financial Networks

## Spreading Phenomena and Portfolio Selection

*Authors: Khoa Pham Dinh, Duy Vu*

## Introduction:

In this report, we try to construct financial networks from the daily price movement of 900 securities in S&P500 and S&P400. The period of consideration is from 23 March 2020 to 22 April 2023.

## Constructed Networks:

**Node:** Stocks

**Link:**

First, we calculate the logarithm price change of every stock over date

Using an unweighted Pearson correlation of the log price change between 2 shares and link the most 1.1% uncorrelated/ correlated stocks. We found that the correlations of stocks follow a normal distribution.

We created the network ourselves by crawling data from yahoo finance. Attribute of a node: its daily price

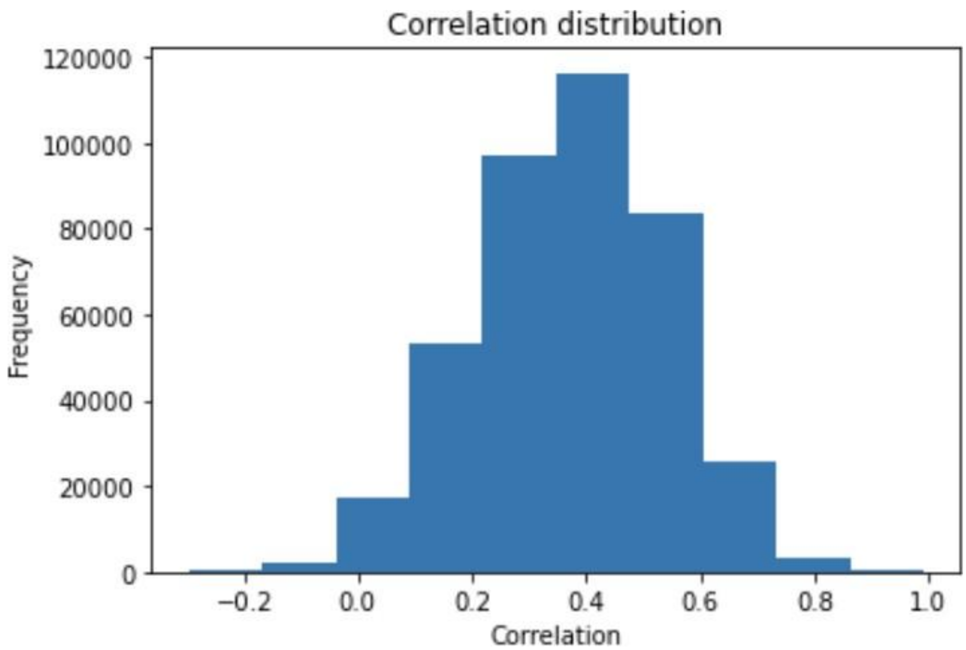


*Figure 1:* Distribution of correlation

From that, we construct 2 networks:

- Network 1: choose 1.1% most correlated pairs of nodes and link each of them
- Network 2: choose 1.1% least correlated(uncorrelated) pairs of nodes and connect each of them

## Networks statistics:

**<u>Network 1 statistics:</u>**

Nodes involve in the connected component of network 1: 524
Networks nodes: 894
Network links: 4391

Average degree: 9.82

Density of network 1 and 2: 1.1%

Size of most significant component compared to the network, in network 1: 38.1% Average clustering of network 1: 0.345

Degree correlation of network 1: 0.487
So, network 1 is clearly disassortive

**Network 2 statistics:**

Nodes involve in the connected component of network 2: 752
Networks nodes: 894
Network links: 4391
Average degree: 9.82
Density of network 1 and 2: 1.1%
Size of most significant component compared to the network, in network 2: 84% Average clustering of network 2: 0.0018
Degree correlation of network 2: -0.503
Network 2 is assertive

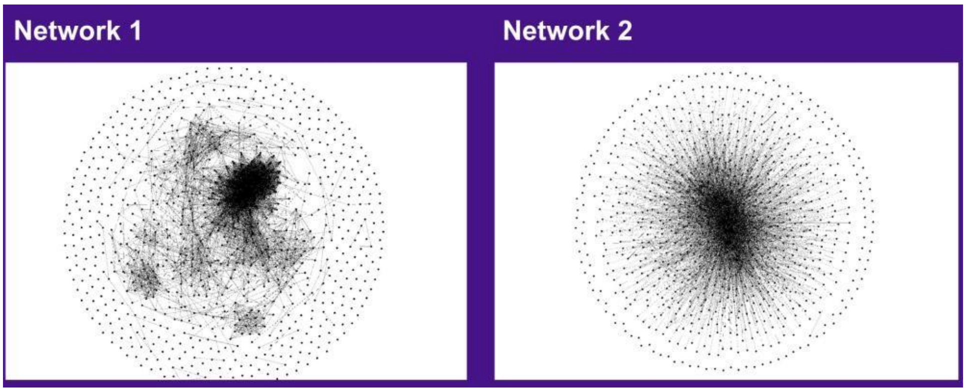


Figure 2: Network 1 and Network 2

**ER model statistics:**

Density of ER model: 1.1%
Size of most significant component compared to the network, in ER model: 99.9% Average clustering of ER model: 0.0107
Degree correlation of ER model: 0.0358

**BA model statistics:**

Density of ER model: 1.1%
Size of most significant component compared to the network, in ER model: 100% Average clustering of BA model: 0.042
Degree correlation of BA model: -0.0485

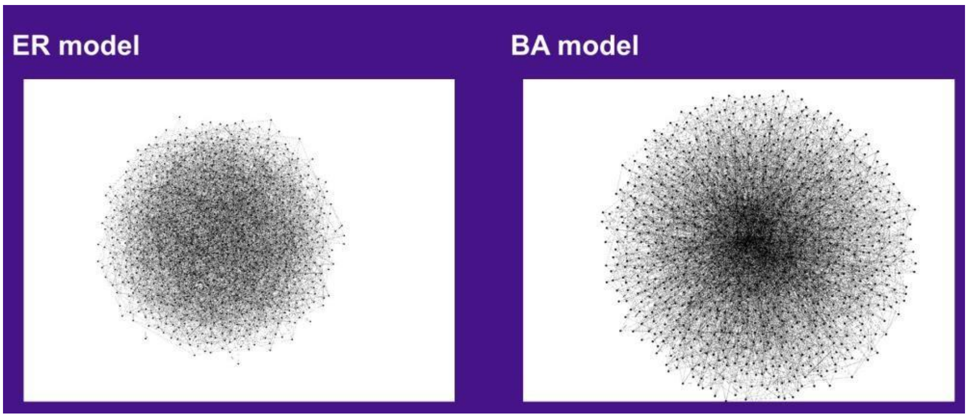


Figure 3: ER model and BA model

So, BA and ER model are neutral to slightly assortative

We compare other statistics:

**<u>Shorted path length distribution:</u>**

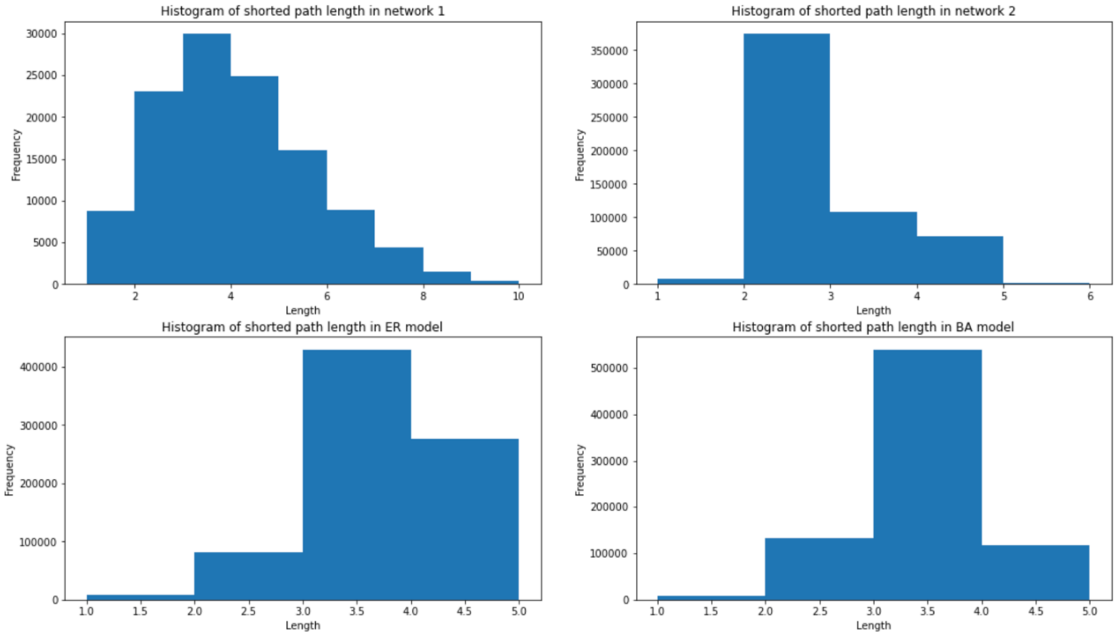


Figure 4: Shorted path length distributions

From figure 2, the shorted path length of network 1 strongly follows the Poisson distribution while that of network 2 follows an exponential distribution. This raises a hypothesis that network 2 is more likely a scale-free network.

**<u>Clustering distribution:</u>**

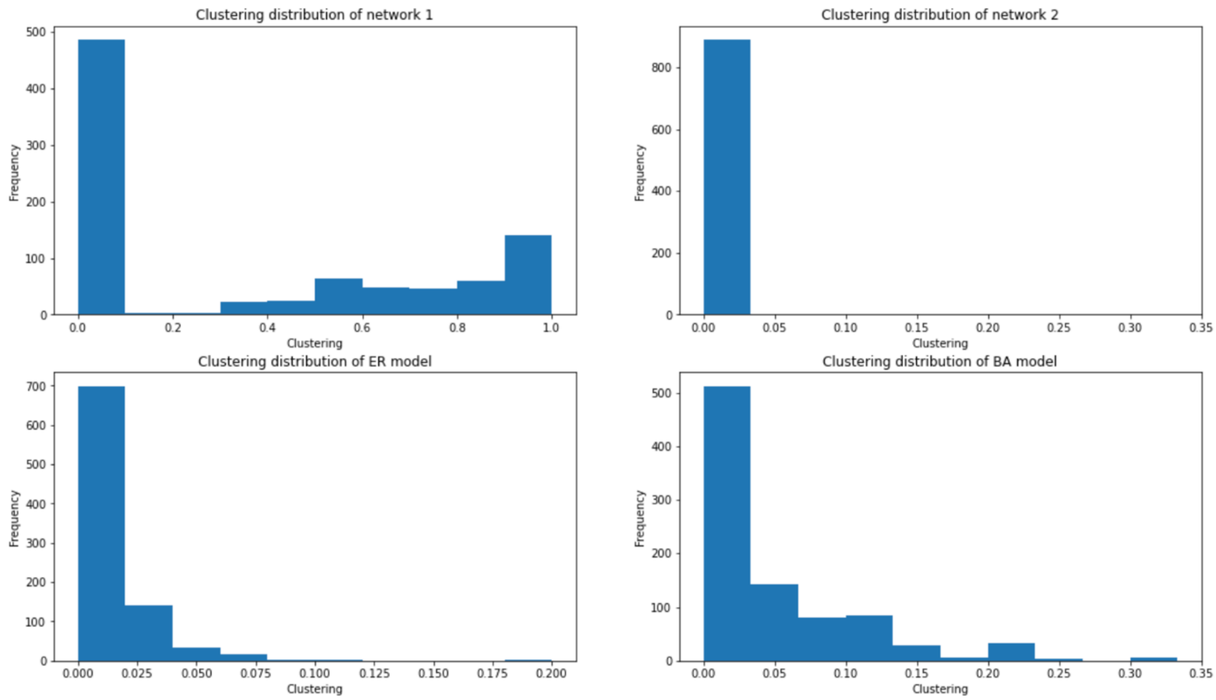


Figure 5: Clustering distributions

The clustering distribution of network 1 is special, there are many nodes with large clustering coefficient indicating very dense hubs.

## Degree distribution:

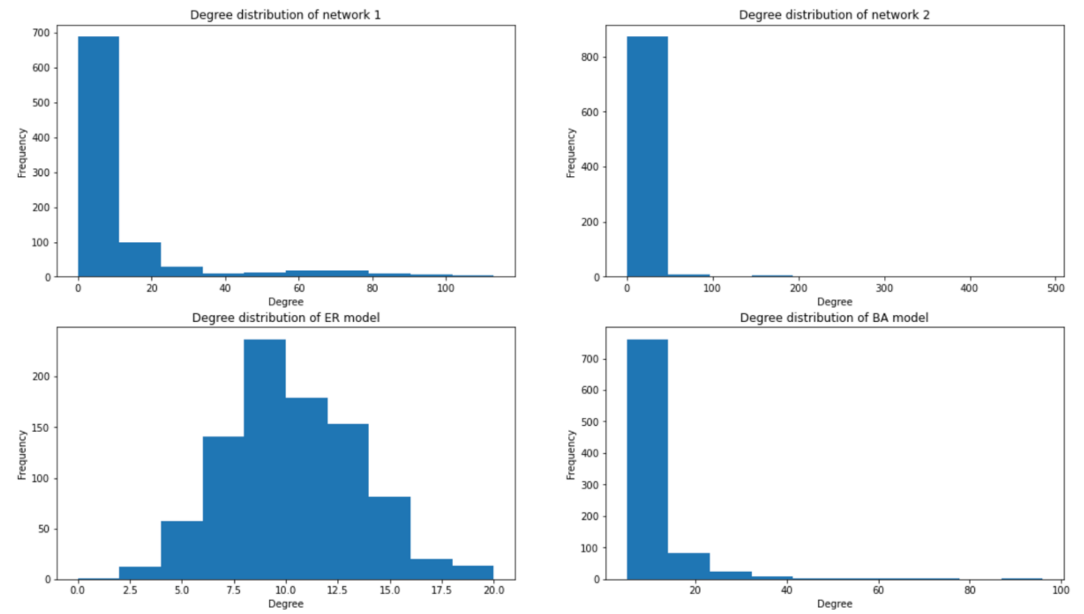


Figure 6: Degree distributions

Both networks 1 and 2 have many nodes with low degrees. But while network 1 has many nodes with medium degrees, network 2 has many nodes with degree 1 and a node with a large degree. This is a sign that network 2 is a scale-free network.

## Network robustness:

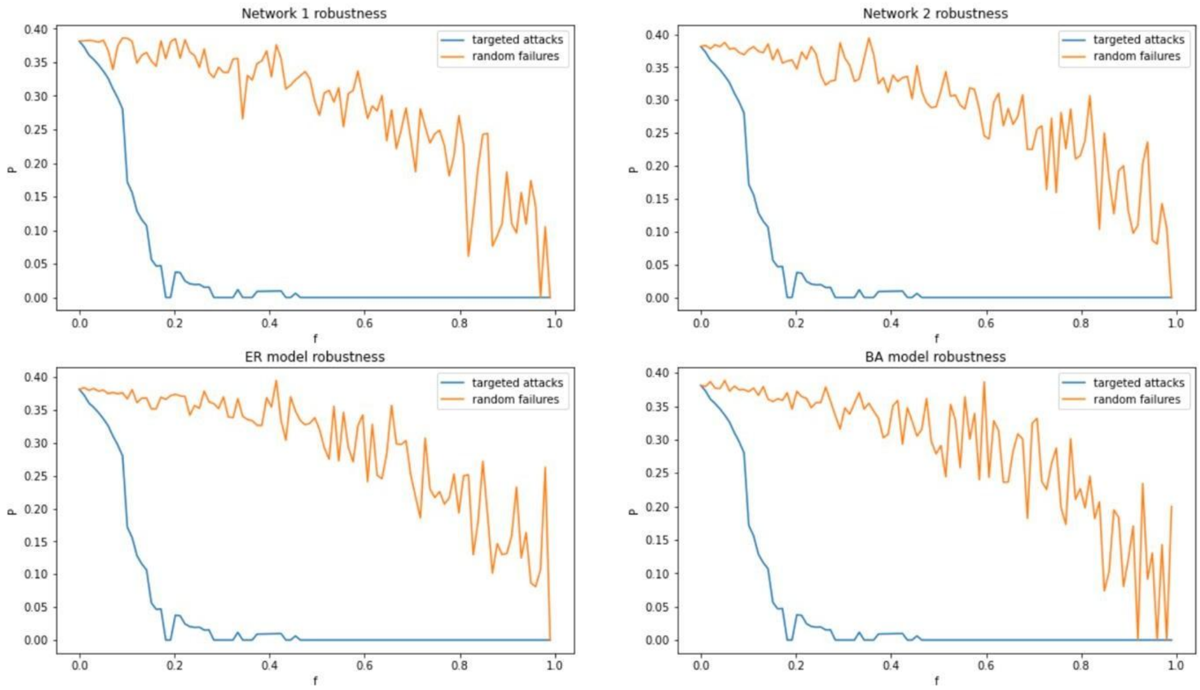


Figure 7: Network robustness

Both network 1 and network 2 show strong resistance to random failure, vulnerable to a targeted attack.

## Degree assortative:

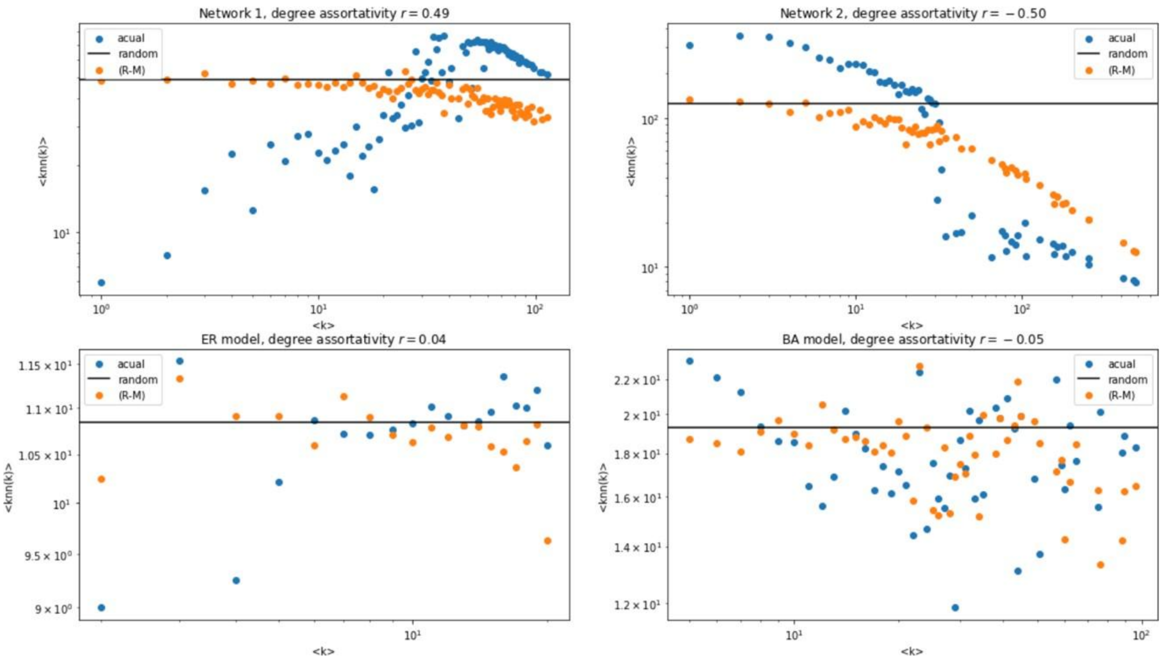


Figure 8: Degree assortative

Network 1 is disassortative, and network 2 is assortative

## Spreading phenomena:

We simulate the spreading of an event or information in the network. For financial networks, it may be bad or good news in the market. The spreading phenomena represent how market price in the news. The model utilized is the SIR model.

Parameter:

Initial source: 5

P (spreading): 0.01

mu (recovered): 0.01

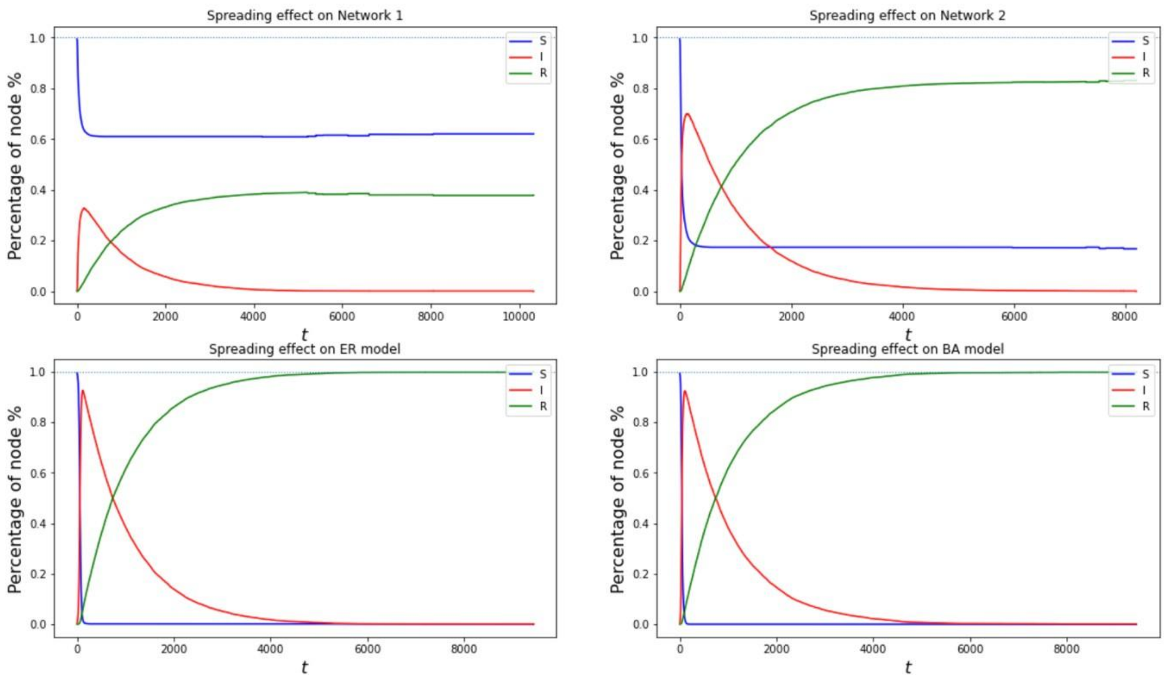

Figure 9: Spreading phenomen

Interestingly, networks 1 and 2 have an infectious size significantly smaller than ER and BA models (max It). We explain that networks 1 and 2 have many isolated nodes or isolated small components that are not affected during the spreading phenomena.

The averaging spreading time is also shorter:

Average spreading time of Network 1: 6451.5
Average spreading time of Network 2: 7363.7
Average spreading time of ER model: 7013.0
Average spreading time of BA model: 7342.1

Moreover, we investigate the spread phenomena for different value of mu compared to p
From mu = 10% 20%...100% p

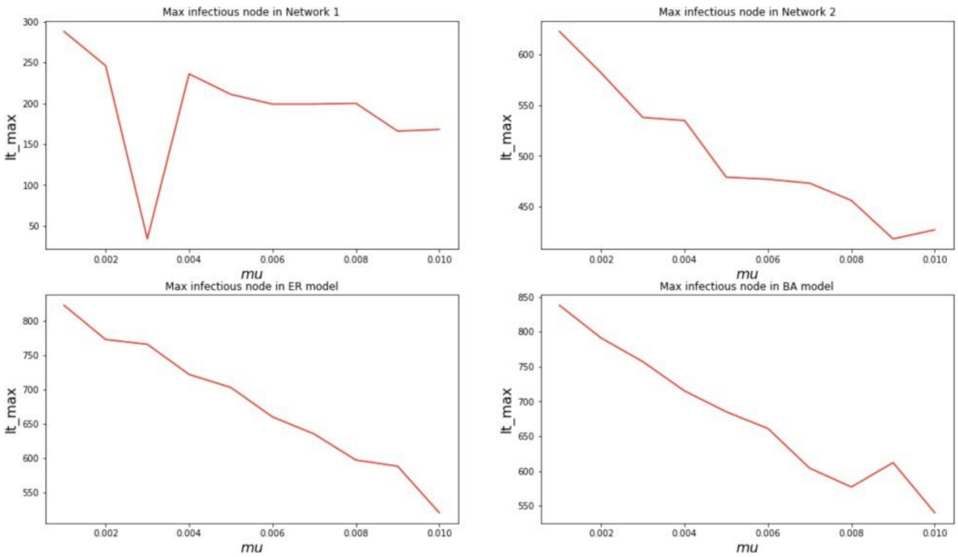


Figure 10: Max infectious size vs. mu

We see that the infectious size degree linearly when mu increase. However, for the contagious time, it decreases exponentially.

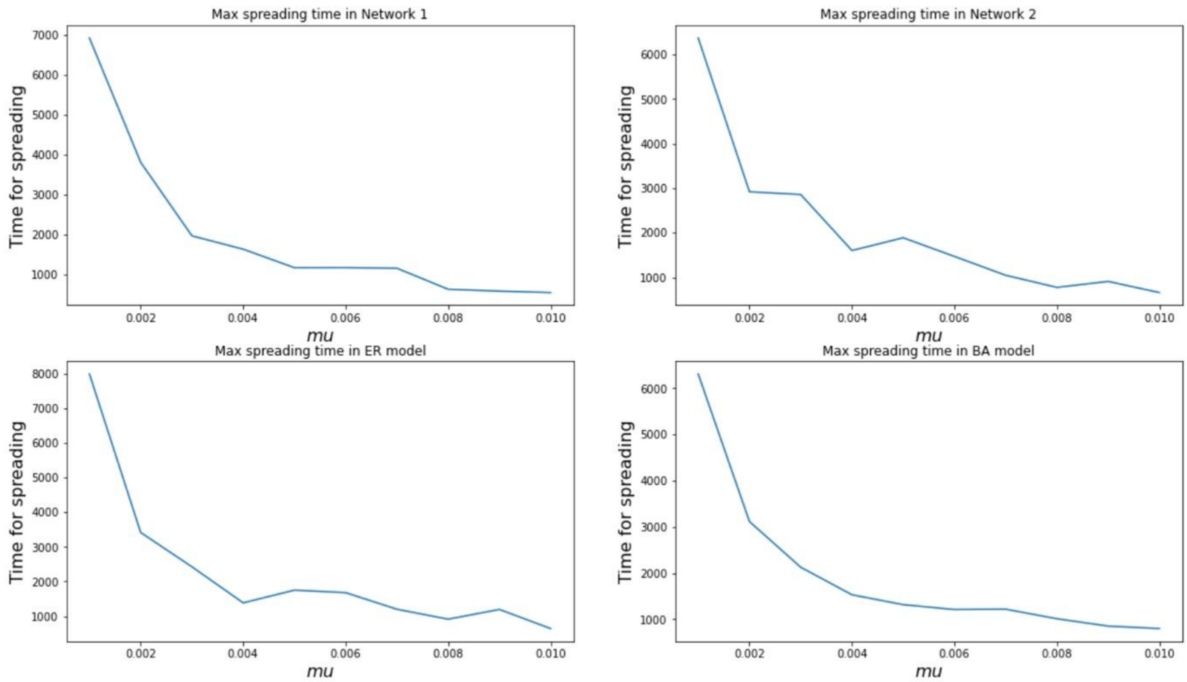


Figure 11: Spreading time decreases exponentially on every network.

# Portfolio Selection

**Problem:** The movement of financial markets is complete stochastics. However, in practice, this idea is not true. The market has momentum. And by following the market's momentum, we can hope to generate profit. So, we will look back at the historical price in a favorable market to see how profit can be made.

**Idea:**

At a specific date X, we calculate the Pearson correlation price movement of 120 trading days before X (120 trading days ~ 6 months). From the collected correlation, we construct networks 1 and 2 as the last parts but with 2% quantile.

We define a score ***xy*** = D + C + BC + EC for each node, with
D: degree centrality
C: closeness centrality
BC: betweenness centrality
EC: eigenvector centrality.

In each network 1 and 2 of date X, we construct a portfolio with 10 nodes that have the highest score ***xy.***
In network 1, nodes with large ***xy*** are more correlated to the overall nodes. While in network 2, nodes with large ***xy*** are more uncorrelated to the overall nodes.

**Method**

To have an inference that date X, we also construct a portfolio with 10 random stocks and a portfolio with all the stocks in the market.
All portfolios have equal weight.
After 80 trading days (~ 4 months), we evaluate the profit of this portfolio in percentage.

**Tactics:**

We use the Dolar cost averaging strategy to construct a portfolio in this problem. It means we buy portfolios on the same weekly date with the same amount.
Over 2 years, from 23/03/2020 to 22/04/2022, we can construct 64 portfolios (left alone some time to construct Pearson correlation and the maturity date for the last portfolio)

**Priori Knowledge:**

We have deliberately chosen an uptrend period, so we expect the most correlated stocks will perform better.

**Result:**

In total, we have 64x4 = 256 portfolios.

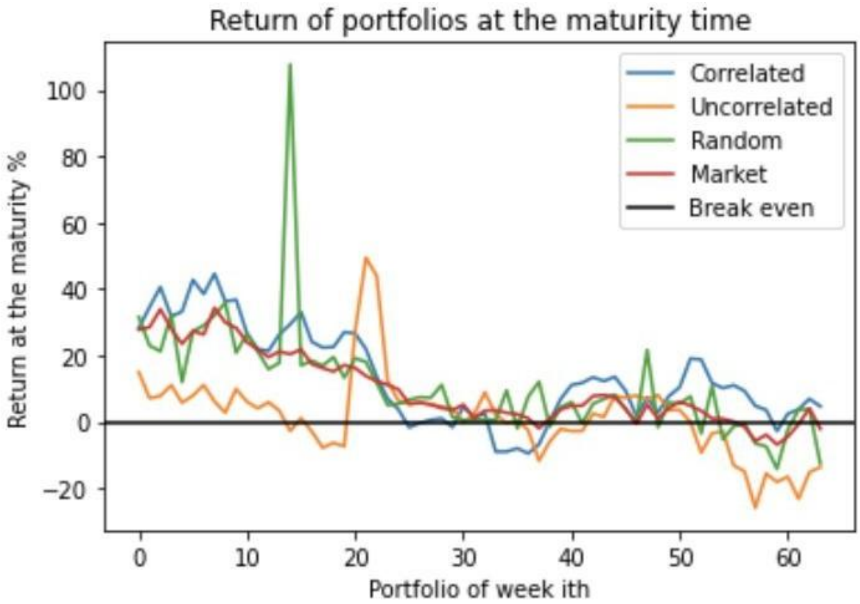


Figure 12: Performance of portfolios chosen each week.

Portfolios constructed by correlated stock consistently outperforms the market

The overall result of 4 strategies

The expected return of correlated portfolio strategy: 13.56%

The expected return of uncorrelated portfolio strategy: 1.57%

The expected return of random portfolio strategy: 10.86%

The expected return of market portfolio strategy: 9.84%

To better estimate the performance, we calculate the Expected return/Risk ratio

**Evaluation criteria: information ratio**

$$r_t(\tau) = [Price(t+\tau) - Price(t)]/Price(t), \text{ over a year } (\tau = 1, .., 250)$$

Standard deviation $s(\tau)$

Average return $\bar{r}(\tau)$

$$\text{Information ratio} = \frac{\bar{r}(\tau)}{s(\tau)}$$

Figure 13: Information ratio to be calculated for each strategy (1)

The larger the ratio, the better, since we want to maximize return while minimizing risk.

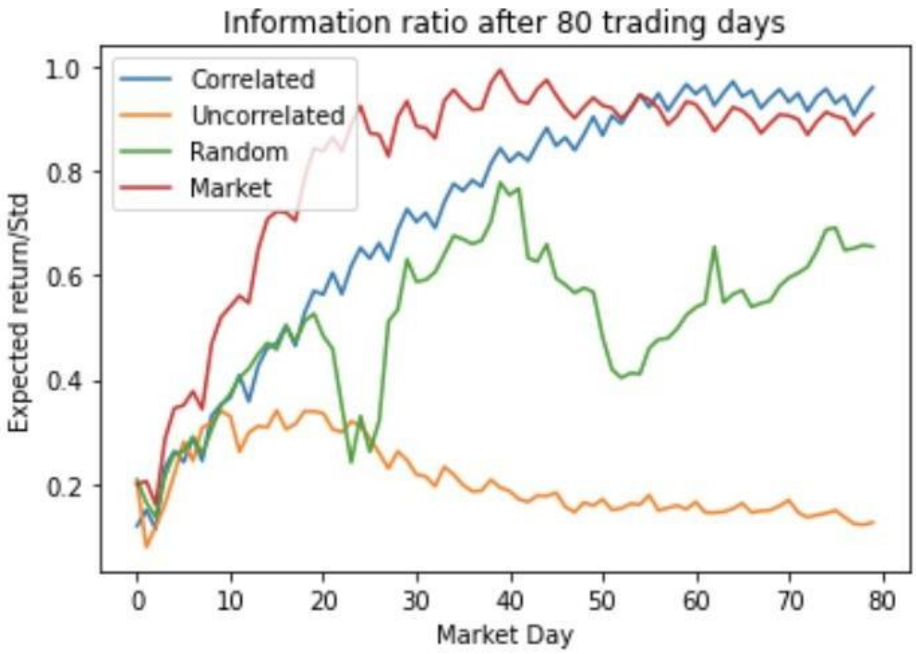


Figure 13: information ratio

From the plot, in the first month after constructing portfolios. However, with the maturity approach, the correlated strategy tends to outperform the market, and the uncorrelated strategy performs badly. This experiment only holds under the condition that the market is favorable, which is not always true.

## What can we improve?

There are multiple aspects to improve:

- Using different correlation methods
- Using exponential smoothing weight
- Apply Minimum spanning tree or Planar Maximally Filtered Graph to construct the network
- Code optimization, the current code is inefficient and take a long time to run
- Using dynamics weight for the portfolios, current portfolios using only equal weight
- Increase data to include bear market

## Conclusion:

In summary, we have successfully constructed financial networks and evaluated their statistics of them. It is interesting to find out that network 2 is a scale-free network or network 1 is disassortative, and we found out the network is resistant to random failure. In the spreading phenomena, we found that networks 1 and 2 have significantly smaller infectious sizes than random models as they have many isolated nodes and small components. Also, the infectious size degree and spreading time degree exponentially with mu increases linearly. In portfolio selection, we have confirmed prior knowledge that a group of central stocks performs better in a favorable market.

## Inference:


(1) Pozzi, F., Di Matteo, T. & Aste, T. Spread of risk across financial markets: better to invest in the peripheries. Sci. Rep. 3, 1665; DOI:10.1038/srep01665 (2013).

(2) Tumminello, M., Lillo, F. & Mantegna, R. N. Correlation, hierarchies, and networks in financial markets. Journal of Economic Behavior & Organization 75, 40–58 (2010).