

ANALYSIS OF STOCK MARKET USING COMPLEX NETWORKS AND MACHINE LEARNING

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF MASTER OF SCIENCE
IN THE FACULTY OF SCIENCE AND ENGINEERING

2018

By
Lingjie Zhang
School of Computer Science

Contents

Abstract	6
Declaration	7
Copyright	8
Acknowledgements	9
1 Introduction	11
1.1 Motivation	11
1.2 Objectives	12
1.3 Proposed methodology	12
1.3.1 Construction of the stock network	12
1.3.2 Logarithmic return of stock prices	13
1.3.3 Topological properties of directed networks	13
1.3.4 Robustness test	15
1.4 Summary of results	15
1.5 Outline of the thesis	15
2 Short Chap title	16
3 Short Chap title	18
4 Short Chap title	20
4.1 Data	20
4.2 Network construction	20
4.3 Analysis of the directed unweighted network	23
4.3.1 Power-law distribution	23
4.3.2 Small-world property	25

4.3.3	Community feature	25
4.4	Analysis of the directed weighted network	30
4.4.1	Analysis on the relationship between price return and clustering coefficient	30
4.5	Stability of network	30
5	Conclusions and future watch	34
6	Making a bibliography	35
	Bibliography	36
A	Example of operation	37
A.1	Example input and output	37
A.1.1	Input	37
A.1.2	Output	38
A.1.3	Another way to include code	38

Word Count: 1626

List of Tables

4.1	Main properties of stock network, small-world network, and random network	23
4.2	Main topologies of conventional stock price network	26
4.3	Main topologies of directed weighted stock price network	26

List of Figures

4.1	Transaction densities	21
4.2	Stock price return correlation coefficient distribution	21
4.3	22
4.4	Numbers of directed edges per EIO-threshold and correlation-coefficient-threshold	22
4.5	Transaction densities	24
4.6	Community structure of the 2016 US stock price return network. Five distinct communities are detected represented by different colours of nodes. The direction of edge is clockwise. The size of nodes and thickness of edges are related to the value of degrees and weights. The grey nodes do not belong to any communities and most of them have zero degree.	27
4.7	Solely community views of the directed stock network. Stock tickers are displayed for the sparsely distributed communities.	28
4.8	Count of stocks of each communities by industrial sectors. Sectors are arranged alphabetically.	29
4.9	Out-degree distribution of stock network	30
4.10	Out-degree distribution and P-P plot of small-world network	31
4.11	Out-degree distribution and P-P plot of random network	32
4.12	Bivariate distribution	33

Abstract

**ANALYSIS OF STOCK MARKET
USING COMPLEX NETWORKS
AND MACHINE LEARNING**

Lingjie Zhang

A dissertation submitted to the University of Manchester
for the degree of Master of Science, 2018

Due to the complexity of financial market and the interconnect- edness and inter-dependencies of industry sectors in the economy, the price returns of each coupling stocks might have certain underlying economic link. Such behaviours can hardly be explained by traditional financial models and theories. This project combines machine learning techniques, individual stock features, and empirical data of Industry Economic Accounts (IEAs) from Bureau of Economic Analysis (BEA) in the US to predict Granger causality of coupling US stocks. Limited Granger causalities are calculated as a small sample set compared to the target date set. A directed weighted complex network (DWCN) is constructed by considering companies as nodes, correlations of abnormal stock returns (α) as weights of links, and predicted Granger causalities indicate directions of links. The generated DWCN is visualised and its topological properties, sta- bility, and effects on individual stocks and industries are researched in this paper. Suggestions towards financial market investment are provided based on the results of this research.

Declaration

No portion of the work referred to in this dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on presentation of Theses

Acknowledgements

I would like to thank...

Contents

Chapter 1

Introduction

1.1 Motivation

Financial markets are complex systems, the interconnectedness and interdependencies of industry sectors in the economy are highly inter-coupled with strong correlations with stock price fluctuations, i.e., the price returns of each coupling stocks underlying certain economic link, e.g. two companies that manufacture similar products, or both in one supply chain. Such behaviours can hardly be explained by traditional financial models and theories.

During recent times, weighted but undirected complex network models have been applied to study the correlations of stock prices. Prevailing approach is using companies as nodes, and correlations between each pair of stock price series, return series, or fluctuation patterns as links, e.g., much of the previous research have proved the represented complex networks of worldwide stock markets are scale-free and small-world [1, 2].

Theoretically, directed complex network for stock market can be achieved therefore more potential information can be produced which is helpful for investment decision and financial market supervision. Conducting Granger causality test between stock pairs is straightforward but not feasible due to the heavy-precondition and time-complexity in programme. Compared to the large order of magnitude of total stock pairs, manually calculated Granger causalities are too less to be used as training samples. Enlightened by the recent work by Jean et al. [3] which uses transfer learning and noisy proxy information performed very well at predicting poverty, demonstrating that machine learning techniques is powerful to be applied in a setting with limited training data, so an exploration towards the directed network of stock market

is motivated in this project, combines machine learning techniques, transfer learning, individual stock features, and empirical data of Industry Economic Accounts (IEAs) from Bureau of Economic Analysis (BEA) in the US to predict Granger causality of coupling US stocks. Therefore, a directed weighted complex network (DWCN) is constructed by considering companies as nodes, correlations of abnormal stock returns (alpha) as weights of links, and predicted Granger causalities as the indications of directions of links.

The goal of this project is to reveal the Granger causality of price return series and utilise them into the topological analysis and visualisation of DWCN as so far no previous work has attempted to construct a directed network about stock price. In addition, suggestions for stock market are provided according to the results and findings.

The outline of this document is as follows. In Section 2, specific objectives and deliverables are listed. In Section 3, relevant previous researches on financial market, complex networks, and machine learning are reviewed. Section 4 describes the analytical methods. Ethics and professional considerations and risk considerations are respectively discussed in Section 5 and 6. Section 7 describes project evaluation approaches and finally, planning Gantt chart is presented in Section 8.

1.2 Objectives

1.3 Proposed methodology

1.3.1 Construction of the stock network

EIO (Economic Input-Output) table

The Bureau of Economic Analysis (BEA) in the US publishes Economic Input-Output (EIO) tables each year, which are the transaction matrices of all purchases and sales between sectors in a certain industry group level of a year, i.e. depict how industries provide input to, and use output from, each other to produce Gross Domestic Product (GDP).

This paper uses the use table of 2016. The Among the transaction matrix (Z) there are Total Industry Input (I) row at the bottom and the Total Industry Output (O) column at the right are the statistics of total purchase by each sectors and total

sales from each sectors respectively.

$$a_{ij} = -\log_{10}(z_{ij}/I_j)^{-1} \quad (1.1)$$

$$b_{ij} = -\log_{10}(z_{ij}/O_i)^{-1} \quad (1.2)$$

1.3.2 Logarithmic return of stock prices

$$r_i(\tau) = \ln P_i(\tau) - \ln P_i(\tau - \Delta t)$$

$$c_{i,j} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)}} \quad (1.3)$$

Correlation coefficient

1.3.3 Topological properties of directed networks

Degree, out-degree, and degree distribution

The degree of a node k represents the number of its neighbours. In terms of the directed stock price return network, this paper mainly focuses on the out-degree analysis on the nodes.

Out-degree k_{out} is the number of edges which start from the given node and end at others, while in-degree k_{in} is the number of edges which end at the given node and start from others. In directed graph, there is relationship between k_{in} and k_{out}

$$k = k_{in} + k_{out}. \quad (1.4)$$

The degree distribution of stock price return network $p(k)$ can be defined as:

$$p(k) = \frac{N_k}{N}, \quad (1.5)$$

while N_k represents the number of nodes whose out-degree value is k .

Path length

Clustering coefficient

Clustering coefficient is a measure of the degree to which nodes in a network tend to cluster together. Concerning the clustering coefficient of the complex networks, it is defined as:

$$C_i = \frac{2E_i}{(k_i(k_i - 1))}, \quad (1.6)$$

where k_i is the degree of a given node v_i , E_i is the real existing edges among the nearest neighbour nodes of the given node v_i , and $k_i(k_i - 1)/2$ means the maximum possible edges existing between its nearest neighbours of the node v_i . Besides, the clustering coefficient of a node accounts for the extent to which the transmission relationship between the given node and its neighbours also exists between its neighbours, and the clustering coefficient may be given by:

$$C = \frac{3 \times \text{number of triangles in the networks}}{\text{number of connected triples of nodes}}. \quad (1.7)$$

This measure gives an indication of the clustering in the whole network, and can be applied to both undirected and directed networks.

Degree centrality

Degree centrality can be used to determine the critical nodes among the entire network and to recognise the most associated firms in the chosen stock market. When it comes to directed network, two separate measures to calculate degree centrality are defined, which are in-degree and out-degree. In-degree represents the number of edges incoming to a node and out-degree represents the number of edges outgoing from a node.

As one of the most widespread measures to calculate network centrality, degree centrality can be described as the number of direct links that relate to a specific node [?]. While the degree of centrality has generally been accumulated to the sum of weights to form the weighted networks. The equation of this measure is shown as bellow:

$$C_D^W(i) = \sum_j^N w_{ij} \quad (1.8)$$

where W represents the matrix of weighed adjacencies, and w_{ij} represents the weight of the link between node i and j .

1.3.4 Robustness test

1.4 Summary of results

1.5 Outline of the thesis

Chapter 2

Background

The revolutionary who pioneered the theory of portfolio, Markowitz [?] proposed the mean-variance model and established a clear mathematical definition of the two vague concepts of risk and return. Sharp [?] and Linter [?] added two key assumptions based on the mean-variance model to enable the portfolio mean-variance valid, forming a capital asset pricing model (CAPM) with the support of economic theories. The CAPM believes that only non-dispersible systemic risks can be compensated off, while non-systematic risks can be eliminated by effectively decentralised investments. Investors could only assume systemic risks through decentralised investment. The systematic risk of a single security or portfolio can be characterised by beta, which represents the extent to which a single security or portfolio is affected by the overall market volatility.

Worldwide scholars had been actively conducting empirical tests towards the practicality of the CAPM then, while early results show that beta is able to explain return movements of stocks. However, in the late 1970s, some empirical studies upon the CAPM began to show that a large part of the changes in the stock returns cannot be explained by beta, with increasingly market return anomalies were found.

Researchers had proposed models and theories considering the individual stock features to explain. For instance, Fama and French [?, ?] proposed a three-factor model based on the inter-temporal capital asset model (ICAPM) [?] and the arbitrage pricing theory (APT) [?], which reveals a large part of the cross-section of the stocks' average return that cannot be explained by CAPM, can be explained using firm size, book-to-market equity ratio, and overall market return.

While traditional stock pricing models still capture limited forms of financial behaviour, the premises of standard financial theory contradict the modern notion of financial markets are complex systems [?], by which many statistical niceties such as stationarity no longer can be taken for granted. Recent researches have implemented the network theory to reveal the underlying factors of price movements. Huang et al. [?] implemented the threshold method to build 's correlation network in China's A-Share stock market and studied the networks topological properties and topological stability. Namaki et al. [?] utilised Random Matrix Theory (RMT) to specify the biggest eigenvector in the complex network of price correlations. Yu [?] studied the evolution of gold price from a network perspective using the visibility network approach. Chopra and Khanna [?] developed a framework which associates the economic input–output (EIO) model with techniques for understanding interdependencies and interconnectedness in the economy of US, based on complex networks theory. Boginski et al. [?] identified cliques and independent groups among stock networks. Chen et al. [?] studied the inter-stock and inter-industry effects towards stock returns based on the topological properties of a complex network of correlations.

In a nutshell, prevailing complex network approaches to analyse stock markets are almost all about investigating weighted or unweight but undirected networks. To the best knowledge of mine, no previous work has attempted to construct a directed network so far.

Chapter 3

Community detection

This paper considers an approach to detect possible communities in directed networks based on modularities [LN08].

Algorithm 1 Community detection

```

1: procedure COMMUNITY( $G, nNode, nEdge, EntireModMat, EntireNodeSpace$ )
2:   procedure CALDELTAQ( $s, \mathbf{B}$ )
3:     return  $Q \leftarrow 1 / (4 * nNode) * s^T (\mathbf{B} + \mathbf{B}^T) s$ 
4:   procedure UPDCOMMUNITYASSIGNMENT( $NodeSpace, UpdAssign$ )
5:      $Mark1, Mark2 \leftarrow \max(Assignment) + 1, \max(Assignment) + 2$ 
6:     for each  $node \in NodeSpace$  do
7:       if  $node \in UpdAssign > 0$  then node of Assignment  $\leftarrow Mark1$ 
8:       if  $node \in UpdAssign < 0$  then node of Assignment  $\leftarrow Mark2$ 
9:     return Assignment
10:    procedure SUBDIVIDECOMMUNITY( $\mathbf{B}$ )
11:       $SymmetricMatrix \leftarrow \mathbf{B} + \mathbf{B}^T$ 
12:       $eigv \leftarrow \text{eigenvector as } \max(eigenvalues) \text{ in } SymmetricMatrix$ 
13:      return sign( $eigv$ )
14:    procedure CALMODULARITY( $assignment$ )
15:      for each  $node1 \in Nodes of G$  do
16:        for each  $node2 \in Nodes of G$  do
17:          if assignment of  $node1 \leftarrow$  assignment of  $node2$  then
18:             $Q \leftarrow Q + HasEdge - (nIn(node1)) * (nOut(node2)) / (nEdge)$ 
19:      return  $Q / (nEdge)$ 
20:    procedure GENMODULARITYMATRIX( $NodeSpace, ModMat$ )
21:      for each  $node1 \in NodeSpace$  do
22:        for each  $node2 \in NodeSpace$  do
23:           $B \leftarrow HasEdge - (nIn(node1)) * (nOut(node2) / nEdge)$ 
24:          if Assignment of  $node1 =$  Assignment of  $node2$  then
25:            for each  $node \in NodeSpace$  do  $C \leftarrow C + HasEdge1 +$ 
26:             $HasEdge2 - (nIn(node1) * nOut(node) + nIn(node) * nOut(node1)) / nEdge$ 
27:    procedure INTERATEBISECTION( $ModMat, NodeSpace$ )
28:       $UpdAssign \leftarrow \text{SubdivideCommunity}(ModMat)$ 
29:       $DeltaQ \leftarrow \text{CalDeltaQ}(UpdAssign, ModMat)$ 
30:      if  $DeltaQ > 0$  then
31:         $Assignment \leftarrow \text{UpdCommunityAssignment}(NodeSpace, UpdAssign)$ 
32:        for each  $side \in UpdCommunityAssignment$  do
33:           $ModMat \leftarrow \text{GenModularityMatrix}(NodeSpace)$ 
34:           $\text{InterateBisection}(ModMat, NodeSpace)$ 
35:      return Assignment
  
```

Chapter 4

Empirical study and results

4.1 Data

This paper considers 1,418 stocks of listing US companies that traded consecutively in the NYSE and NASDAQ stock market of US from January 4, 2016 to December 30, 2016 and uses daily closing price during this period.

4.2 Network construction

This paper first calculated normalised direct requirement and normalised direct demand values for every sectors using formula 1.1 and formula 1.2 and cross-correlation coefficients between every stock pairs using formula 1.3.

According to the figure 4.1, the transaction densities decrease as threshold of normalised direct requirement and normalised direct demand increase, and their patterns are exactly similar with the same inflection point at around $threshold = 0.136$ where the densities begin to decline. Therefore, the values of thresholds for normalised direct requirement and normalised direct demand are set to be same, i.e., $\theta_{DD} = \theta_{DR} = \theta_{EIO}$, to filter the directed edges among the stock network.

Figure 4.2 shows the distribution of stock price correlation coefficients has a shape complies to the normal distribution for the long tails. The correlation coefficients are vary from -0.687 to 0.977 with the mean of 0.265. It implies that the prices of most stocks traded in NYSE and NASDAQ usually fluctuate to the same direction.

Figure 4.4 shows the number of directed edges remain in the conditions of different value combinations of θ_{EIO} and θ_{corr} . When both of the thresholds set to be minimal at their own value range, i.e., $\theta_{EIO} = 0$ and $\theta_{corr} = -1$, the number of directed edges

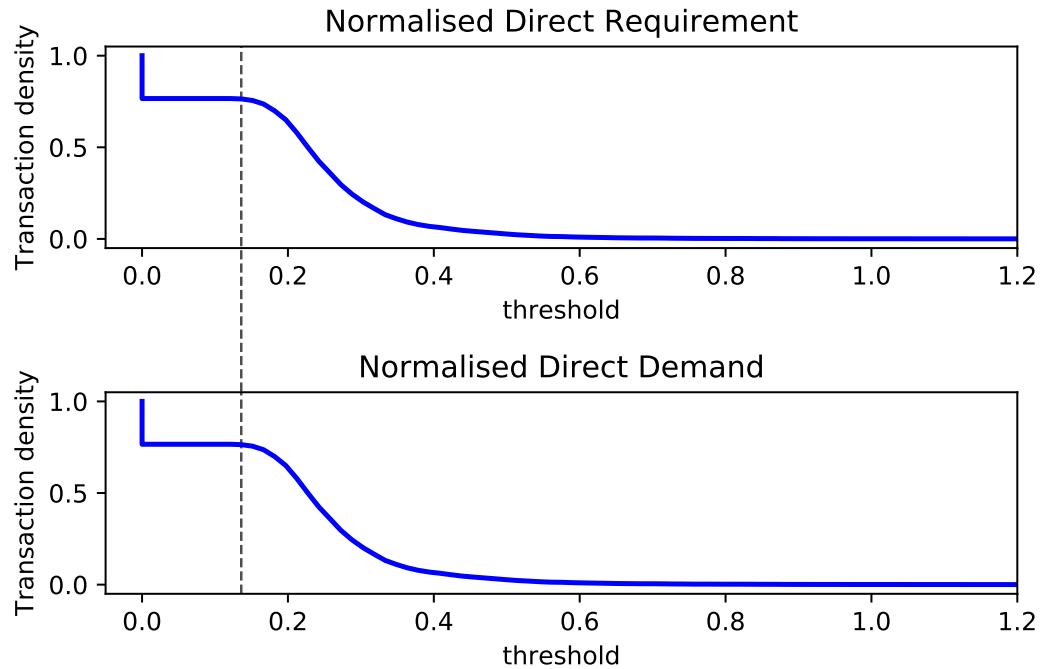


Figure 4.1: Transaction densities

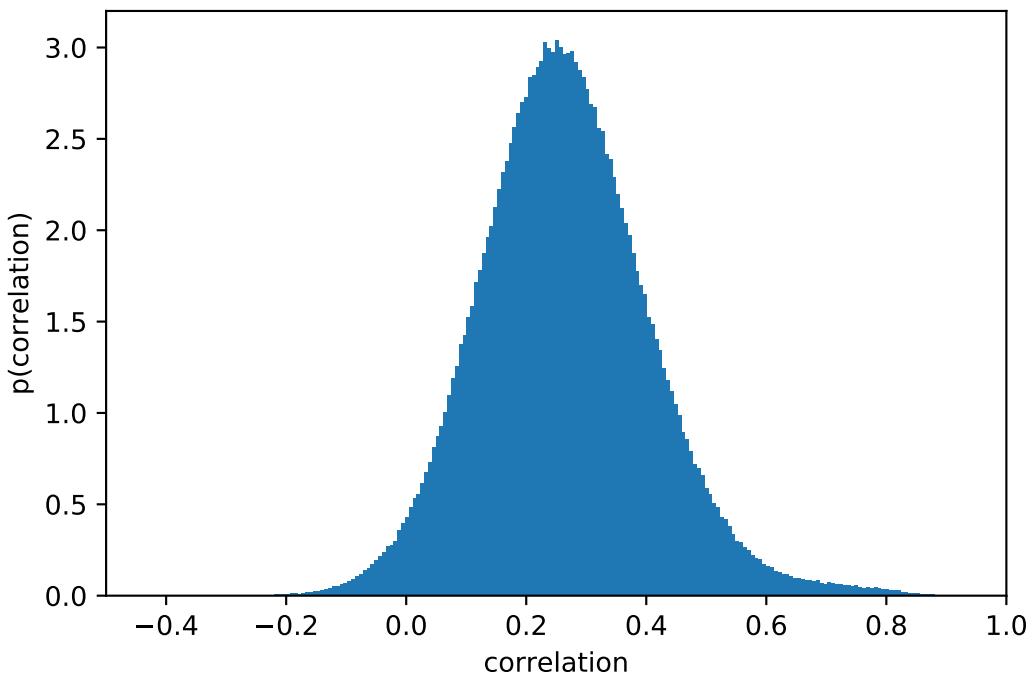


Figure 4.2: Stock price return correlation coefficient distribution

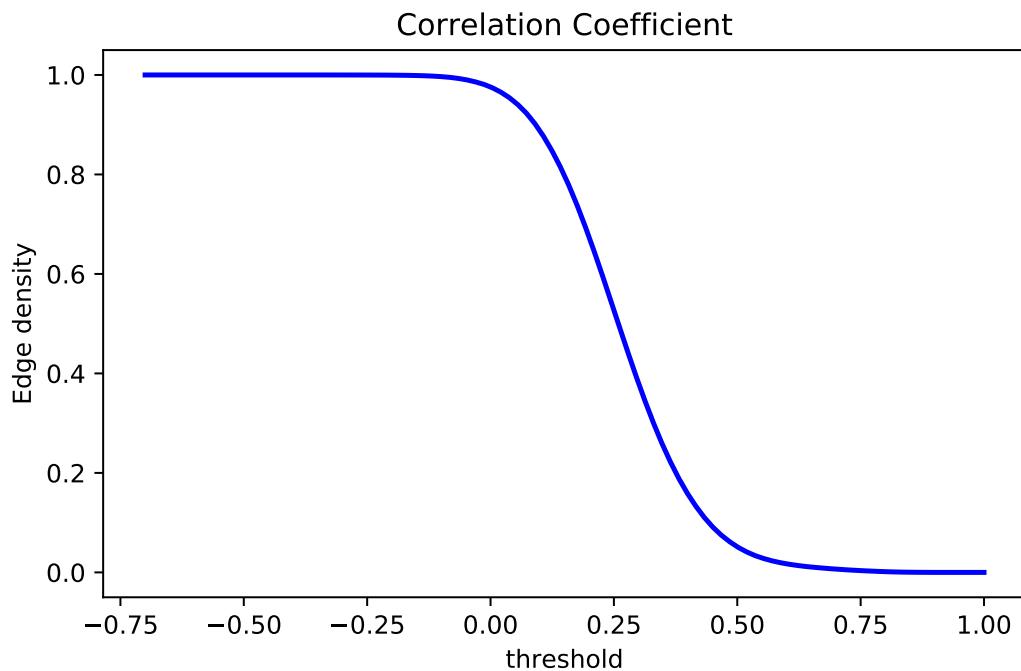


Figure 4.3

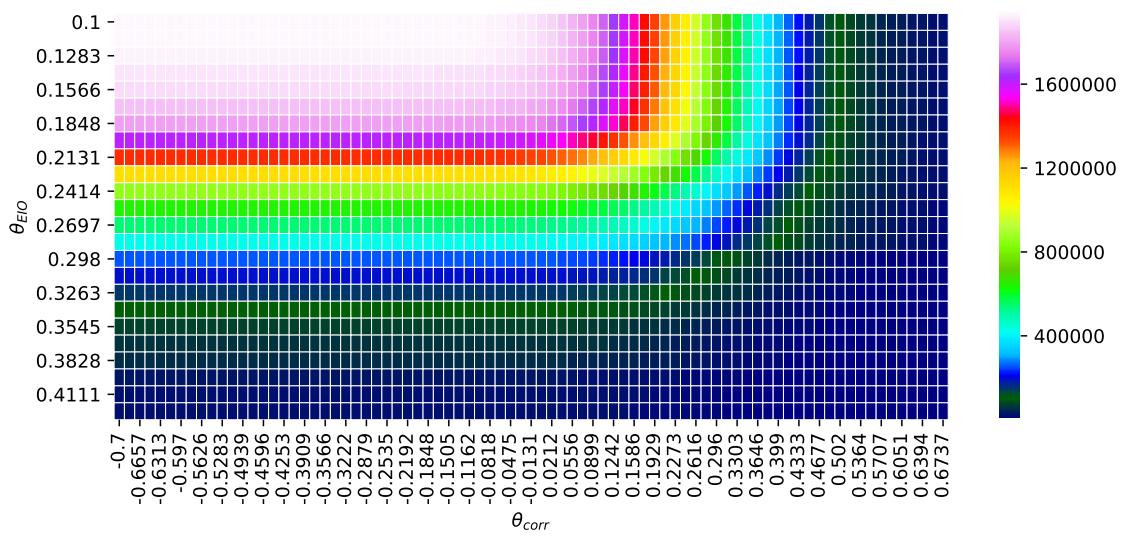


Figure 4.4: Numbers of directed edges per EIO-threshold and correlation-coefficient-threshold

Directed networks	Stock network	Small-world network	Random network
Number of nodes	1418	1418	1418
Number of edges	102051	102088	102097
Out-degree distribution	Power-law	Normal	Normal
Average out-degrees	143.94	143.99	144.00
Average path length	2.775	2.005	1.973
Clustering coefficient	0.4675	0.1367	0.05105
Global efficiency	0.2563	0.5161	0.5216
Local efficiency	0.6276	0.5027	0.4456
Assortativity	0.02004	-0.002180	0.001452

Table 4.1: Main properties of stock network, small-world network, and random network

is $N \times (N - 1) = 2,009,306$, while N indicates the total number of nodes. According to the figure 4.4, the number of edges will be less than 100,000, in which case the network has a density of less than 5%, if θ_{EIO} is larger than 0.3545 or θ_{corr} is larger than 0.5020.

It is obvious that the larger values assigned to θ_{EIO} and θ_{corr} , the more significant for the weights and directions of the remaining edges. But if the network is too sparse, it can not be strongly or weakly connected hence the network becomes inefficient. As a result, this paper selects the threshold-value-pair $\theta_{EIO} = 0.292$ and $\theta_{corr} = 0.379$ to construct directed unweighted stock network and directed weighted stock network.

4.3 Analysis of the directed unweighted network

A directed Watts Strogatz small-world network and a random network with the same number of nodes and edges with the stock directed unweighted network are generated. Table 4.1 compares the main topologies of the three different networks.

4.3.1 Power-law distribution

According to the table 4.1, the values of average out-degrees of stock network, small-world network and random network are almost the same due to the similar number of nodes and edges, but in terms of the distributions of out-degrees, stock network is different from the others.

Figure 4.10 and 4.11 show clearly that the out-degree distribution of small-world

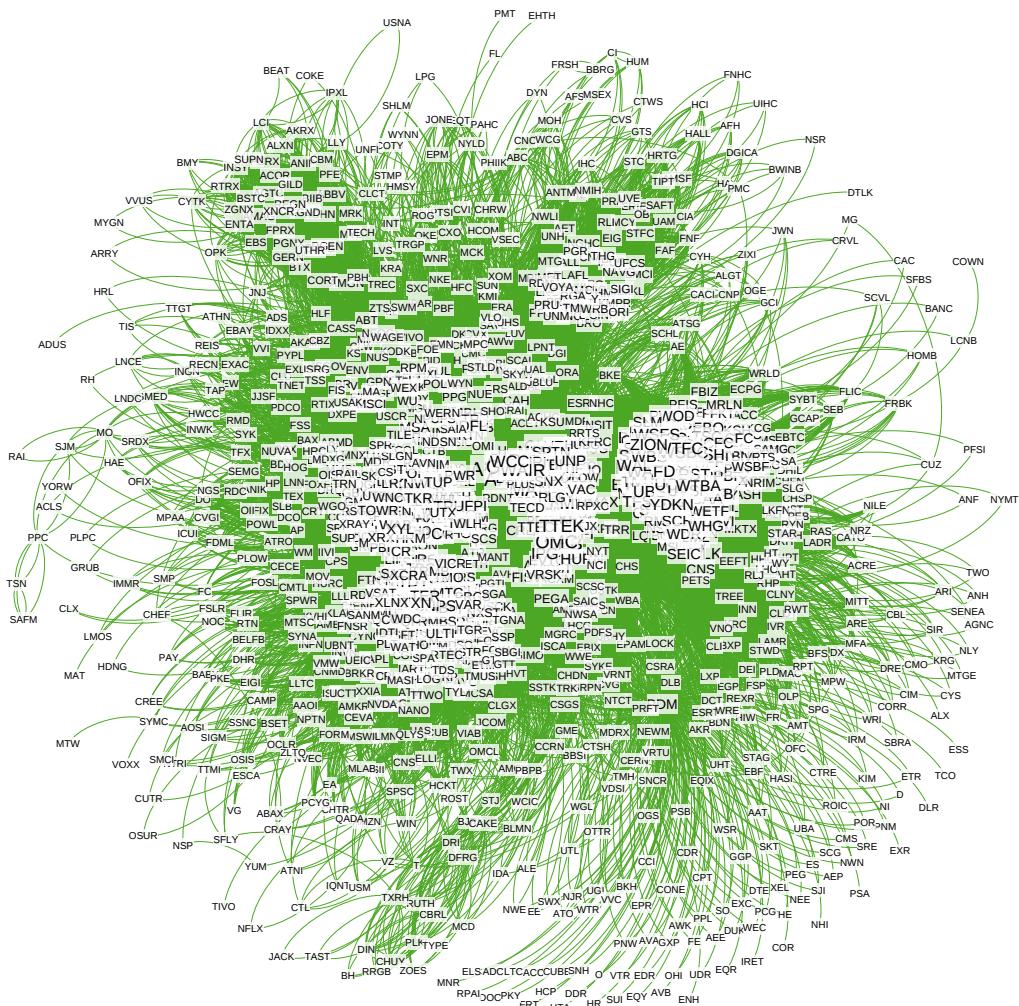


Figure 4.5: Transaction densities

network and random network follow the normal distribution, i.e, most degree numbers of nodes fall in the middle range, while figure 4.9a illustrates that for the stock network, only a few number of nodes show higher out-degree, while most nodes are in the positions of low out-degree level. The distribution of the directed stock price return networks follows power-law distribution with the exponent of 4.057.

4.3.2 Small-world property

The average path length of small-world network and random network are both close to 2, indicates that taking any node in the network, it can expect to reach any other nodes just through one node as medium. For stock network, the expectation number of medium nodes is 1.775, which is also a small number for connection, so like small-world and random networks, stock network also has the small-world property.

4.3.3 Community feature

The larger value of clustering coefficient for stock network than the other two networks indicating that the nodes in stock network tend to cluster together. Therefore, communities of stock network will be identified implementing the community detection algorithm for directed networks in this section. According to the industrial sectors to which most component stocks are belong, as figure 4.8 shows, following five communities are identified: (1) Production (2) Finance (3) Livelihood (4) Insurance and chemical products (5) Utilities and financial vehicles.

It can see from figure 4.6 that the community of finance (green) has a very dense structure, connected closely inside and completely exclusive from other communities. The communities of production (purple) and livelihood (blue) are sparsely distributed while there are some large-sized nodes acting as hubs of the overall network. The other two communities are more interesting because they have peculiar structures.

Every industrial sectors of individual stocks in community are identified to investigate the properties of the community of insurance and chemical products (yellow). As figure 4.7d illustrates and through the investigation, most firms in the upper and lower clusters are in the sectors of "chemical products" and "Insurance carriers and related activities" respectively, while firms between the two big clusters, like "MCK" (McKesson) and "CAH" (Cardinal Health), are large medical supplier, pharmaceutical and healthcare service companies with high out-degrees to both of the two clusters. Apart from that, there are also considerable number of links from the upper cluster

	Undirected stock network
Degree distribution	Power-law
Average out-degrees	Power-law
Average path length	2.775
Clustering coefficient	0.4675
Global efficiency	0.2563
Local efficiency	0.6276
Assortativity	0.02004

Table 4.2: Main topologies of conventional stock price network

	Undirected stock network
Strength distribution	Power-law
Average centrality	1
Weighted assortativity	0.1244

Table 4.3: Main topologies of directed weighted stock price network

to the hubs of chemical companies. Thus, it is reasonable to infer that the prices of medicines have significant influence to medical insurance industry, additionally the purchases of chemical products of pharmaceutical firms and the sales of chemical products have made pharmaceutical and chemical companies influence to each other.

Another investigation towards the community of utilities and financial vehicles (orange) is conducted by the same measure. As figure 4.7e illustrates, there is one huge hub (PDM) which is the company "Piedmont Office Realty Trust" among the community while all the others are one-degree nodes located remotely. More links from the hub to the rest than the opposite direction, and also the weights of the former links are generally higher. The hub, "Piedmont Office Realty Trust", is a real estate investment trust company, and the rest in the community contains 59 "Funds, trusts, and other financial vehicles" firms and 44 "Utilities" firms.

Although the assortativity values for the three networks are all non-significant,

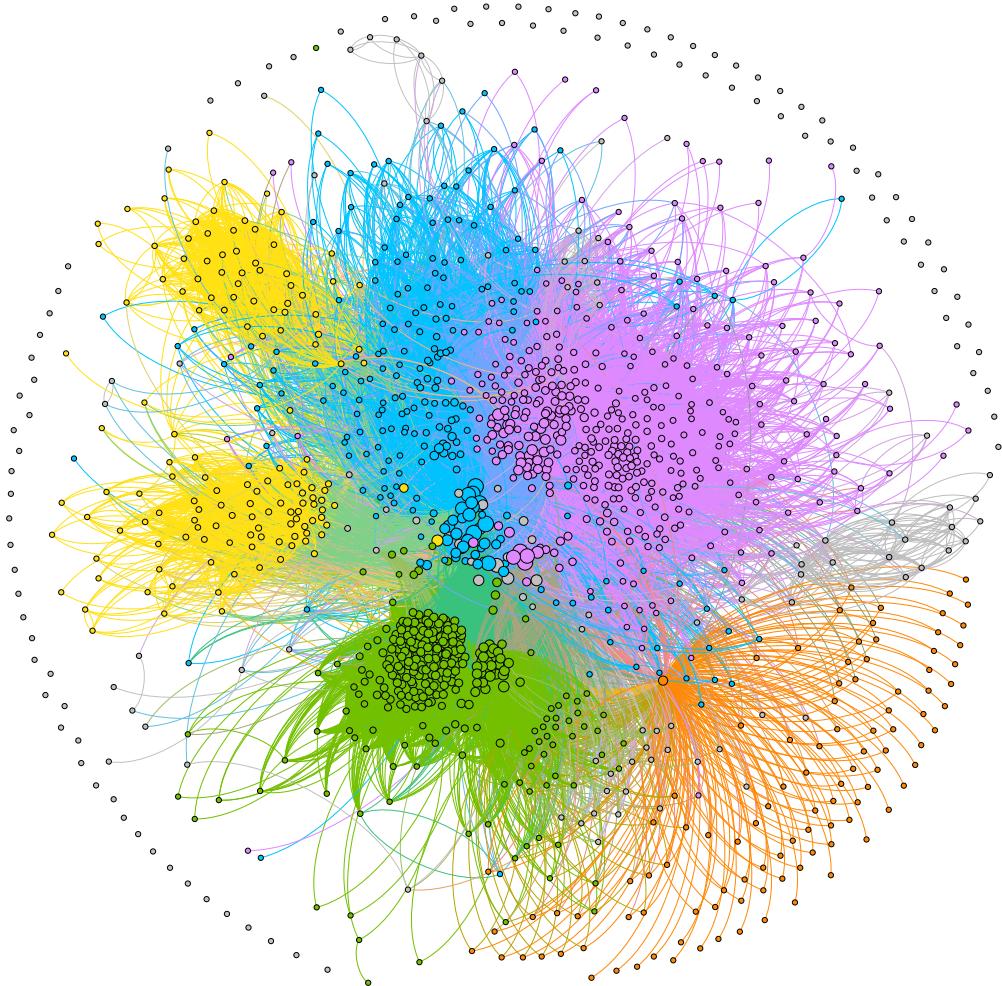


Figure 4.6: Community structure of the 2016 US stock price return network. Five distinct communities are detected represented by different colours of nodes. The direction of edge is clockwise. The size of nodes and thickness of edges are related to the value of degrees and weights. The grey nodes do not belong to any communities and most of them have zero degree.

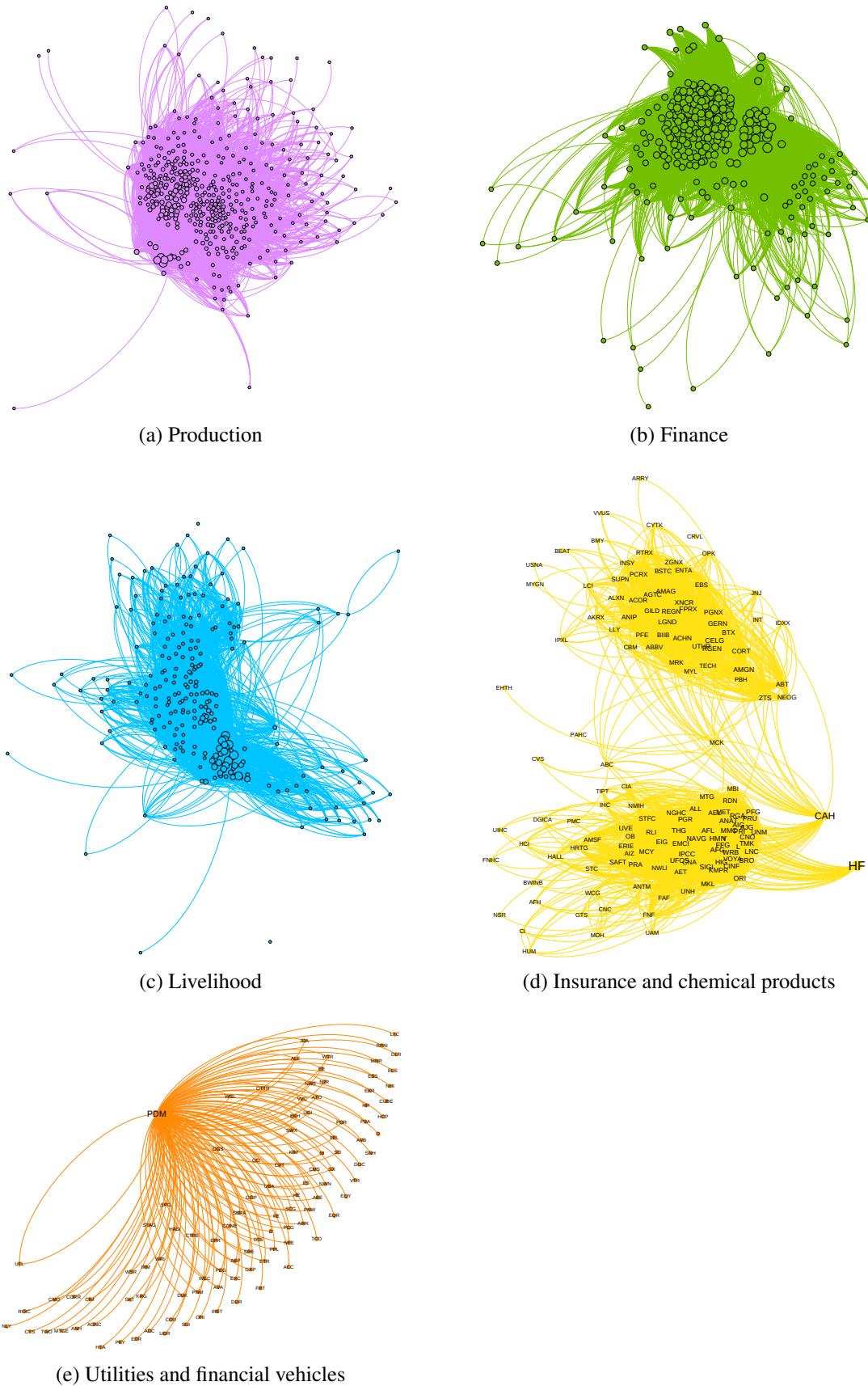


Figure 4.7: Solely community views of the directed stock network. Stock tickers are displayed for the sparsely distributed communities.

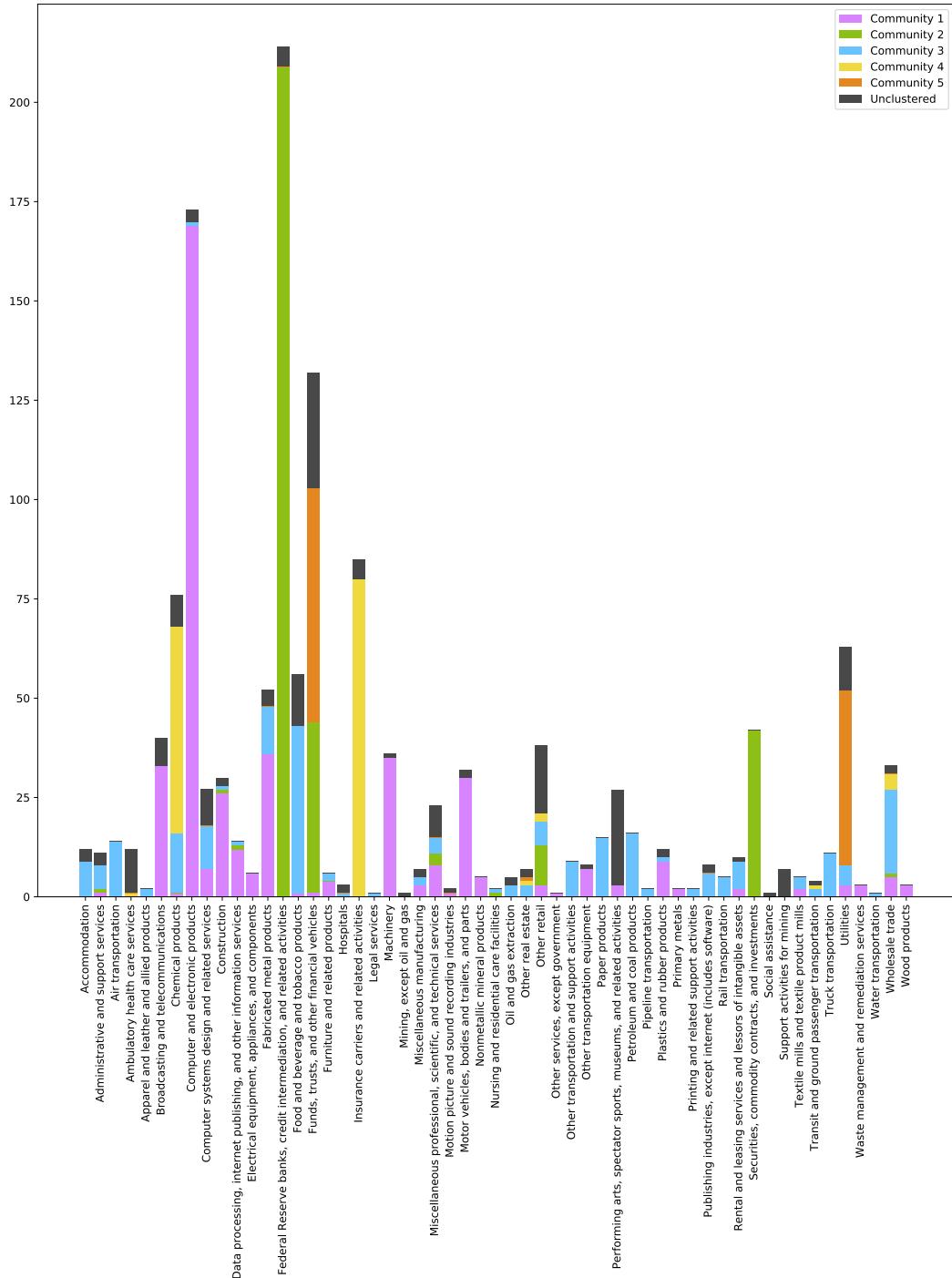


Figure 4.8: Count of stocks of each communities by industrial sectors. Sectors are arranged alphabetically.

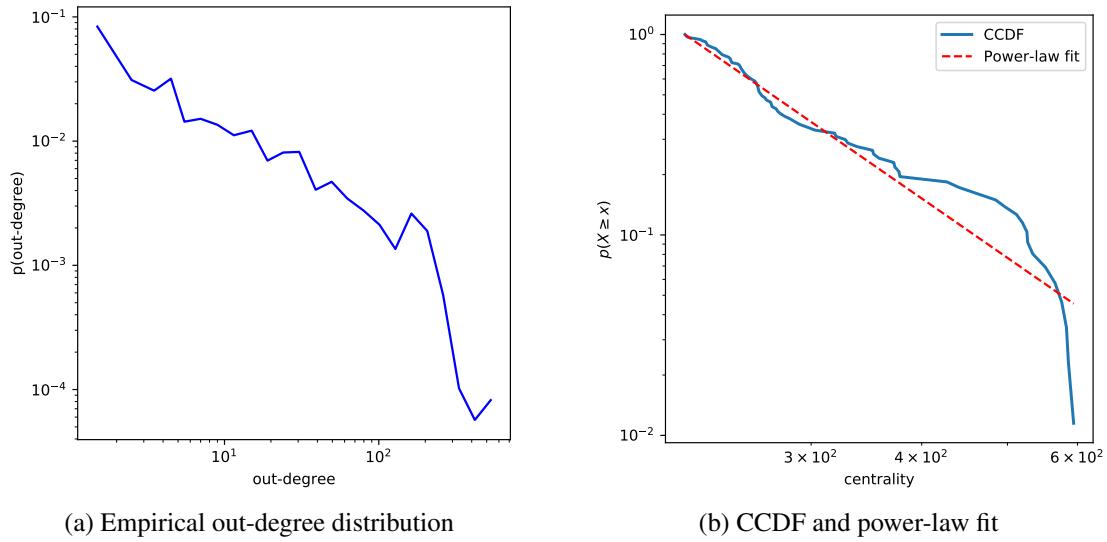


Figure 4.9: Out-degree distribution of stock network

4.4 Analysis of the directed weighted network

4.4.1 Analysis on the relationship between price return and clustering coefficient

4.5 Stability of network

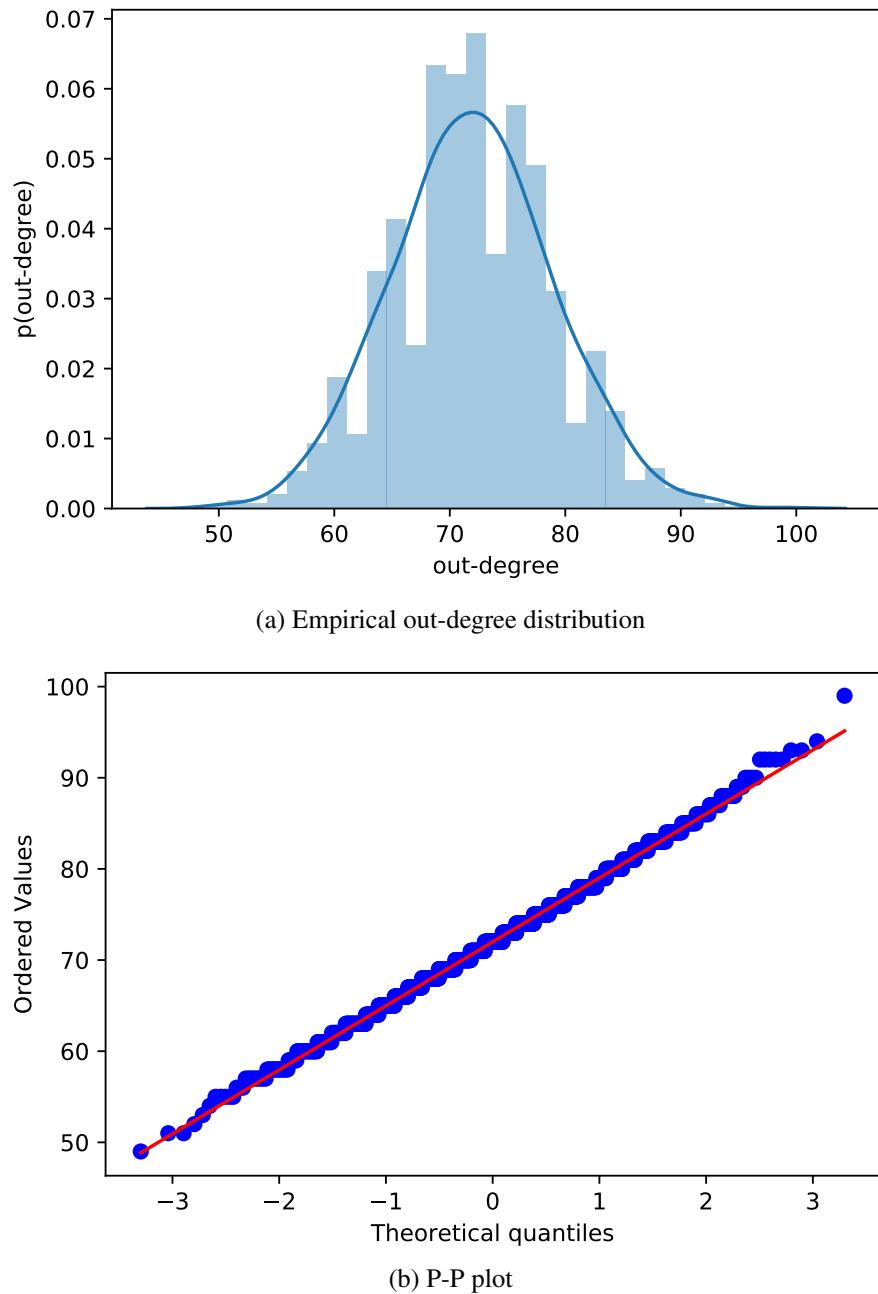


Figure 4.10: Out-degree distribution and P-P plot of small-world network

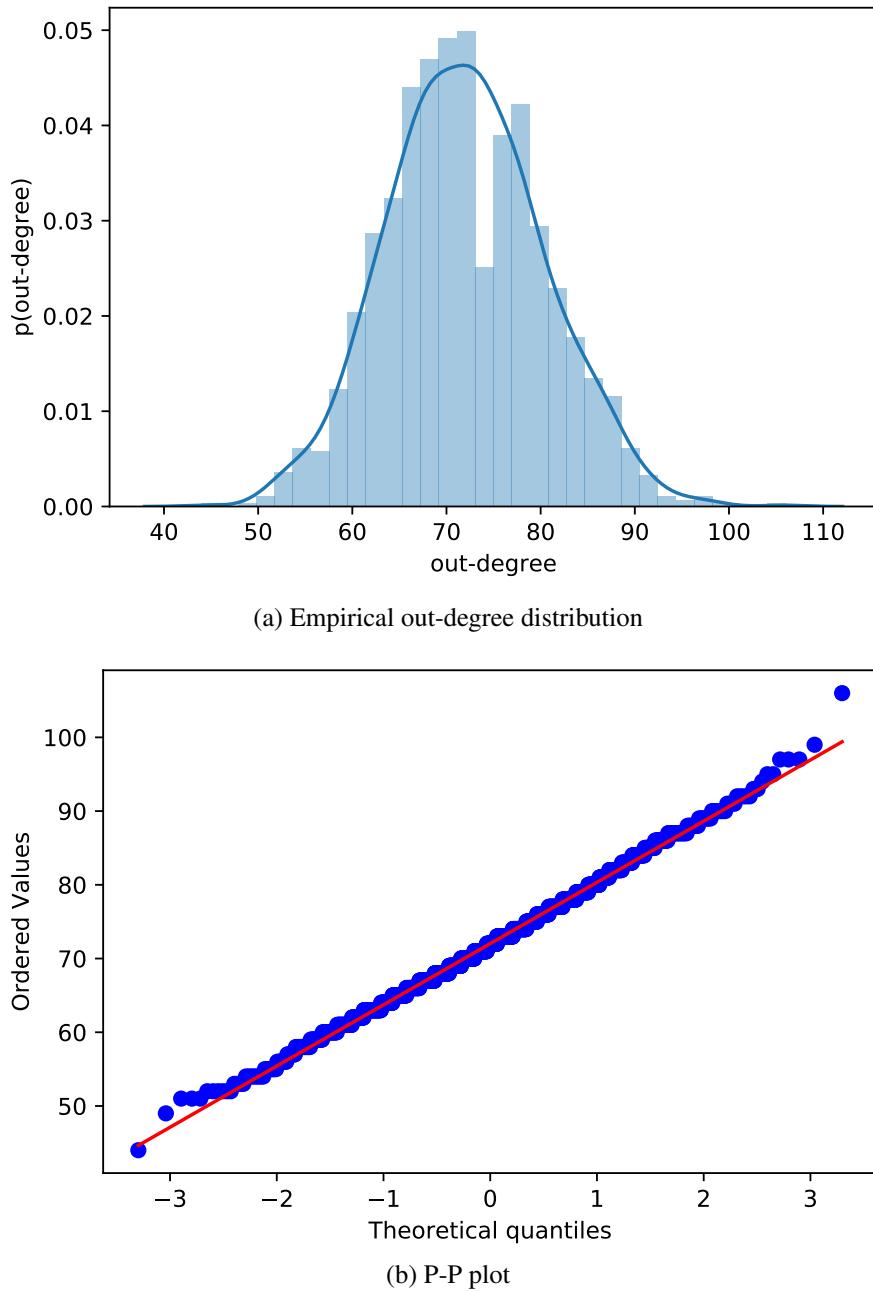
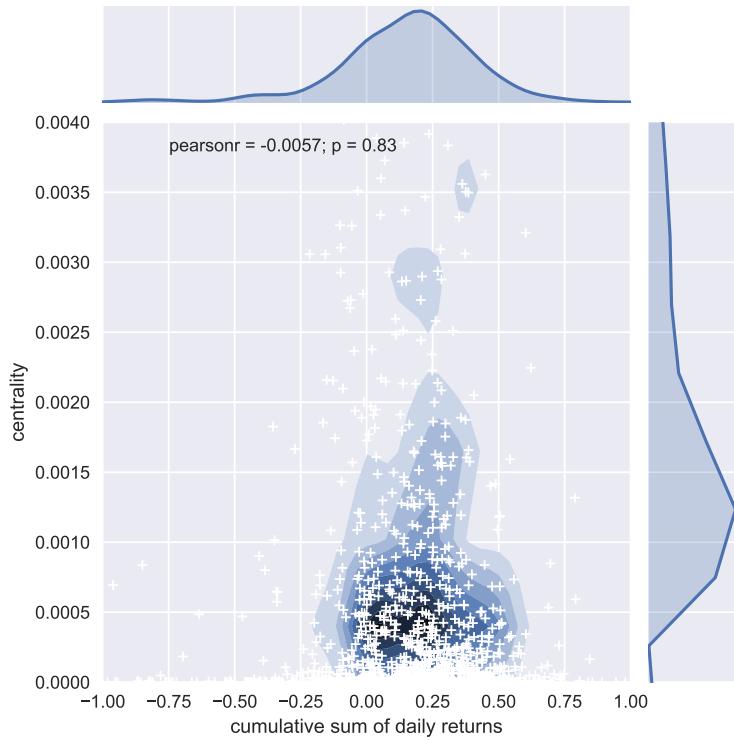
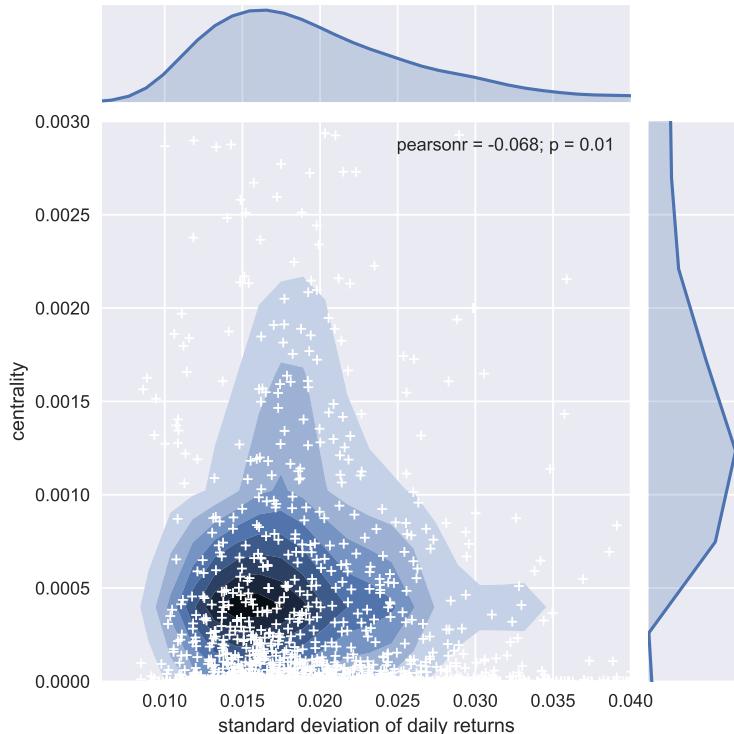


Figure 4.11: Out-degree distribution and P-P plot of random network



(a) Bivariate distribution of cumulative sum of daily return



(b) Bivariate distribution of standard deviation of daily return

Figure 4.12: Bivariate distribution

Chapter 5

Conclusions and future watch

Chapter 6

Making a bibliography

Whenever you wish to refer to books or articles relevant to your report you should use a citation such as [?]. You can also force entries to appear in the bibliography without a citation appearing in the document, by using \nocite.

Each document cited must have an entry in a .bib file. For this document we have only one, called `refs.bib`. These files are listed in the \bibliography command at the end of `report.tex`. Note that the .bib files can (and often do) contain many more entries than are actually cited in a particular document; the only ones that appear in the bibliography are those that have been referenced using \cite or \nocite.

In order to generate the appropriate reference entries, you will need to run `bibtex` after `latex` has been run, using the command `bibtex report`. This will generate a file `report.bbl`, which contains the bibliography entries. Once that file is there, you do not need to run `bibtex` again unless you add new citations, but you will probably have to run `latex` twice after running `bibtex` the first time.

The `TEX` FAQ ([Tea12]) gives tips on how to cite URLs.

The file `refs.bib` provides an example of what can be done with Bib`TEX`. You can find much more information in any book on `LATEX`, for example [HZY09, LN08, CK15, CLSW15]

Bibliography

- [CK15] Shauhrat S. Chopra and Vikas Khanna. Interconnectedness and interdependencies of critical infrastructures in the us economy: Implications for resilience. *Physica A: Statistical Mechanics and its Applications*, 436:865 – 877, 2015.
- [CLSW15] Kun Chen, Peng Luo, Bianxia Sun, and Huaiqing Wang. Which stocks are profitable? a network method to investigate the effects of network structure on stock returns. *Physica A: Statistical Mechanics and its Applications*, 436:224 – 235, 2015.
- [Dyk94] William Dyke. The nutter’s guide to L^AT_EX. *Read Only*, (7), May 1994.
- [HZY09] Wei-Qiang Huang, Xin-Tian Zhuang, and Shuang Yao. A network analysis of the chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388(14):2956–2964, 2009.
- [LN08] E. A. Leicht and M. E. J. Newman. Community structure in directed networks. *Phys. Rev. Lett.*, 100:118703, Mar 2008.
- [Tea12] L^AT_EX Team. URLs in BibTeX bibliographies. <http://www.tex.ac.uk/cgi-bin/texfaq2html?label=citeURL>, 2012.

Appendix A

Example of operation

An appendix is just like any other chapter, except that it comes after the appendix command in the master file.

One use of an appendix is to include an example of input to the system and the corresponding output.

One way to do this is to include, unformatted, an existing input file. You can do this using \verbatiminput. In this appendix we include a copy of the C file hello.c and its output file hello.out. If you use this facility you should make sure that the file which you input does not contain TAB characters, since L^AT_EX treats each TAB as a single space; you can use the Unix command expand (see manual page) to expand tabs into the appropriate number of spaces.

A.1 Example input and output

A.1.1 Input

(Actually, this isn't input, it's the source code, but it will do as an example)

```
/* Hello world program */

#include <stdio.h>

int main(void)
{
    printf("Hello World!\n") ;
    return 0 ;
}
```

```
}
```

A.1.2 Output

Hello World!

A.1.3 Another way to include code

You can also use the capabilities of the `listings` package to include sections of code, it does some keyword highlighting.

```
/* Hello world program */

#include <stdio.h>

int main(void)
{
    printf("Hello_World!\n");
    return 0;
}
```