Complex Network Of Interbank Market And Its Application In Neural Network

G14PJA = MATH4041

Mathematics 3rd Year Project

Autumn 2021/22

School of Mathematical Sciences

University of Nottingham

Zhiyuan Lei

Supervisor: Dr. Reuben Odea

Assessment type: Investigation

I have read and understood the School and University guidelines on plagiarism. I confirm that this work is my own, apart from the acknowledged references.

Abstract

The Interbank market is an essential part of the modern financial system. Complex interbank networks are formed by inter-bank lending, payment and settlement, discount and guarantee. On the one hand, banks' networks do provide channels for interbank exchanges. However, the exchange of the interbank liquidity assets also make banks in danger, the inter-bank lending behaviour become potential paths for risk contagion when defaults occur, and those defaults may cause a domino effect. Terefore, it is necessary to analyze the structure of the interbank network and mechanism of risk contagion. Consequently, I will concentrates on the characteristics of the structure of interbank network, the mechanism of interbank risk contagion and how to do the risk prediction in this report

Firstly, I build a dynamic network model for interbank market, which is based on the interbank-lending behaviors of the bank descirbed with balance sheet. Then, I analyze the features of interbank market networks. Simulation results indicate that the interbank network is a small-world network and has currency center structure in most of countries.

Secondly, I build the stochastic SIR model on interbank market networks to simulate risk contagion, and analyze the how characteristics of network influence the scale of banks which is exposed to risk. The simulation results show that exposure of bank assets increase when the lines of network increase or the average shortest path decrease.

Thirdly, I use graph neural networks (GNNs) to predict the label of banks in a specific graph based on the results of Monte-Carlo method. And then analyze how the structure of graphs influence the accuracy of prediction. Simulation results show that the accuracy of prediction is acceptable and accuracy increases when the lines of network increase or the average shortest path decrease.

Contents

1 Introduction

Inter-bank behavior forms complex relationships among banks through market transactions, payment and guarantees. On the one hand, Inter-bank lending among banks can spread risks and provide efficient cash flow. On the other hand, it also provides the way for risk exposure among banks. When one bank faces with the finacial crisis, it may not be able to pay creditor's debt and cause creditor banks to also face more difficulties, and its dominoes effect will cause collapse of interbank network.

Therefore, studying the network structure characteristics of the bank market is helpful for administrator to correctly understand the risk contagion path and propose relevant policies. Besides, studying the risk contagion mechanism of the bank market and designing effective risk contagion control strategies can help the managers of commercial banks. In addition, quik method predictng the risk of banks should be found to save time though the accuracy of this method is less than Monte-Carlo method.

1.1 Background

1.1.1 Current Reasearch On Interbank Network

Current reasearch on interbank network can be divided into two categories. One is the mechanism of constructing network based on Watts-Strogats graph. The other one is the characteristics analysis of exist networks. So far, scholars have conducted empirical research on the network structure of the interbank market in Japan, Austria, Italy, Brazil, Germany, the United Kingdom, the United States, Hungary, Belgium, Switzerland, Finland, Russia, Mexico, the Netherlands and other countries.

Firstly, Watts and Strogatts (1998) propse Watts-Srogats network and analyse its main characteristcs. Empirical research on the network structure of the interbank market finds that it has small-world and scale-free characteristics as Watts-Srogats network. Boss (2004) found out that the network degree of Austrian interbank market obeys a double power-law distribution, and the network has small-world characteristics. Kanno (2015) conducted an analysis of the Japanese interbank market and found that the network structure of the Japanese interbank market has small-world and scale-free characteristics. When Soramaki (2007) studied the interbank debt linkages in the Federal Reserve's electronic transfer payment system (fedwire), they found that the interbank market network in the United States has the characteristics of a small world network. Becher et al. (2008) found that the number of banks in the UK is far less than that in the US, but the average path length of the interbank capital network is similar to that of the US interbank network, and the UK interbank network is also a small-world network.

Secondly, interbank market network also has currency central structure. Lubloy (2005) analyzed the structure of the interbank market based on the transaction data of the interbank market composed of 39 banks in Hungary in 2003, and found that there are multiple currency center structures in the Hungarian interbank market. Most of banks in are connected to the several big banks. Similarly, there are multiple currency center structures in the Belgian interbank market (Degryse and Nguyen, 2004). The four major banks account for 85% of the total assets of all banks, and 35% of the interbank market transactions occur among these four major banks. And 90% of the inter-bank market transactions are related to the big 4 banks. Craig and Peter (2014) found that there is a currency center structure in the German interbank market.

In addition, the interbank market network has dynamically evolving. Iori et al. (2008) conducted a study on the Italian interbank market and found that the structure of the Italian interbank market was evolving year by year from 1999 to 2002, and the network structure of the Italian interbank market was a random network. Martiliez-Jaramill et al. (2010) studied the Mexican inter-bank capital network and found that most of the inter-bank debt links are stable, while a few, local links will change.

1.1.2 Risk Infection Mosel

The existing research on the relationship between the network structure of the bank market and risk contagion is mainly divided into two categories: one is the analysis and comparison of network stability between different network structures (such as small-world network, scale-free network and random network, etc.). The other one is to analyze the relationship between network topology characteristics and network stability on the network with the same structure.

Firstly, based on the classic epidemic model such as SI, SIS and SIR model, scholars built their ODE form on social network. Stefano (2016) built a new stochatic infection model on social network, and found out that the stochatic infection rate will converge to its ODE form. Based on the epidemic model, Scholars found that the inter-bank lending relationship provides a potential path for inter-bank market risk contagion (Gai and Kapadia. 2010. (Krause and Giansante, 2012). Ladley (2013) revealed that the conditions for inter-bank lending to spread risks provide ways for risk contagion.

Based on network theory, Aleksiejuk and Holyst (2001) analyzed the contagion and scale of bank default when a single bank failed on the bank network. Moreover, Nier et al. (2007) studied the systemic risk caused by default in the interbank market based on random network;. Thumer et al. (2003) firstly applied SI model on reasearcing the contagion and scale of the bank default.

Allen and Gale (2000) studied the risk contagion problem in the interbank market based on different assumptions about the structure of the interbank market. The results of the Allen-Gale model show that the risk contagion among banks is an equilibrium phenomenon. When the inter-bank market is a complete market structure, the system can achieve optimal risk sharing, under an incomplete market structure, the system can also achieve optimal risk sharing, but the market structure is more fragile. However, Cassar et al. (2001) assumed that the network structure of the interbank market is a local network and a global network, respectively. The study found that when the interbank market network is locally connected, the speed of bank risk contagion is relatively low, when the inter-bank market is globally connected, the speed of bank risk propagation is relatively high.

The network characteristics (such as network centralization, network connectivity, average shortest path length and network clustering coefficient, etc.) will also influence the scale of bank default. Krause and Giansante (2012) simulated interbank networks with different scaling parameters, and their research believed that the smaller the interbank network scale, the higher the concentration, the less likely the default contagion would occur. Battiston (2012) analyzed the impact of network density on bank failure in individual banks and in interbank market networks. Tabak et al. (2014) proved that the clustering coefficient can be used as a measure of systemic risk, and that the clustering coefficient is negatively related to interest rates.

1.1.3 Classic Prediction Method: Monte Carlo And Neural Network

When analyze a specific graph structure make some banks more likily to expose to risk, Monte-Carlo method must be a effective method because frequency can represents probablity well when we run the model for many times. However, Monte-Carlo always wastes lots of time when network is large. Therefore, it is naturally to consider use neural network to do the prediction if the accuracy is acceptable. Bryan et al. (2014) gives a method to transfer the information of graph to data embedding vector named DeepWalk based on the language model Word2Vec (Kyunghyun et al., 2014). Jizhe Wang in Alibaba (2018) gives another method for graph embedding named EGES, their results shows that this method is more accurate and runs more swift than DeepWalk. Thomas and Max (2017) built graph convolutional networks (GCNs) for classification and shows that GCNs work well on Zacharys karate club network calssification.

1.2 Structure Of The Report

The current research have a lot of great results on interbank network and risk model. However, there are still some further works to do, which are shown as follows:

1)Most of the inter-bank market network structure constructed by the existing simulation models is static, and most of them simply assume that the inter-bank market network is a specific structure, such as random network, scale-free network, and so on. However, the actual inter-bank market network is very complex, and the inter-bank relationship depends on the inter-bank lending behavior of banks, and the behavior of banks is constantly changing with time. Therefore, the actual network structure of the inter-bank market should also renew dynamically.

2)Most of risk infection model applied on network is their ODE form. Few of them applied stochastic model on inter-bank network. However, the stochastic model performs better when describes single bank's infection rate.

3)Few of current model research how the graph neural network works on inter-bank network, but it is a effective and efficient way todo the prediction and we will not train a new model in the future when analyse a new graph structure.

1.2.1 Content Of Report

In this paper, I will do some further work of interbank network and risk model, the main part of my work is shown as follows

(1) Construction of dynamic interbank market network:

I build a dynamic bank balance sheet, and then build a dynamic interbank market network based on the dynamic bank balance sheet especially inter-bank lending. Apply the model to simulate interbank network in many countries, and analyze the characteristics of inter-bank market network structure.

(2) Construction of stochastic rate model on interbank network:

Based on the interbank market network, I build the stochastic SIR model on networks. Apply the model to simulate its results in many different graphs, and analyze how the characteristics of graph influence the scale of banks easily exposed to risk.

(3) Prediction of banks which is in danger on a specific graph structure:

I use Monte-Carlo method to train many graphs to label the banks on a specific graph. They are labeled with 3 categories: Easily get infected, recovered and susceptible. Then I treat the results as the sample of graph neural network, and train the sample to get the prediction model. At last, I analyze how the characteristics of inter-bank market network influence the accuracy of prediction.

1.2.2 Structure Of Dissertation

In section 1, I introduce the backgroud and current research of interbank network and risk model, and analyze the further work of current research. Then introduce the content of the report.

In section 2, I introduce some basic knowledge of graph including the characteristics of a graph such as average clustering coefficient, averag shortest path and centrality e.t.c. Then I introduce some essential network such as ER graph and Watts-Strogatts graph, and analyze their characteristics.

In section 3, I build dynamic interbank network in many countries based on dynamic bank balance sheet. Then I built stochastic SIR model on interbank market network.

In section 4, I use Monte-Carlo method to label the banks in many graphs. Then I use those banks' label to train a specific graph neural network for prediction. In addition, I introduce the graph embedding method in this section.

In section 5, I summarize the conclusion of all section and give some shortcomings and further work of this report.

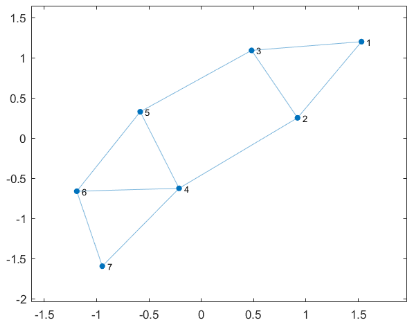
2 Basic knowledge of Interbank market network

2.1 Complex network

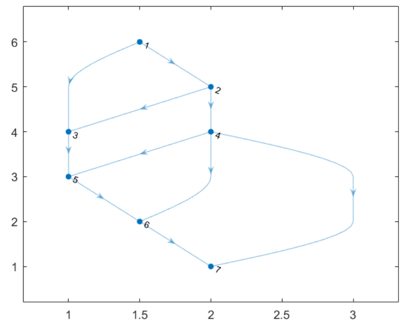
2.1.1 Directed and Undirected Graphs

Since Complex Networks are special Graphs, graph theory can be used in decribe interbank market network (Wasserman and Faust, 1994. Scott, 2012. Cormen, 2001). Basically, a graph $G = (V, E)$ is a mathematical structure consisting of a nodes set $V = \{v_1, v_2, \cdots, v_N\}$ and an edges set $E = \{e_1, e_2, \cdots, e_K\}$. The number of nodes N $= |$ V $|$ and the number of edges K$= |$ E $|$ are called the order and size of the graph G, respectively.

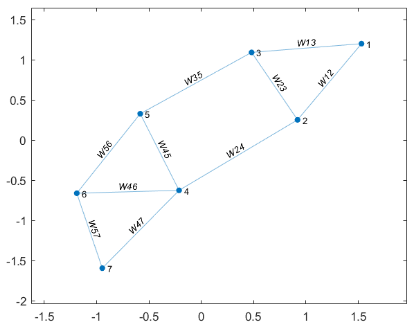
A graph is called undirected graph if there is no ordering in the vertices defining an edge. A graph G for which each edge in E has an ordering to its nodes is called a directed graph. Obviously, in undirected graph $e_{ij} = e_{ji}, (i \neq j)$, but in directed graph $e_{ij} \neq e_{ji}, (i \neq j)$. Additionally, If edge E has weight, then the graph is also called weighted graph.



(a) Undirected Graph



(b) directed Graph



(c) Weighted Graph

Figure 1: Directed and Undirected Graphs in one example, $N = 7$, $K = 10$

Figure 1 shows a graph G with $N = |V| = 7$ nodes and $K = |E| = 10$ edges, where Figure 1(a) is undirected graph because there is no ordering in the vertices defining an edge, and 1(b) is directed graph, you can see an edge always comes from a node and points to another node (i.e. from node 1 to 2). 1(c) is weighted graph with edge weight W_{ij} between node v_i, v_j.

As for all the graph G, it always holds that $0 \leq K \leq N(N-1)/2$. When each node in G has connection with all other nodes, the number of edges come from this node is ()

A graph G is called sparse graph if $K << N^2$. And graph G is called dense graph when $K = O(N^2)$. If $K = N(N-1)/2$, graph G is a complete graph. Actually, the interbank market network is formed as dense graph, this will be introduces in section 2.3.2 and section 3.2.3.

Generally, graph G can be represented by its adjacency matrix $A = (a_{ij})_{N \times N}$. For example, matrix A_a is adjacency matrix of Figure 1(a) and 1(c), matrix A_b is adjacency matrix of Figure 1(b).

$$A_a = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, A_b = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $a_{ij} = 1$ means node v_i, v_j is connected while $a_{ij} = 0$ means node v_i, v_j is not connected.

2.1.2 Degree and its distribution

In undirected graph G, the degree of node v_i is the number of links that are connected to v_i, thus it can be defined as

$$k_i = \sum_j a_{ij}, \tag{2.1}$$

where adjacency matrix $A = (a_{ij})_{N \times N}$.

In directed graph, The indegree is number of edges going into node v_i and the outdegree is the number of edges going out of v_i. In-degree is defined as

$$k_i^{in} = \sum_j a_{ji},$$

and out-degree is

$$k_i^{out} = \sum_j a_{ij},$$

and the degree of v_i is $k_i = k_i^{\text{in}} + k_i^{\text{out}}$

Now we define the moment of degree as follow

$$\langle k^n \rangle = \sum_k k^n P(k). \tag{2.2}$$

When $n = 1$, equation (2.2) is the average degree of graph G.

In graph G, $P(k' \mid k)$ is the probablity that a node with degree k points to a node with degree k'. Obviously, $\sum_{k'} P(k' \mid k) = 1$, and it also satisfies (Boguna and Pastor-Satorras, 2002)

$$kP(k' \mid k) P(k) = k' P(k \mid k') P(k'), \tag{2.3}$$

where $P(k)$ is the probablity that a node has degree k.

In large graph, it is difficult to calculate $P(k' \mid k)$. Generally, we calculate average degree which are connected to v_i instead, which is the eqution (2.4)

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N_i} k_j = \frac{1}{k_i} \sum_{j=1}^{N} a_{ij} k_j, \tag{2.4}$$

where N_i is the set of node v_i's neighbor nodes.

In Random graphs, assume that the probablity that node v_i and v_j is connected is p. Then the degree of node v_i always follows the Bernoulli distribution, which is

$$P(k_i = k) = C_{N-1}^{k} p^k (1-p)^{N-1-k}. \tag{2.5}$$

Define X_k is the number of nodes whose degree is k, then

$$E\left(X_k\right) = NP\left(k_i = k\right) := \lambda_k.$$

When N is very large, X_k follows Poisson distribution, so we have

$$P\left(X_k = r\right) = e^{-\lambda_k}\frac{\lambda_k^r}{r!}. \tag{2.6}$$

Meanwhile

$$\begin{aligned} E\left(X_k\right) &= \operatorname{Var}\left(X_k\right) = \lambda_k, \\ P\left(k_i = k\right) &= C_{N-1}^k p^k (1-p)^{N-1-k}. \end{aligned} \tag{2.7}$$

When $N \to +\infty$

$$P\left(k_i = k\right) = e^{-pN}\frac{(pN)^k}{k!} = e^{-\langle k\rangle}\frac{\langle k\rangle^k}{k!}. \tag{2.8}$$

We set N=100 and p=0.15, the density curve of degree is shown in Figure 2. As it shows, the degree is basically remains in [5,30] in this example.

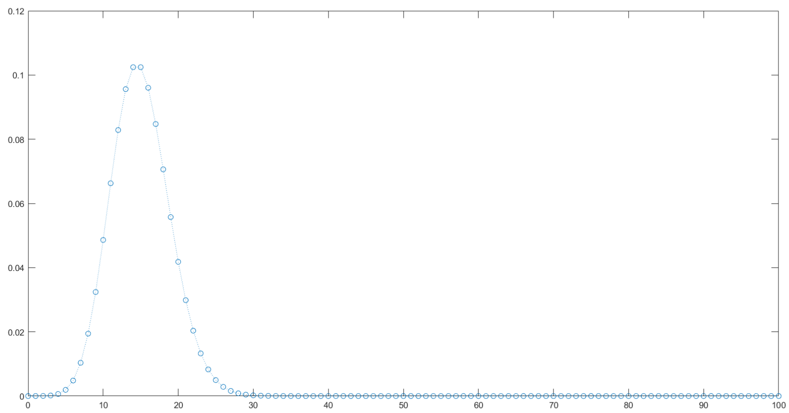


Figure 2: Density function: N=100, p=0.15

Now we get the distribution of degree. This is the fundation of building both random graph and Watts-Strongats network which are introduced in section 2.3.

2.2 Characteristics Of Graph

In this section, I will introduce some important characteristics of a graph. Those characteristics will influence how we build the interbank market network and how the model works in graphs.

2.2.1 Local cluster coefficient

Local clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together. Precisely, it is given by a proportion of the number of links between the vertices within its neighbourhood divided by the number of links that could possibly exist between them (Watts and Strogatz, 1998). As for node v_i, its cluster coefficient c_i is given by

$$c_i = \frac{2e_i}{k_i\left(k_i-1\right)} = \frac{\sum_{j,m} a_{ij}a_{jm}a_{mi}}{k_i\left(k_i-1\right)}, \tag{2.9}$$

where $A = \left(a_{ij}\right)_{N\times N}$ is the adjacency matrix.

And the network average clustering coefficient is given by

$$C = \langle c \rangle = \frac{1}{N}\sum_{i\in N} c_i. \tag{2.10}$$

2.2.2 Network Density

Network density describes the portion of the potential connections in a network that are actual connections, which is given by

$$D = \frac{|E|}{N(N-1)}. \tag{2.11}$$

2.2.3 Shortest path

In a undirected graph, a subset $Pa = v_i, v_k, v_m, \ldots, v_j$ in which v_i is connected to v_k, v_k is connected to v_m e.t.c, rewrite Pa as $Pa = v_1, v_2, v_3, \ldots, v_n, n \leq N$, then v_k is connected to v_{k+1} for all $1 \leq k < n$, then Pa is called a path with length n-1 from v_1 to v_n. The shortest path from v to u is path $Pa = v_1, v_2, v_3, \ldots, v_n, v = v_1, u = v_n$ that over all possible n minimizes the sum

$$\sum_{i=1}^{n-1} W_{i,i+1}$$

, where $W_{i,i+1}$ is weight between node v_i and v_{i+1}. It equals to 1 in undirected graph.

The average shortest path of a graph is defined as follow

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij}, \tag{2.12}$$

where $d_{ij} = d(v_i, v_j)$ is the shortest path between v_i and v_j

The diameter of a graph is the length of the shortest path between the most distanced nodes.

Figure 3 shows a graph with 40 nodes in total. The shortest path from 6 to 36 is highlighted by red line and one of the diameter (From 1 to 28) is highlighted by green line.

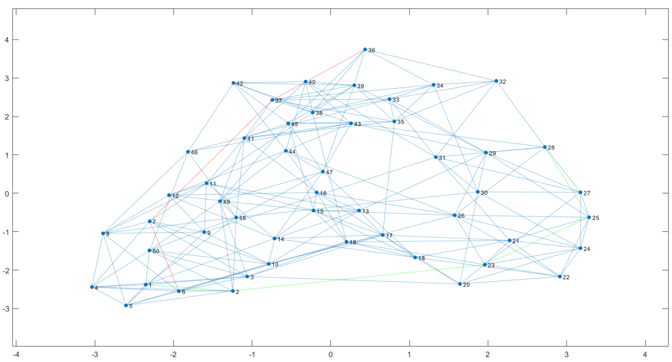


Figure 3: Example: Shortest path from 6 to 36 and Diameter

Shortest path is a very important characteristics in interbank market network because it directly influence the length that the risk is spreaded from one bank to another. We will discuss it further in section 4.1.

2.2.4 Centrality

The degree centrality of node v_i is defined as

$$cd_i = \frac{k_i}{N-1}. \tag{2.13}$$

Graph centralization describes how tightly the nodes connect with each other. Freeman(1996) defined the graph centralization as below

$$cd = \frac{\sum_{i=1}^{n} (k_{i^*} - k_i)}{(N-1)(N-2)}, \tag{2.14}$$

where i^* is the node with highest degree in graph g.

Closeness centrality of a node is the average length of the shortest path between the node and all other nodes. It is defined as below

$$C(v_j) = \frac{N-1}{\sum_i d(v_i, v_j)}, \tag{2.15}$$

where $d(v_i, v_j)$ is the distance between vertices v_i and v_j.

Betweenness centrality is the number of times the shortest path walk through a node between two other nodes. It can be represented as

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{nd_{st}(v)}{nd_{st}}, \tag{2.16}$$

where nd is total number of shortest paths from node s to node t and $nd(v$ is the number of these paths that pass through node v.

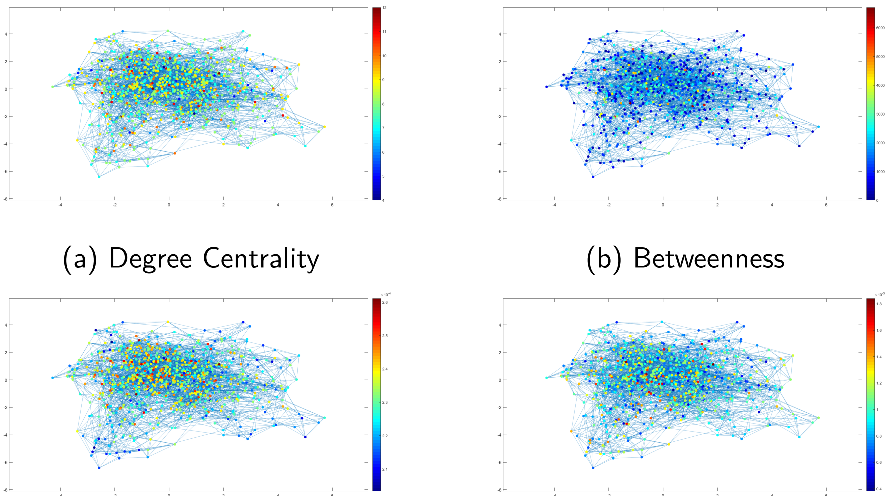
Eigenvector centrality is a measure of how important a node is in graph g. Let A is adjacency matrix of graph G, then the Eigenvector centrality vector $X = (x_1, x_2, \ldots, x_N)$ is given by

$$\mathbf{Ax} = \lambda \mathbf{x} \tag{2.17}$$

Equation (2.17) is the eigenvector equation of matrix A.

Those 4 centrality all describe some information of each node and how important it is in the graph. However, none of them describe the whole information of the nodes. Is there any method can record nodes information in a big matrix? The method solving this problem is called Graph Embedding which will be introduced in section 4.2.1.

Figure 4 shows four different centrality of a same graph with 500 nodes. The nodes with large centrality will be colored by dark color and they tend to be the 'bridges' among the connections. Thus, those nodes often are the currency center in interbank network and some of them is more likely to be dangerous in finacial crisis because the risk tend to pass through those 'bridges' during the crisis. This will be discussed in section 4.1.



(a) Degree Centrality

(b) Betweenness

(c) Clossness

(d) Eigenvector

Figure 4: Examples of centrality of the same graph: N=200, p=0.4

2.3 Graph Model

2.3.1 Random Graph

Erdos and Renyi (1959) created a graph with N nodes and K edges model, which is called ER Random Graph. There are two methods to creat a ER graph.

1)Randomly select a pair of nodes and connect them. Repeat the step over and over again until the number of edges equals to K. This model is called $G^{ER}_{N,K}$.

2)The probablity of two nodes connected to each other is p. Then the probablity that the graph has K edges is $p^K(1-p)^{N(N-1)/2-K}$. This model is called $G^{ER}_{N,p}$.

When $N \to \infty$, the average degree of $G^{ER}_{N,K}$ is $\langle k \rangle = 2K/N$ (see equation (2.2)) and the average degree of $G^{ER}_{N,p}$ is $\langle k \rangle = p(N-1)$. Newmann(2001) gives the formula that computes average shortest path, which is given by

$$L = \frac{\ln(N/z_1)}{\ln(z_2/z_1)} + 1. \tag{2.18}$$

In ER model, $z_1 = \langle k \rangle, z_2 = \langle k \rangle^2$.

And the cluster coefficient of the graph is given by

$$C = \frac{\langle k \rangle}{N} \left[\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \right]^2. \tag{2.19}$$

2.3.2 Watts-Strogats Network

Watts and Strogats bulit a new network which is similar to the ER model, but it always has higher cluster coefficient. The basic algorithm of generating this network is shown below:

Given the N nodes, the mean degree K (assumed to be an even integer), and the probablity of rewire p, where $N \gg K \gg \ln N \gg 1$. Then the model constructs an undirected graph as follow

1)Construct a regular ring lattice, a graph with N nodes each connected to K neighbors ($K/2$ on each side). There is an edge (i,j) if and only if

$$0 < |i-j| \bmod \left(N - 1 - \frac{K}{2}\right) \leq \frac{K}{2}.$$

2)For every node $V_i = 1, \ldots, N$ take every edge connecting i to its $K/2$ rightmost neighbors, that is every edge (i,$jmodN$) with all j that $i < j \leq i + K/2$.

3)Rewire the nodes with probablity p, which means every edge (i,$jmodN$) is replaced by (i,m) where m is chosen uniformly at random from all possible nodes and $m \neq i$.

Albert and Barabasi (2002) shows the structure of Watts-Strogats network with different p, which is shown in Figure 5.

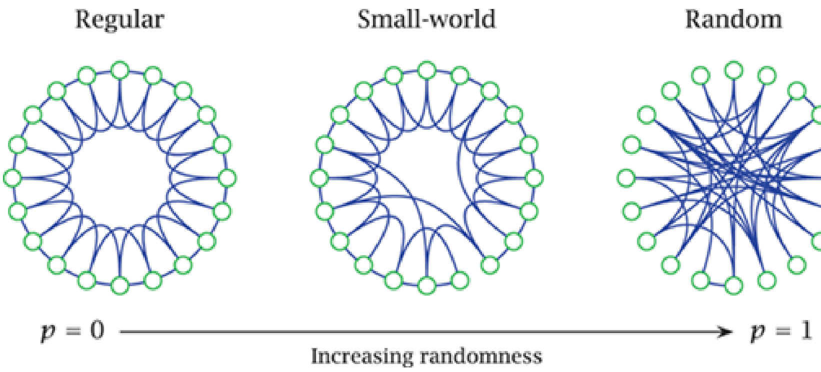


Figure 5: Example: Watts-Stongats network, Albert and Barabasi, 2002

If we draw the scatter figure of average cluster coefficient C(p) to see how it is influenced by probablity p. The result is shown below

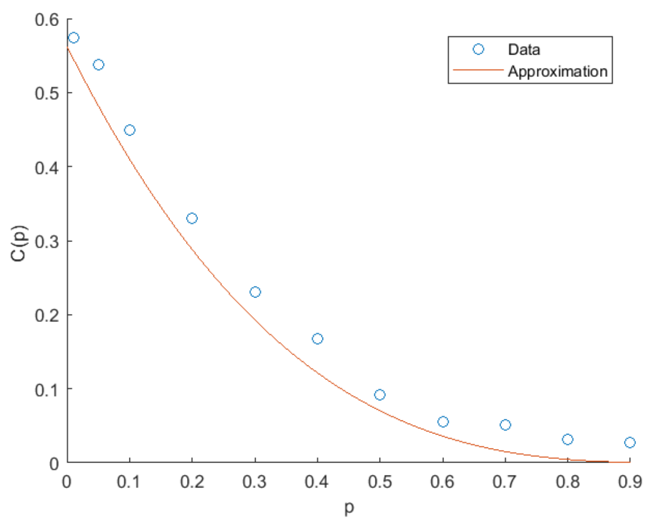


Figure 6: Example: Average cluster coefficient C(p) against p. N=200, K=5.

By equation(2.9) we have cluster coefficient of node i is given by $c_i = \frac{2e_i}{k_i(k_i-1)} = \frac{\sum_{j,m} a_{ij}a_{jm}a_{mi}}{k_i(k_i-1)}$, We use the expectation of e_i and $k_i(k_i-1)$ to represent average cluster coefficient $C(p; N, K)$. It is obvious that expectation of k_i is mean of degree: K. Since the node i is not rewired with probablity $1-p$, and probablity that the two neighbor nodes of i before rewiring are still the neighbor nodes after rewiring is $(1-p)^2$, the expectation of e_i is given by

$$e_i = M_0(1-p)^3 + O(1/N), \tag{2.20}$$

where M_0 is the number of edges between K (mean of degree of node i) neighbor nodes. It is given by $M_0 = 3K(K-2)/8$.

Then C(p) can be approximated as $\tilde{C}(p)$, which is given by

$$\begin{aligned}\tilde{C}(p) &\triangleq \frac{M_0(1-p)^3 + O(1/N)}{K(K-1)/2} \\ &= \frac{3K(K-2)/8}{K(K-1)/2}(1-p)^3 \\ &= \frac{3(K-2)}{4(K-1)}(1-p)^3 + O(1/N).\end{aligned} \tag{2.21}$$

The comparison between appoximation $\tilde{C}(p)$ and real Data see Figure 6.

As for regular network (p=0), equation(2.21), average cluster coefficient C(p) of W-S graph is given by

$$C(0) = \frac{3(K-2)}{4(K-1)} \approx \frac{3}{4}, K \to \infty. \tag{2.22}$$

We draw the scatter figure of average shortest path of W-S network against p to get their relationship, the result is shown below

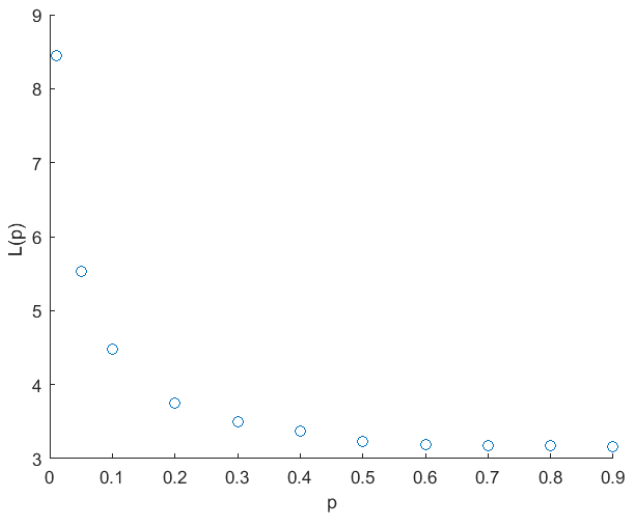


Figure 7: Example: Average shortest path L(p) against p. N=200, K=5.

Newman and Watts (1999), Barrat and Weigt (2000) introduce the average shortest path $L(p)$ can be approximated as fpllpws:

$$L(p) = \frac{N}{K} f(NKp), \tag{2.23}$$

where $f(x)$ can be approximated by

$$f(x) = \frac{2}{\sqrt{x^2+4x}} \operatorname{arctanh} \sqrt{\frac{x}{x+4}}. \tag{2.24}$$

Since

$$\operatorname{artanh} u = \frac{1}{2} \ln \frac{1+u}{1-u}, \tag{2.25}$$

then

$$f(x) \simeq \frac{1}{\sqrt{x^2+4x}} \ln \frac{\sqrt{1+4/x}+1}{\sqrt{1+4/x}-1} \simeq \frac{\ln x}{x}, \quad x >> 1. \tag{2.26}$$

By equation (2.21) and equation (2.24), we get

$$L = \frac{\ln(NKp)}{K^2 p}, \quad NKp >> 1. \tag{2.27}$$

Thus when $p = 0$ and 1, the value of $L(p)$ is

$$L(0) \approx N/2K >> 1, L(1) \approx \frac{\ln N}{\ln K}. \tag{2.28}$$

As algorithm of W-S network shows, each node is still connected to at least K/2 original nodes, that is, the degree of each node is at least K/2. Thus we rewrite the degree as

$$c_i = s_i + \frac{K}{2}$$

Moreover, s_i can be represented by

$$s_i = s_i^{(1)} + s_i^{(2)},$$

where $s_i^{(1)}$ is the initial neighbor nodes of i that do not rewired, and the probablity of not rewiring is $1-p$. And $s_i^{(2)}$ is neighbor nodes of i that rewired.

Then for nodes i whose degree is bigger than $K/2$, and $\mathrm{s_i^{(1)}} \in [0, \min(\mathrm{k} - \mathrm{K}/2,\ \mathrm{K}/2)]$, we have

$$P_1\left(s_i^{(1)}\right) = \begin{pmatrix} K/2 \\ s_i^{(1)} \end{pmatrix} (1-p)^{s_i^{(1)}} p^{K/2-s_i^{(1)}}, \tag{2.29}$$

and when N is very large

$$P_2\left(s_i^{(2)}\right) \simeq \frac{(pK/2)^{s_i^{(2)}}}{\left(s_i^{(2)}\right)!} \mathrm{e}^{-pK/2}. \tag{2.30}$$

Consequently, we have

$$P(k) = \sum_{n=0}^{\min(k-K/2,K/2)} \begin{pmatrix} K/2 \\ n \end{pmatrix} (1-p)^n p^{K/2-n} \frac{(pK/2)^{k-(K/2)-n}}{(k-(K/2)-n)!} \mathrm{e}^{-p/2}. \tag{2.31}$$

We set N=500, K=6, and draw the scatter figure of degree distribution with different rewire probablity p and compare them with Poisson distribution with parameter K. The result is shown in Figure 8

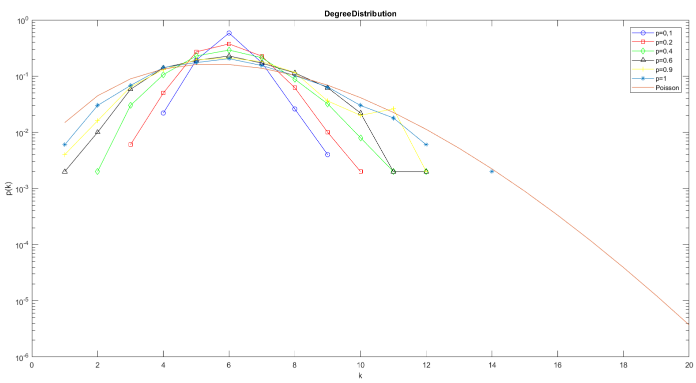


Figure 8: Example: Degree Distribution with different p, N=500, K=6

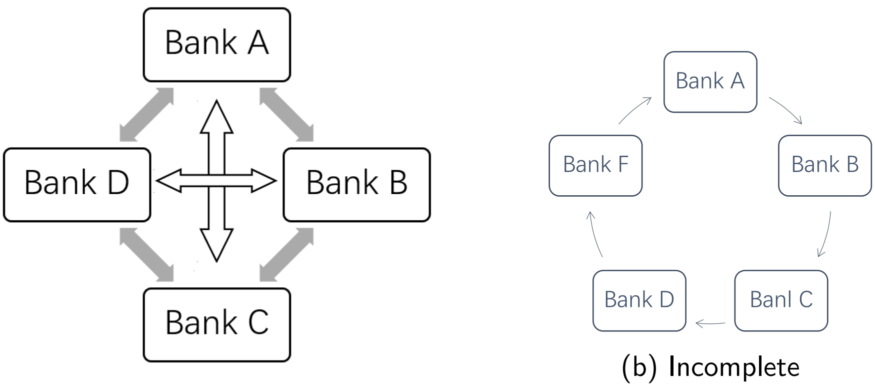
As shown in Figure 8, when p increases, the distribution of degree beocomes more close to its expectation. This will influence the average cluster coefficient and average shoretest path of the graph (see Figure 6 and 7). In next section, We will find out that interbank network also has small-world characteristics as Watts-Strogats network shows.

3 The Model Of Interbank Risk Market

The interbank market is an important part of the modern financial system. Inter-banks have formed a complex relationship of creditor's rights and debts through lending, payment and settlement, discounting, acceptance, guarantee and other forms. Actually, it can be seen that the network structure of the interbank market has many complex characteristics such as small world, scale-free, hierarchical structure, currency center structure, cluster structure, dynamic evolution and so on. On the basis of the existing research, this chapter constructs a balance sheet-based interbank market network model, explains the formation mechanism of the interbank market network dynamics, which is conducive to a deeper understanding of the formation process of the interbank market, and also provides insights into the formation process of the interbank market. It lays a foundation for analyzing the risk contagion mechanism of the inter-bank market and designing risk control strategies.

3.1 Relationship Among Banks

Allen and Gale (2002) assume that the market network structure of banks has two forms. One is complete market structure (as shown in Figure 9(a)) and the other one is incomplete market structure (as shown in Figure 9(b))



(b) Incomplete

(a) Complete

Figure 9: Two forms of bank structure

Obviously, complete market is a special underected graph where any two nodes have a bidirectional edge connection, while incomplete market is a directed graph in which there are only directed connections between banks.

Moreover, Freixas (2000) proposed currency center structure (as shown in Figure 10). In this structure, there is a undirected connection between the currency center (Bank A) and other banks, while there is no connection between other banks.

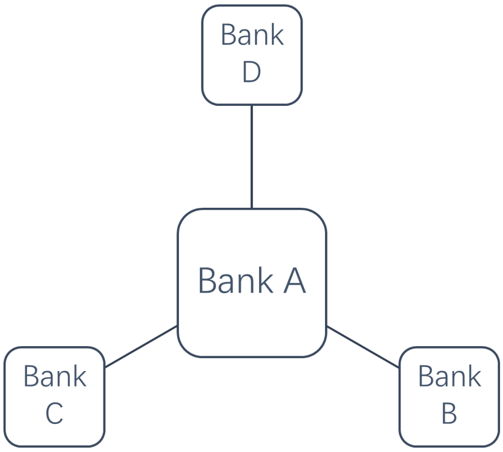


Figure 10: Currency center structure

In addition, the interbank market structure has complex network characteristics, such as small-world characteristics, scale-free characteristics, hierarchical structure and currency center structure characteristics, cluster structure. Basically, Watts-Strongats network works well in simmulating the relationship among banks, one of the example is shown in Figure 11.

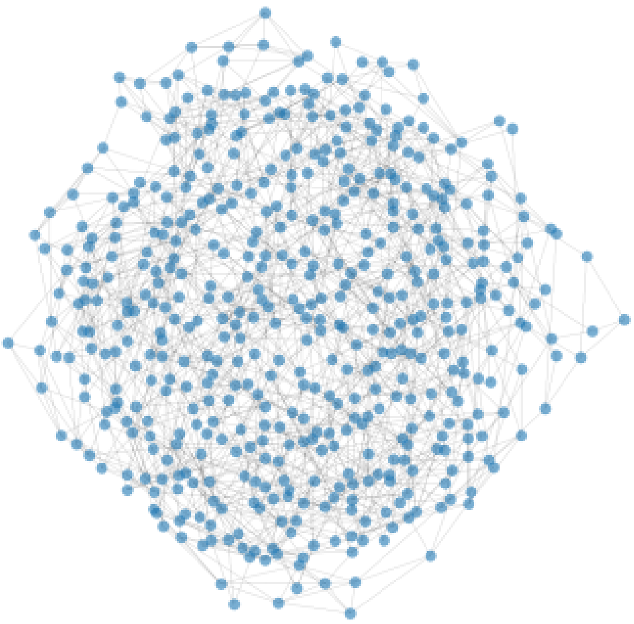


Figure 11: Example: W-S Network, N=500, K=4, p=0.2

However, Random graph does not fit other characteristics such as currency center structure and incomplete market though it do represents some key features of interbank market. For example, Through analysing 39 banks in Hungary, Lublóy (2005) found out that there were several currency center in Hungarian bank market. Figure 12 shows his result

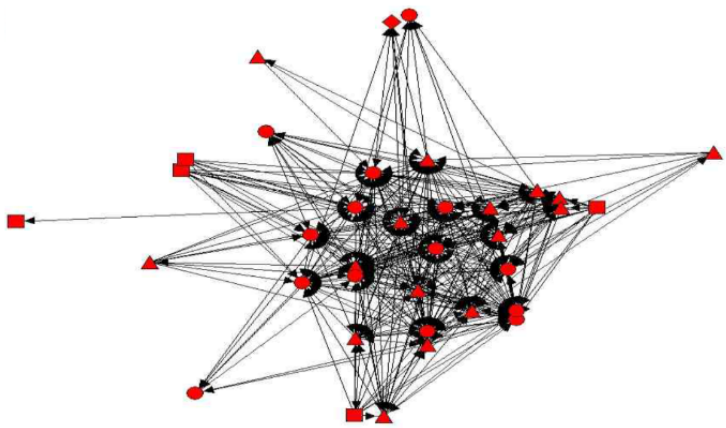


Figure 12: Hungarian Market: Several currency center, built by Lublóy in 2005

Through the examples above, it can be found that the interbank market network constructed by the existing simulation models is static, and most of them simply assume that the interbank market network is a specific structure, such as random network, scale-free network, etc. However, the actual inter-bank market network is highly complex. The inter-bank creditor-debt relationship depends on the behavior of banks, and the behavior of banks is constantly changing over time. Therefore, the actual network structure of the inter-bank market should also evolve dynamically. Therefore, the existing inter-bank market network structure model cannot accurately reflect the formation mechanism of the inter-bank market network. In view of this, this chapter firstly constructs the bank's dynamic balance sheet; secondly, it designs the interbank connection probability based on the bank lending scale and credit lending risk appetite, and finally constructs the interbank market dynamic network.

3.2 Dynamic Network Of Interbank Market

Simplicitly, I assume that weighted mixed graph (directed and undirected connection both exists) $G = (V, E, X)$ is the Interbank market network, in which bank i and bank j have connection if and only if they have Inter-bank Lending relationship. V is the set of banks and E is the set of Inter-bank Lending relationship. If bank i is not the currency center, then bank i is creditor and bank j holds the debt, which means the directed edge e_{ij} is from i to j. On the contrary, there is a undirected edge between i and j if i is a current center. Matrix $X = (x_{ij})_{N\times N}$ is the amount of Inter-bank Lending, it can be regarded as the weight of edge between i and j. Then we build the model based on the Inter-bank Lending and Balance sheet as follows

3.2.1 Balance Sheet

The balance sheet of a bank is very complex. In this section, I just choose a very simple form to discuss. Assume that asset of bank j includes external investment I_j, interbank borrowing L_j and current assets M_j, debt of bank j includes deposit D_j, interbank lending B_j and owners' equity E_j. By accounting equation (David Romer, 2009), we have

$$I_j + L_j + M_j = D_j + B_j + E_j. \tag{3.1}$$

Suppose that at time zero, the asset of bank j is $BA_j^0 = BA_j(t = 0)$ and $E_j^0 = \alpha BA_j^0$, $B_j^0 = \beta BA_j^0$, and $M_j^0 = \gamma BA_j^0$.

If Bank i lends money from bank j, then

$$x_{ij}^0 = B_i^0 B_j^0 / \sum_{k:j\in N_k} B_k^0, \tag{3.2}$$

Otherwise $x_{ij}^0 = 0$. Thus L_j is given by

$$L_j^0 = \sum_k x_{jk}^0. \tag{3.3}$$

Since $E_j^0 = \alpha BA_j^0$ and $B_j^0 = \beta BA_j^0$, we have

$$E_j^0 = \alpha \cdot BA_j^0 = B_j^0 \cdot \alpha/\beta. \tag{3.4}$$

Similarly,

$$M_j^0 = \gamma \cdot BA_j^0 = B_j^0 \cdot \gamma/\beta. \tag{3.5}$$

Since $BA_j = D_j + B_j + E_j$, we have

$$D_j^0 = BA_j^0 - E_j^0 - B_j^0 = B_j^0 \cdot (1 - \alpha - \beta)/\beta. \tag{3.6}$$

By equation (3.1), we have

$$I_j^0 = BA_j^0 - L_j^0 - M_j^0 = B_j^0 \cdot (1 - \gamma)/\beta - L_j^0. \tag{3.7}$$

In addition, it requires that the investment should be more than the borrowing, which means $I_j^0 > B_j^0 - L_j^0$, since $I_j^0 = BA_j^0 - L_j^0 - M_j^0 = D_j^0 + E_j^0 + B_j^0 - L_j^0 - M_j^0$, we have $D_j^0 + E_j^0 - M_j^0 > 0$, which implies $\beta + \gamma < 1$.

3.2.2 Renew Balance Sheet

I build a stochastic model for deposit D_j:

$$dD_j(t) = r(t)D_j(t)dt + \sigma_j D_j(t)dW_j(t), \tag{3.8}$$

where $r(t), \sigma_j$ are real constants representing the short interest rate, and volatility respectively. $(W_1(t), \ldots, W_N(t))$ are indepenent wiener process.

For example, the short interest rate for every 3 month in UK is shown in Figure 13.

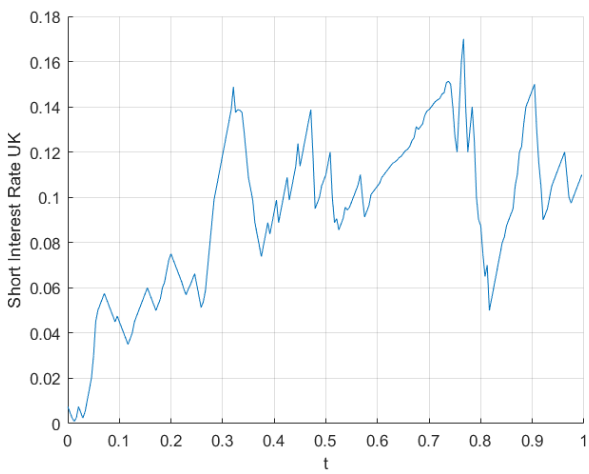


Figure 13: Yahoo Database: UK Short Interest Rate, time gap: 0.01 means 3 months in reality

I replot $r(t)$ with different time gap. In Figure 13, $0.01t$ means 3 months in real world.

By equation (3.8) and data shoen in Figure 13, I use Weak-Euler scheme[**?**] to simulate D_j in UK market, the simulation of Halifax Bank (UK) is shown in Figure 14.

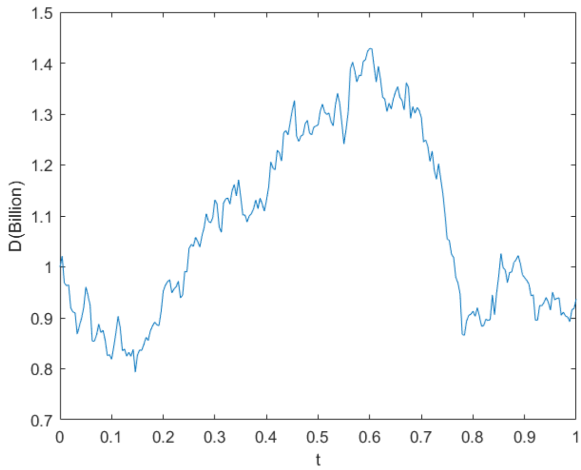


Figure 14: Example: Deposit, $\sigma = 0.36$, time gap: 0.01 means 3 months in reality

Then the owner asset is given by

$$E_j^t = E_j^{t-1} + \mu I_j^{t-1} + r_t^b \left(L_j^{t-1} - B_j^{t-1} \right) - r_t^d D_j^{t-1}, \tag{3.9}$$

where μ is the average return of investment, r_t^b is the interest rate of borrowing and r_t^d is the interest rate of deposit. I assume the $r_t^d = r(t)$ and in $r_t^d = r(t) + 0.02$ in (3.9). By equation(3.4) and (3.6), we get

$$B_j^t = \beta \cdot BA_j^t = \left(D_j^t + E_j^t \right) \cdot \beta / (1 - \beta). \tag{3.10}$$

Similar to equation(3.2) and (3.3),

$$x_{ij}^t = B_i^t B_j^t / \sum_{k:j \in N_k} B_k^t, \tag{3.11}$$

and

$$L_j^t = \sum_k x_{jk}^t \tag{3.12}$$

By equation(3.10), we have

$$M_j^t = \gamma \cdot BA_j^t = \left(D_j^t + E_j^t\right) \cdot \gamma/(1-\beta). \tag{3.13}$$

Since $BA_j = I_j + L_j + M_j$, we get

$$I_j^t = BA_j^t - L_j^t - M_j^t = \left(D_j^t + E_j^t\right) \cdot (1-\gamma)/(1-\beta) - L_j^t. \tag{3.14}$$

In addition, it also requires $\beta + \gamma < 1$ when renew the balance sheet.

3.2.3 Build Interbank Market Network

The construction process of the inter-bank market network is mainly based on the following two assumptions:

1)The banks with small amount of asset always tend to lend money to large banks because it can lower the risks. However, large banks always consider the interest rate of borrowing, and they prefer to spread their funds to different small banks, so as to spread risks as much as possible and obtain higher returns. This assumption means if bank i's borrowing B_i is much bigger than bank i's borrowing B_j, then i tend to lend money to j. If B_i is much smaller than B_j, i still tend to lend money to j.

2)Inter-bank lending is generally short-term, and can only be used as a short-term investment for currency profit and lower risk, not as a long-term investment and financing method. Therefore, the structure of the banks market network should change simultaneously with the inter-bank lending.

Based on assumption 1) and 2), I build the model as follows:

At time t ($(0 \leq t \leq T)$, the probablity that there is an edge between bank i and bank j is given by

$$p_{i,j}^t = 1 - \exp\left(-\lambda\left(B_i^t/B_j^t + B_j^t/B_i^t - 2\right)\right) \tag{3.15}$$

In equation (3.15), we use B_i^t/B_j^t to assess whose interbank lending is bigger. Based on assumption 1), $p_{i,j}^t$ should satisfies when $B_i^t \gg B_j^t$ or $B_i^t \ll B_j^t$, $\quad p_{i,j}^t \to 1$. If we draw the figure of equation (3.15), we will see it satisfies assumption 1) very well:

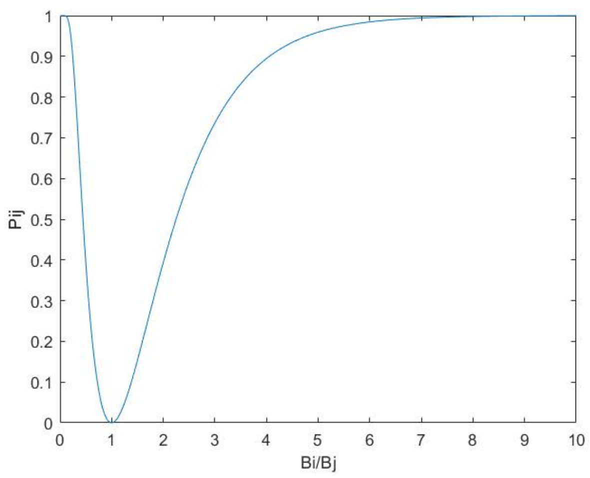


Figure 15: Probablity that i connects j against B_i/B_j, $\lambda = 1$

The reason why I choose $B_i^t/B_j^t + B_j^t/B_i^t - 2$ is $p_{i,j}$ should plunge when $B_i \to B_j$ and increases slowly when $B_i^t \gg B_j^t$. As shown in Figure 15, this model is suitable for describing connection probablity.

I will produce a variable U follows uniform distribution between 0 to 1, if U is samller than $p_{i,j}^t$ then there would be an edge between bank i and bank j and the edge is from i to j when i is not the currency center.

If i is a currency center, normally with a large B_j^t so that $B_i^t << B_j^t$ and $B_i^t/B_j^t + B_j^t/B_i^t \to +\infty, \quad p_{i,j}^t \to 1.$ In this case, the edge between i and j is undirected.

Obviously, $p_{i,j}^t$ is monotonically increasing with the parameter λ. and we have

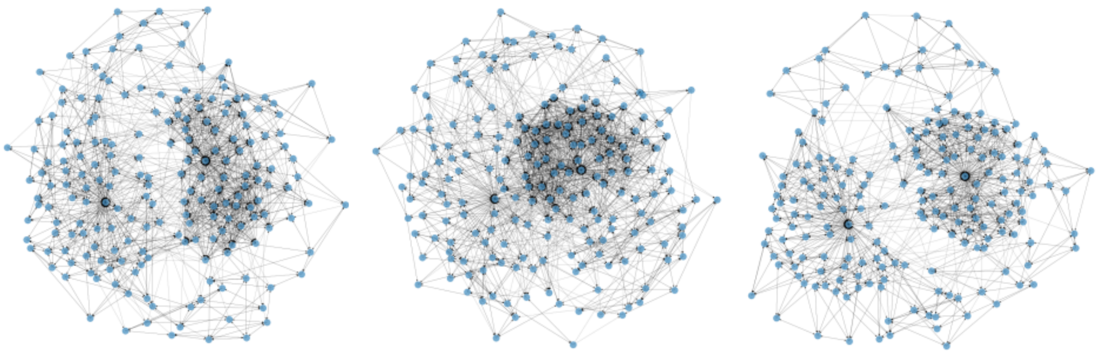
$$B_i^t/B_j^t + B_j^t/B_i^t \geq 2\sqrt{B_i^t/B_j^t \cdot B_j^t/B_i^t} = 2$$

Thus, $\exp\left(-\lambda\left(B_i^t/B_j^t + B_j^t/B_i^t - 2\right)\right) \in (0, 1], \quad p_{i,j}^t \in [0, 1)$, which satisfies the requirements of probablity.

3.2.4 Interbank Market Network

I set $\alpha = 0.05$, $\beta = 0.1$, $\gamma = 0.05$, $\lambda = 1$, and $T = 240$, time gap is 1, which means 3 month in reality. the gap between (a), (b), and (c) is 120.

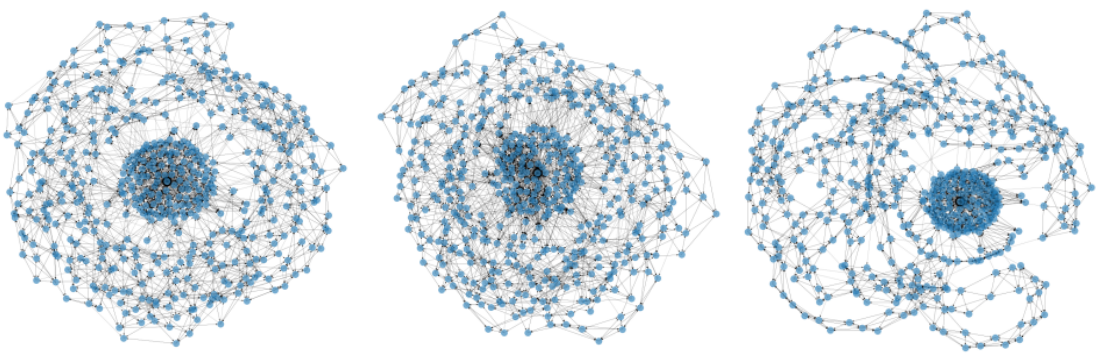
Figure 16 shows that Swiss market is a two-center market with sparse-cluster structure. This network includes 150 largest bank in Switzerland, and the data (Database: Yahoo Database) indicated that there are two currency center in Switzerland. One is UBS whose amount of asset is 9780 billion dollars in 2021 and the other one is Credit Suisse whose amount of asset is 8580 billion dollars while Raiffeisen ranks third with 2290 billion dollars.



(a) t=0 (b) t=120 (c) t=240

Figure 16: Swiss Market, time gap: 1 means 3 month in reality

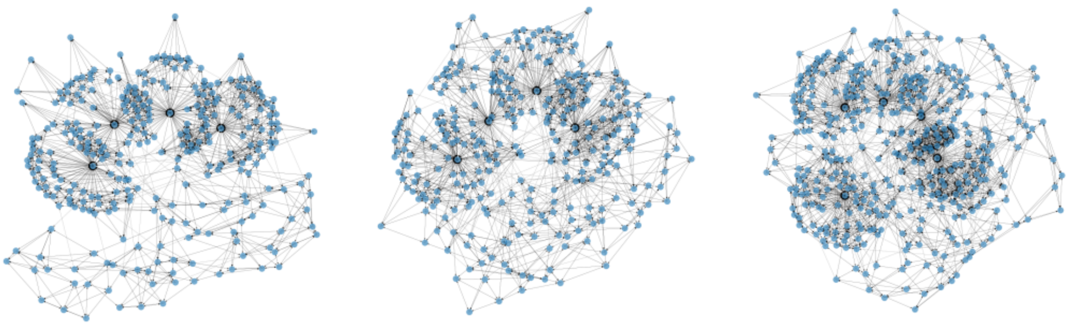
Figure 17 shows the structure of Chinese bank market. This network contains 500 largest banks in CHina. There is a large bank named China bank almost dominates the bank market, its amount of asset is 35578 billion dollars. When t=240, you can see the most of banks' probablity of connection is lower because the Covid-19 in China while the edges coming from China Bank is more than usual.



(a) t=0 (b) t=120 (c) t=240

Figure 17: Chinese Market, one-center, N=500

Figure 18 shows the structure of UK bank market. This network includes 400 largest banks in UK. There are 4 currency centers in UK. They are The Royal Bank of Scotland, HSBC bank, Barclays and Lloyds Banking Group. In recent years, Standard Chartered grows rapidly, ranking fifth in UK market with 6479 billion dollars in 2020, so you can see there are 5 centers in 18(c).



(a) t=0 (b) t=120 (c) t=240

Figure 18: UK Market, four-center, N=400

3.3 Risk Spreading Model

When risk walks along the network through the Inter-bank Lending behavior, it is naturally to consider its 'Infection' as an epidemic. Thus some model of epidemics is suitable for simulating the risk spreading on interbank network. In this section, I will introduce a classic model named SIR model and simulate it on network shown in section 3.2.4.

3.3.1 SIR Model

Assume that total population N is a constant number in short term, and $S = S(t)$ is the number of susceptible individuals, $I = I(t)$ is the number of infected individuals, and $R = R(t)$ is the number of recovered individuals.

Moreover, assume that $s(t) = S(t)/N$ is the susceptible fraction of the population, $i(t) = I(t)/N$ is the infected fraction of the population, and $r(t) = R(t)/N$ is the recovered fraction of the population.

Generally, SIR model is written as ODE form:

$$
\begin{aligned}
&\frac{ds}{dt} = -bs(t)i(t) \\
&\frac{di}{dt} = bs(t)i(t) - ki(t) \\
&\frac{dr}{dt} = ki(t) \\
&s + i + r = 1
\end{aligned}
\tag{3.16}
$$

In equation(3.16), We assume that the population is fixed and ignore births, death and immigration. On average, each infected individual generates $bs(t)$ new infected individuals per day. We also assume that a fixed fraction k of the infected group will recover during any given day.

Actually, when N is very large, $i(t)$ and $r(t)$ can be regarded as the probablity of an individual get infected and recovered respectively. Thus I build another SIR model which is written as stochastic equation.

In a weighted, mixed graph G=(V,E), The banks is described by an $N \times N$ matrix A (the adjacency matrix of G). Assume that node j is the neighbor node of node i and at time t, node j get infected but node i does not, then we obtain the following system of stochastic differential equations:

$$
\begin{aligned}
&\mathrm{d}P_i(t) = \left[\mu s_i(t)\left(1 - P_i(t)\right) - \delta P_i(t)\right]\mathrm{d}t + \sigma_i\left(P_i(t)\right) s_i(t)\left(1 - P_i(t)\right)\mathrm{d}W_i(t), \\
&s_i(t) = \sum_{j \in \Omega_i} \frac{a_{ij} w_{ij} P_j(t)}{\sum_{j \in \Omega_i} a_{ij} w_{ij}}, \quad i \in \{1, \ldots, N\},
\end{aligned}
\tag{3.17}
$$

where μ is a positive constant number, Ω_i is the set of i's infected neighbors. w_{ij} is the weight between i and j. $(W_1(t), \ldots, W_N(t))$ are independent wiener process. $P_i(t)$ is the probablity that node i get infected by its infected neighbors at time t. $s_i(t)$ can be regarded as the strength of the infection that may from infected neighbors of bank i to itself through the network.

During the infection processes, the $P_i(t)$ should strictly remain in $[0, 1]$. System (3.17) has two very good features when simulating $P_i(t)$, the two main theorems is shown below:

Theorem 1. At time zero, for any condition $P(0) = (P_1(0), \ldots, P_N(0))$ such that $P(0) \in [0,1]^N$, there exists a unique solution to equation(3.17) and the solution remains in $[0,1]^N$.

A quik way to see that the solution remains on $[0,1]^N$ is to note that for any $P \in \partial[0,1]^N$ the scalar product between $f(x)$ and the outward normal vector $\nu(x)$ is

$$\langle f(x), \nu(x) \rangle \leq 0, \quad P \in \partial[0,1]^N,$$

and that the diffusion terms is orthogonal to the outward normal vector $\nu(x)$:

$$\langle g(x) \cdot y, \nu(x) \rangle = 0, \quad \text{for all } y \in \mathbb{R}^N, P \in \partial[0,1]^N.$$

STEFANO(2016) gives another method to proof this. The proof gives some stronger theory about stochastic SIR model.

Theorem 2. the null solution for (3.17), is stochastically asymptotically stable in $(0,1)^N$, which means that X_0 is stochastically stable and

$$\lim_{t \to \infty} \mathbb{P}[P(t) = 0] = 1,$$

for all $X(0) \in (0,1)^N$.

The proof of Theorem1 and 2 see appendix A.

Figure 19 shows the probablity of infection of HSBC bank in UK interbank market. As it shows, the probablity of infection will decrease as time goes by and it remains in $[0,1]$. In addition, it has mean-reverting dynamics as well.

At time t, the recovery rate of node i is given by

$$dr_i(t)\,(\theta_1 - \theta_2 r_i(t))\,dt + \gamma\sqrt{r_i}dW_i(t). \tag{3.18}$$

where θ_1, θ_2, γ are positive constants and $2\theta_1 \geq \gamma^2$. Equation(3.18) is a very classic model for short interest rate named CIR model. I choose this model because it has mean-reverting

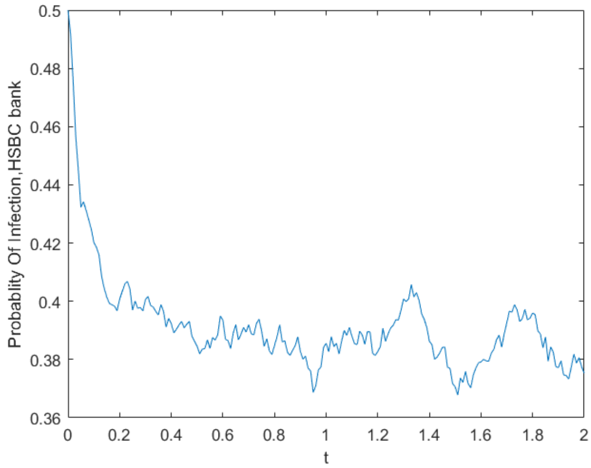


Figure 19: Example: Probablity of infection, HSBC bank, $\mu = 4.1$, $\sigma = 0.36$, T=2

dynamics, and its expectation[**?**] is given by

$$E_Q r(t) = \beta(0)e^{-\theta_2 t} + \frac{\theta_1}{\theta_2}\left(1 - e^{-\theta_2 t}\right). \tag{3.19}$$

Figure 20 shows the recovery rate of HSBC bank in UK interbank market.

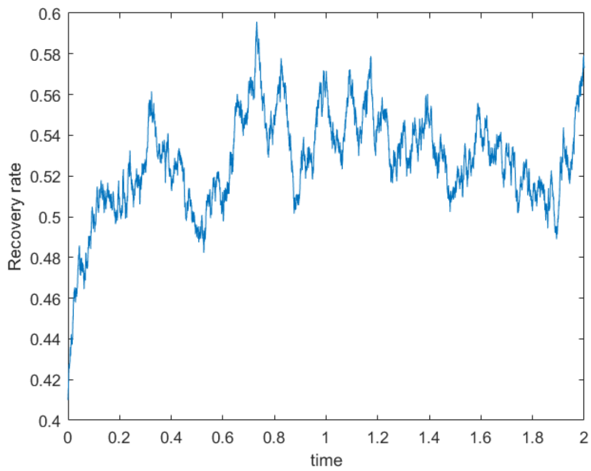


Figure 20: Example: Recovery rate, HSBC bank, T=2

3.3.2 Parameter Estimation

When simulate SIR model in network, another problem is the parameter of the recovery rate. Recall the recovery rate function

$$dr_i(t)\,(\theta_1 - \theta_2 r_i(t))\,dt + \gamma\sqrt{r_i}dW_i(t). \tag{3.20}$$

In order to determine the values for the θ parameters in equation, certain parameter estimation techniques were used. First, the successive $r(t)$ term is defined.

$$r(t+1) = -\beta_1 r(t) + \beta_0. \tag{3.21}$$

Next, an ordinary least squares regression (OLS) process was implemented, resulting in a response vector to build our version of equation (**??**) to use in the parameter estimation, and solve for θ and σ. The OLS scheme gave the following:

$$Y = \beta_1 X + \beta_0 + \epsilon. \tag{3.22}$$

Note that here $Y = (r_1, \ldots, r_n)$ is the response vector, and $X = (r_0, \ldots, r_{n-1})$ is the initial column vector.

Using equation (3.22), the θ parameters were finally estimated as follows:

$$\theta_1 = \frac{\beta_0}{\Delta t}, \theta_2 = \frac{1-\beta_1}{\Delta t}, \sigma = \frac{sd(\frac{\epsilon}{\sqrt{X}})}{\Delta t}. \tag{3.23}$$

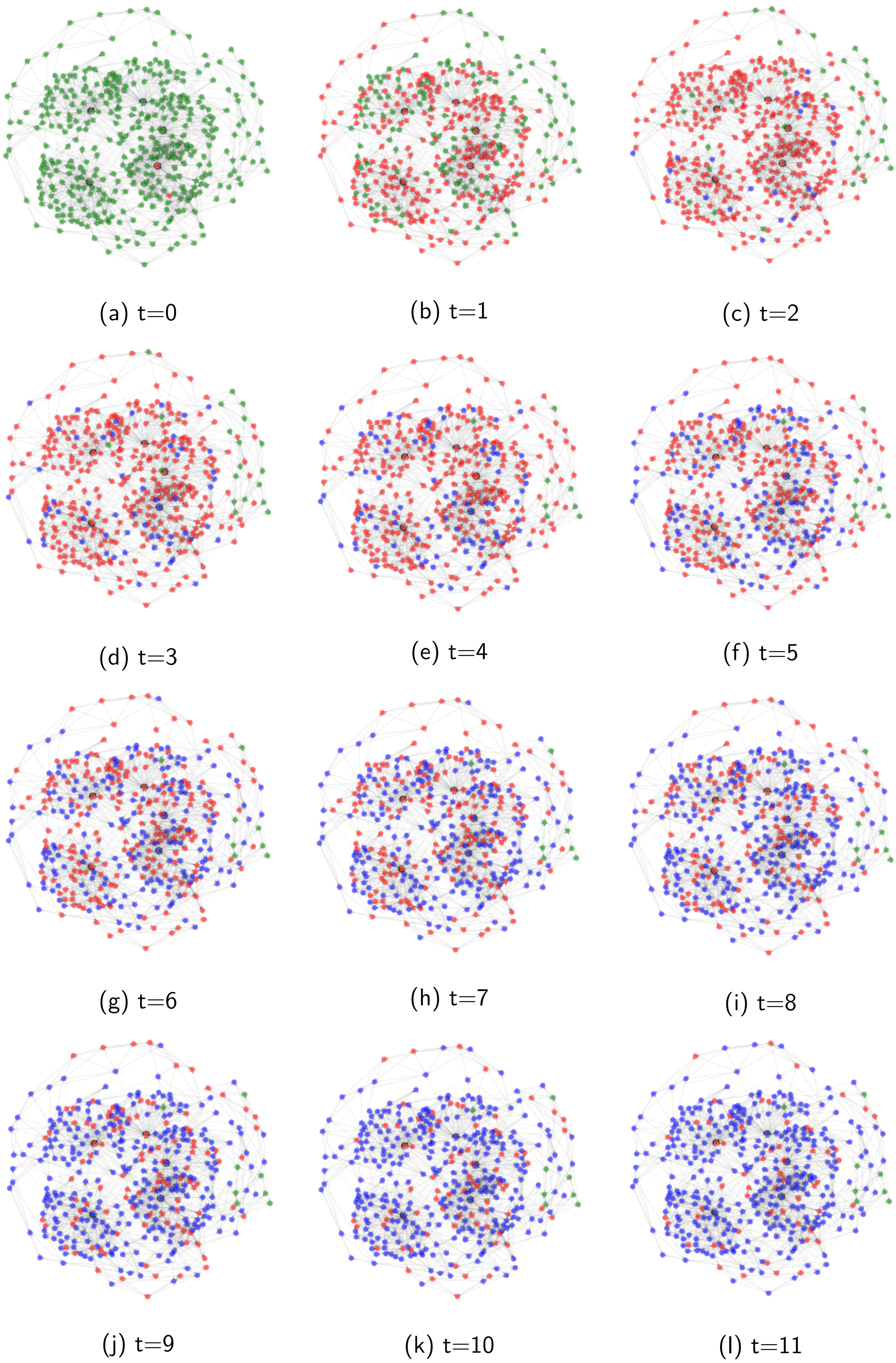
where $sd(\frac{\epsilon}{\sqrt{X}})$ is the standard deviation of vector $\frac{\epsilon}{\sqrt{X}}$ and ϵ is the residual of the OLS estimation.

This estimates the values of our model to be:

$$\theta_1 = 0.24044417;, \theta_2 = 0.4481835, \sigma = 0.025655528 \tag{3.24}$$

3.3.3 Simulation On Network

I choose UK Interbank market as example, select patient zero (get infected at time zero) randomly. Set T=11, P(0)=0.4 for all nodes, and r(0)=0.1 for all nodes. The green node is susceptible banks, red node is infected banks, and blue node is the recovered banks. The results is shown below (time gap between pictures: 1)



(a) t=0 (b) t=1 (c) t=2

(d) t=3 (e) t=4 (f) t=5

(g) t=6 (h) t=7 (i) t=8

(j) t=9 (k) t=10 (l) t=11

Figure 21: Example: SIR model in UK bank market, T=11

The number of three different nodes is shown in Figure 21

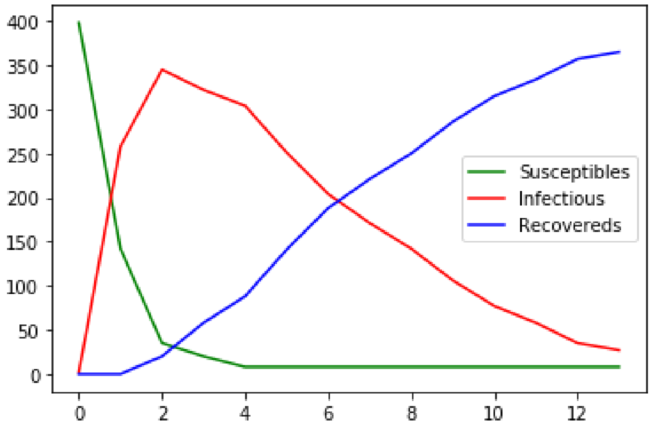


Figure 22: Example: Number Of three nodes

As shown in Figure 21 and 22, most of the infected banks can be recovered itself, but there must be some banks can not recovered before T or get infected again after recovering. Thus in next section, we will predict which bank in graph is more likily to be infected and help government with changing network structure by changing inter-bank lending behavier.

4 Graph Neural Network and Prediction

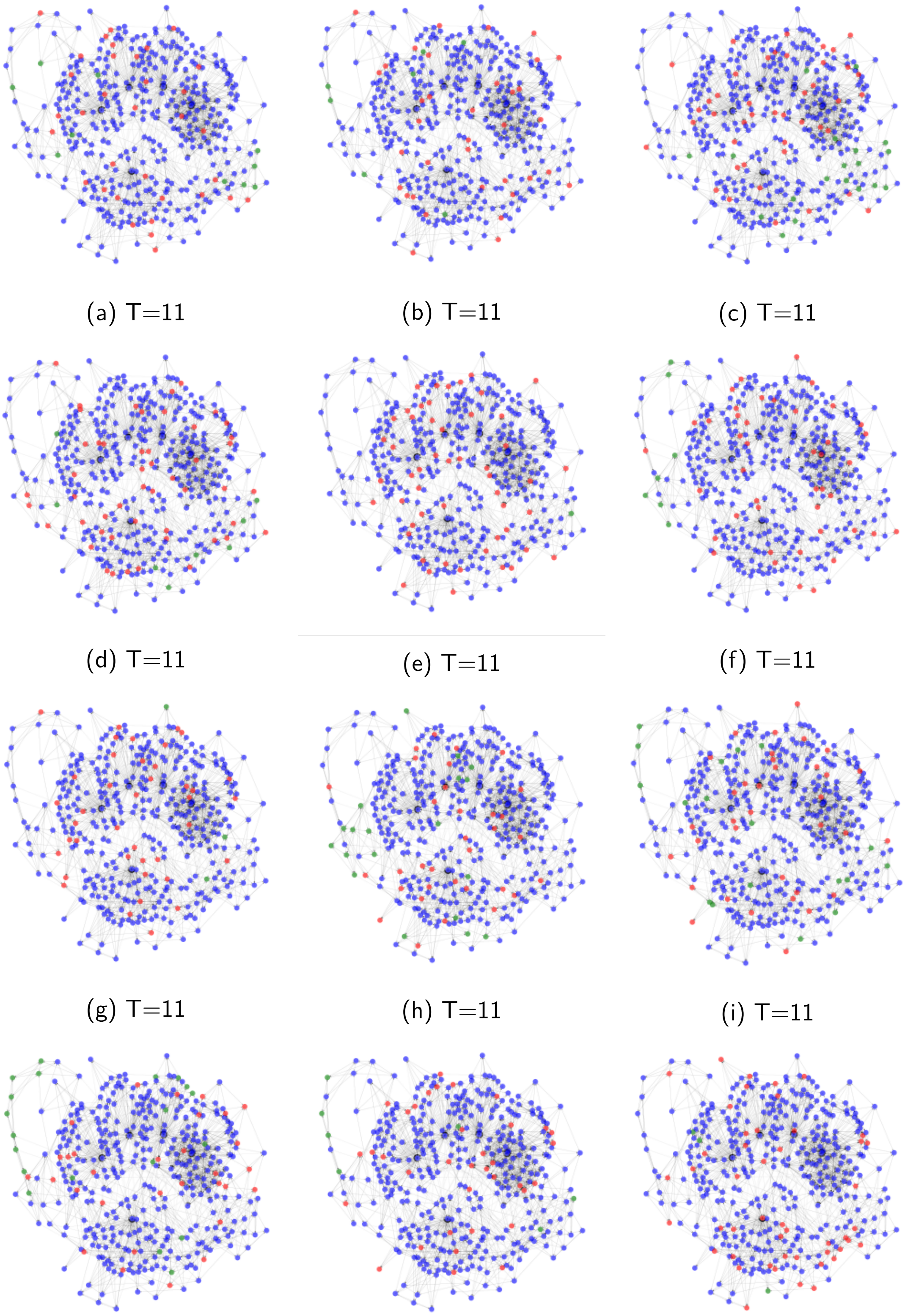
In section 3, I introduce the dynamic network of interbank market and the stochastic SIR model on network. In this section, I will use the infection model to predict which location on network is most likly to be infected when a financial crisis occurs to help government assist the banks in danger or rewire the inter-bank lending connection before the crisis. The first method is classic Monte-Carlo method which is widely used in prediction. The second method is Graph Neural Network. Besides, the training sample of Graph Neural Network is also comes from the result of Monte-Carlo method.

4.1 Monte-Carlo Method

The basic logic of Monte Carlo method is simple. Take UK market network as example, I run the SIR model on network for M=100000 times, and record the proportion of susceptible times, infected times, and recovered times of each node. In this section, the bank is in danger means more than 10000 times (10$\%$ of M times) during the loop, the bank get infected (colored by red in figures). Then if the times of a bank getting infected is less than 10000 and the times of being recovered (colored by blue) is more than 60000 (60$\%$ of total), it would a recovered bank.

Moreover, will the characteristics of graph influences its fraction of dangerous banks when finacial crisis occurs. For example, will the node with larger cluster coefficient becomes easier to be infected during the crisis? And the node with larger centrality (i.e Degree, Closenessand Betweenness e.t.c.) may be in danger in most of cases. The bank structure with smaller average shortest path may be more dangerous. I will discuss more details later.

Figure 22 shows random 12 figures in M=100000 times running. In this example, I choose Uk interbank market network, select two patient zero randomly. Set T=11, $P(0) = 0.4$ for all nodes, and $r(0) = 0.1$ for all nodes. The green node is susceptible banks, red node is infected banks, and blue node is the recovered banks. All figures below is at the same time T=11.



(a) T=11 (b) T=11 (c) T=11

(d) T=11 (e) T=11 (f) T=11

(g) T=11 (h) T=11 (i) T=11

(j) T=11 (k) (l)

Figure 23: Example: Monte-Carlo Method, T=11

Surprisingly, the currency center in UK market is not likily to get infected in the model while the banks which have large Interbank market lending with currency center are in danger.

The main reason of this situation is currency center tends to spread risk to many smaller banks through Interbank market lending so that it always have a larger recovery rate.

In other bank networks, the characteristics of the graph strucure have influence on how dangerous the structure will be. Figure 24 shows the proportion of dangerous banks against average clustering coefficient of graph, T=11. I choose 15 different bank graph to plot the scatter picture

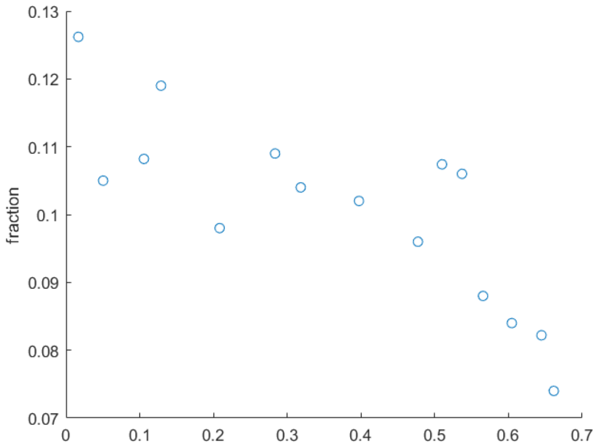


Figure 24: The fraction of dangerous banks against average cluster coefficient, T=11

Figure 25 shows the proportion of dangerous banks against average shortest path of graph.

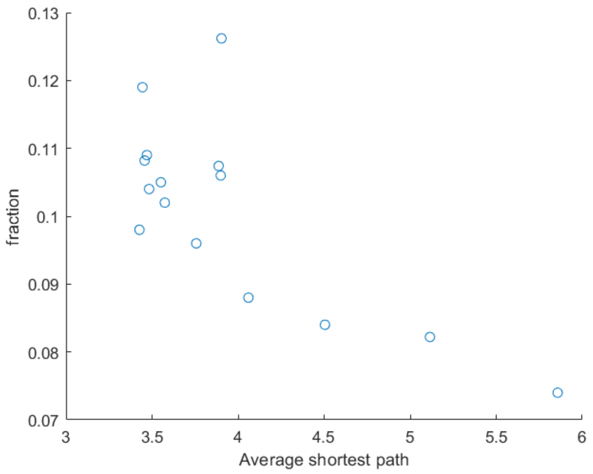


Figure 25: The fraction of dangerous banks against average shortest path, T=11

Basically, when average cluster coefficient and average shortest path decrease, the edges between nodes increases (see Figure6), the graph tends to be more dangerous but this trend is not obvious as the average shortest path decreases.

4.2 Graph Neural Network

In setion 4.1, I introduces the Monte-carlo method to analyse the node in network. It is a very accurate method to predict the infection process. However, Monte-Carlo method always wastes lots of time in application. Naturally, we will use a new method which is not so accurate but more swift to do the prediction. One of the effective method is called Graph Neural Network.

4.2.1 Graph Embedding And Deepwalk

Graph Neural Network is very similar to traditional neural network such as CNNs and RNNs. However, the input layer of neural network is always a data matrix rather than a graph. Thus the first problem is how to learn the structure of the graph and transfer the information of graph to the input vector. The method to solve this problem is called Graph Embedding, and one of the most important method of Graph Embedding is Random Walk.

Figure 26 shows the process of Random Walk. As it shows, Random walk is to perform random walks on the graph structure composed of nodes to generate a large number of nodes sequences, and then input these item sequences as training samples in word2Vec that is introduced in next section.

Random walk is a bit like simulating how the epidemic walks along the network and those sequences is the infection sequences show how the epidemic infect from patient zero to susceptible individuals. For example, one of the sequences 'A-B-E-F' in Figure 26 means B get infected by A, then E get infected by B e.t.c. Then we use those infection sequences to learn the information of the infection network and after that we get the sequences used as input layer in word2Vec.

Let $G = (V, E)$ is the interbank network. Given a partially labeled network $G_L = (V, E, X, Y)$, where $X \in \mathbb{R}^{|V| \times S}$ where S is the size of the feature space for each attribute vector, and $Y \in \mathbb{R}^{|V| \times |\mathcal{Y}|}$ where $\mathcal{Y}$ is the set of labels.

The random walk generator firstly takes a graph G and choose randomly a node v_i as the root of the random walk $\mathcal{W}_{v_i}$. A walk samples uniformly from the neighbors of the last node visited until the length of our walk L is big enough. I will set the length of walk fixed during the simulation.

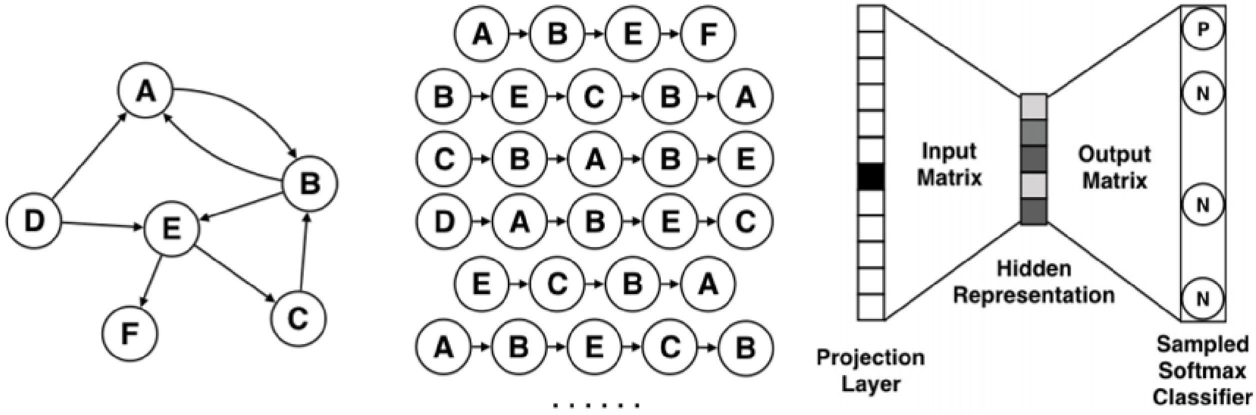


Figure 26: Random Walk

As Figure 26 shows, the Random Walk is divided into 4 parts

1)Constructed the graph g

2)Randomly select the root of walk such as (A,B,C,D,E,A).

3)Then walk from the root for length L to get the sequences of infection. Like 'A-B-E-F' is one of the sequence and length of this walk is 4.

4) inputs these sequences into the word2vec model to generate the final Embedding vector, the word2Vec model will be introduced in next section.

The key problem of Random Walk is on part (3). The probablity of walk from node v_i to node v_j usually is given by

$$P(v_j \mid v_i) = \begin{cases} \frac{\mathbf{W}_{ij}}{\sum_{j \in N_+(v_i)} \mathbf{W}_{ij}}, & v_j \in N_+(v_i), \\ 0, & else \end{cases} \tag{4.1}$$

where W_{ij} is weight of edge between i and j, and $A = (a_{ij})_{N \times N}$, $N_+(v_i)$ means the set of i's neighbor nodes. For example, if i have 2 neighbor nodes, and the weight of them are both 50. We generates a variale U following Uniform distribution in [0,1]. Then divide [0,1] into 2 parts [0,0.5], (0.5,1] and see which part does U locate, then we get a walk sequence from i to one of its neighbor node.

4.2.2 Word2Vec

In last section, we have mentioned embedding vectors is generating by lots of walk sequences. In this section, we will use a very famous language model named Word2vec to solve this problem.

The main goal of Word2vec is to estimate the probablity of a specific sequence of words appearing in a corpus. Assume the given word x_k, and the output will be the whole sentence $\{\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{C}\}$. For example, the sentence is 'I drive my car to the store'. If we input 'car' as the input layer, then the output will be { 'I', 'drive', 'my', 'to', 'the', 'store'}. The Figure 27 shows the its basic logic of Word2Vec.

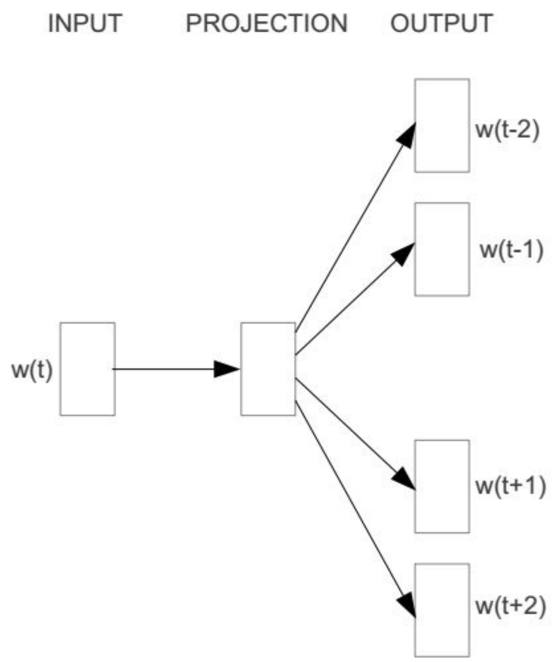


Figure 27: Word2Vec

By Random walk, we get lots of sequences such as 'A-B-E-F', and 'E-C-B-A' (see Figure 26). Those sequences are the sample of the WordVec model. For example, if 'A-B-E-F' is 'I'-'drive'-'my'-'car', then Word2Vec will output 'A,E,F' if we input 'B'.

Figure 28 shows the input and output when traing Word2Vec model. As it shows, We input the one-hot code vector x of a single input sequence. For example, 'A-B-E-F' is the vocubulary sheet. Then one-hot code of B is a 5-dimensional vector with 1 in second position of vector while others are all zero: $(0, 1, 0, 0)$. Similarly, A is $(1, 0, 0, 0)$, E is $(0, 0, 1, 0)$ and F is $(0, 0, 0, 1)$.

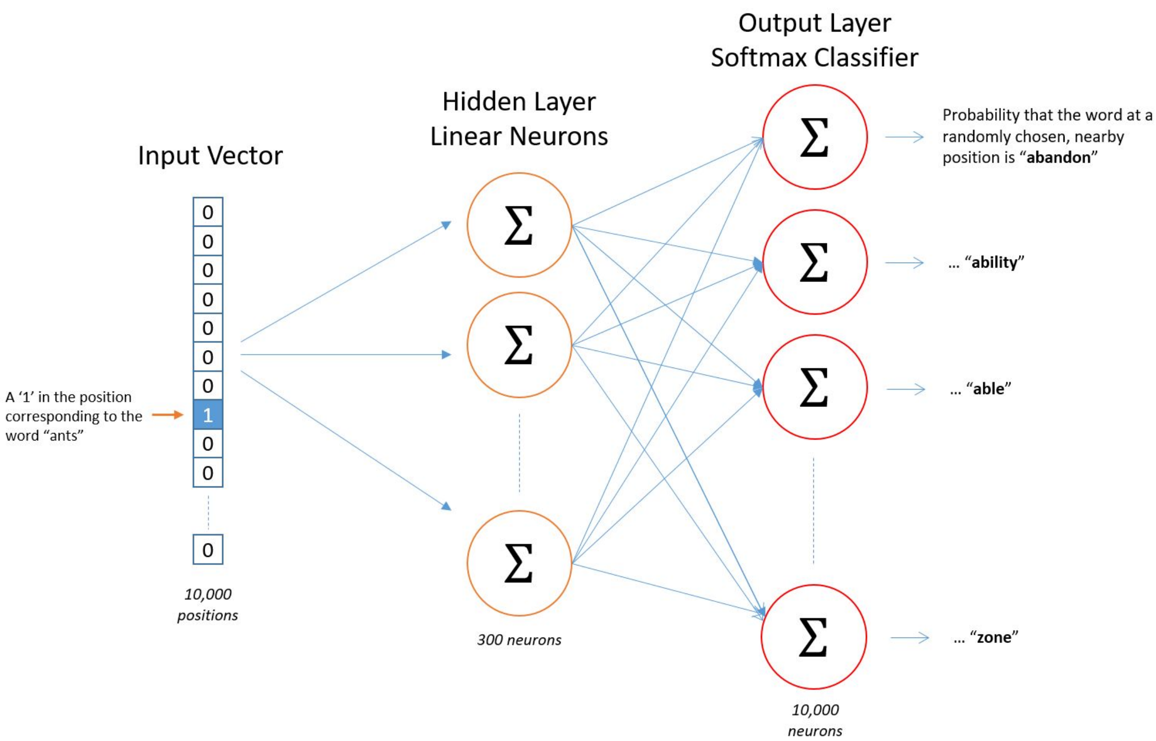


Figure 28: Input and output of Word2Vec

Figure 29 shows the how Wors2Vec generates Embedding vector of graph. Assume output vector is $\{\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{C}\}$ (see Figure 29), where y_c is a V-dimensinal vector whose all elements $y_{c,j}$ remains in [0,1] and $\sum_{c=1}^{C} y_c = 1$. Although Word2Vec is used to output 'A,E,F', what we actually do is estimating probablity for each of words. For instance, If the index of biggist number in y_c is 1, then the c^{th} word is 'A'. If the index of biggist number in y_c is 2, then the c^{th} word is 'B' e.t.c.

$\{\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{C}\}$ is called the Embedding vector of the graph. What we really want is Embedding vector which can be input layer in GNNs (see section 4.2.4) instead of 'A,E,F'. Now we can build Word2Vec model in our specific problem:

Assume the sample is given by the random walk sequences $\{'A - B - E - F - \ldots', 'E - C - D - H - \ldots', \ldots\}$, which includes V different nodes in total. Firstly, we code each node as one-hot code (i.e A for $(1, 0, 0, \ldots)_V$). Those one-hot code are the input vector x shown in figuer 29.

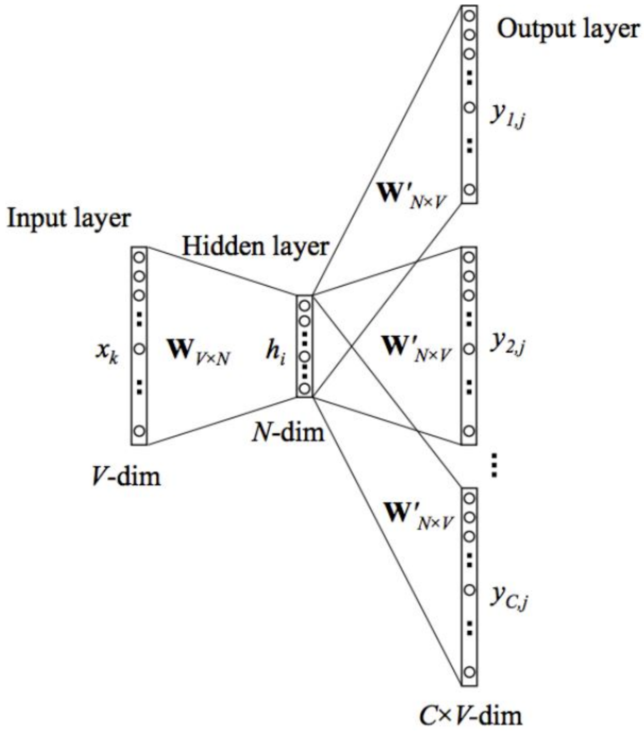


Figure 29: Output layer of Word2Vec

Secondly, we train the model as follows:

Since the i^{th} column of weight matrix W represents the weight of i^{th} code of input word and the zero elements will not influence hidden layer. if $x_k = 1$ then the hidden layer h equals to the k^{th} column of the weight matrix, which is given by

$$\mathrm{h} = \mathrm{x}^{\mathrm{T}}\mathrm{W}_{V\times N} = W_{(k)}. \tag{4.2}$$

For example

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \\ m & n & o \end{bmatrix} = \begin{bmatrix} j & k & l \end{bmatrix}. \tag{4.3}$$

then the output layer $\mathrm{u}_{\mathrm{c,j}}$ is given by

$$\mathrm{u}_{\mathrm{c}} = \mathbf{W}_c^{\prime T}\mathrm{h}. \tag{4.4}$$

At last, we use softmax function to get output $\{\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{C}\}$ by output layer $\mathrm{u}_{\mathrm{c,j}}$.

$$\mathrm{y}_{\mathrm{c,j}} = \frac{\exp\left(\mathrm{u}_{c,\mathrm{j}}\right)}{\sum_{\mathrm{j}'=1}^{\mathrm{V}} \exp\left(\mathrm{u}_{\mathrm{cj}}{}'\right)}. \tag{4.5}$$

Then we can determine which word in vacabulary sheet is most likily on the y_c location, which means

$$\mathrm{p}\left(\mathrm{w}_{\mathrm{O,c}} \mid \mathrm{w}_\mathrm{I}\right) = \mathrm{y}_{\mathrm{c,j}} \tag{4.6}$$

Then we can write down the loss function of this neural network:

$$\begin{aligned} minE &= -\log \prod_{\mathrm{c}=1}^{\mathrm{C}} \mathrm{p}\left(\mathrm{w}_{\mathrm{O,c}} \mid \mathrm{w}_\mathrm{I}\right) \\ &= -\log \prod_{\mathrm{c}=1}^{\mathrm{C}} \frac{\exp\left(\mathrm{u}_{\mathrm{c,j}}\right)}{\sum_{\mathrm{j}'=1}^{\mathrm{V}} \exp\left(\mathrm{u}_{\mathrm{cj}}{}'\right)}. \end{aligned} \tag{4.7}$$

By Backpropagation algorithm (BP) (Gareth James, Statistical Learning) of classic neural network and stochastic gradient descent algorithm (SGD), we get the $W' = (w'_{ij})_{N\times V}$ is given by

$$\mathrm{w}'(\text{ new }) = \mathrm{w}'_{\mathrm{ij}}(\text{ old }) - \eta \cdot \sum_{\mathrm{c}=1}^{\mathrm{C}} \left(\mathrm{y}_{\mathrm{c,j}} - \mathrm{t}_{\mathrm{c,j}}\right) \cdot \mathrm{h}_\mathrm{i}. \tag{4.8}$$

and W is given by

$$\mathrm{w}^{(\text{new })} = \mathrm{w}_{\mathrm{ij}}^{(\text{old })} - \eta \cdot \sum_{\mathrm{j}=1}^{\mathrm{V}} \sum_{\mathrm{c}=1}^{\mathrm{C}} \left(\mathrm{y}_{\mathrm{c,j}} - \mathrm{t}_{\mathrm{c,j}}\right) \cdot \mathrm{w}'_{\mathrm{ij}} \cdot \mathrm{x}_\mathrm{j}. \tag{4.9}$$

where η is study rate.

Then we can write down the algorithm of Word2Vec and Deepwalk. The are all shown below

Deepwalk: (G, d, n, L)

Input: graph $G(V, E)$, dimension of embedding vector d, the number of walks n, Walk length L, and window size w

Output: embeding matrix $X \in \mathbb{R}^{|V|\times d}$ for each node v_j points to a vector $x \in \mathbb{R}^d$

1. Initialization weight matrix W, sample $X \in \mathbb{R}^{|V| \times d}$ to get training set of Word2Vec
2. Build a binary Tree T from V
3. for $i = 0$ to n do
4. $\quad \mathcal{R} = \text{Shuffle}(V)$
5. $\quad$ for each node $v_i \in \mathcal{R}$ do
6. $\quad\quad$ Sequences $\mathcal{W}_{v_i} = \text{Random walk}\,(G, v_i, L)$
7. $\quad\quad$ Word2Vec$(X, \mathcal{W}_{v_i}, w)$

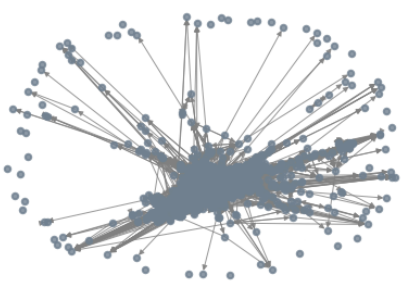
Word2Vec: (X,$\mathcal{W}_{v_i}$)

1. for each $v_j \in \mathcal{W}_{v_i}$ do
2. $\quad$ for each $u_k \in \mathcal{W}_{v_i}[j - w : j + w]$ do
3. $\quad\quad E(W) = -\sum \log \Pr(u_k \mid W, v_j)$
4. $\quad\quad W = W - \eta * \frac{\partial E}{\partial W}$

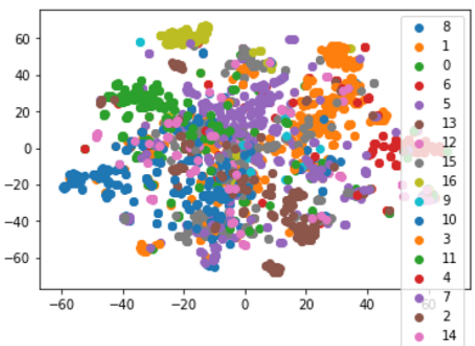
4.2.3 Graph Embedding Results

I shows 3 examples of the results of Graph Embedding inthis section.

Figure 30(a) shows the global interbank market network which contains 2405 banks in total. Figure 30(b) shows the Embedding results of this graph. The Embedding vector is 10-dimensional and I use TSNE function (similar to PCA method) in python to lower the dimensions to draw the embedding picture(b)



(a) Global network

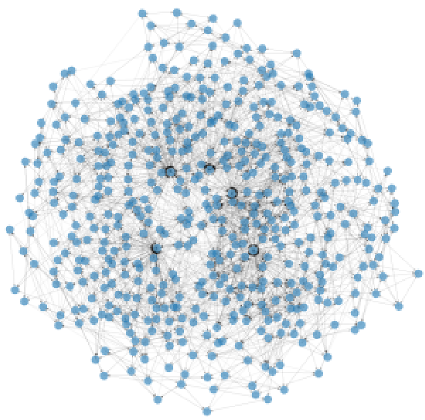


(b) Graph Embedding Results (10-dimension)

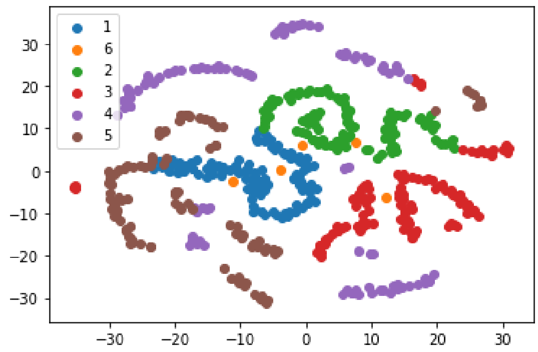
Figure 30: Embedding results of global bank network, N=2405

As shown in Figure 29, the nodes whose partial structure are similar will have similar embedding vector because the random walk sequences of those nodes is very similar, so those nodes will be colored by same color in TSNE model.

Figure 31 shows the result of graph embedding of UK bank market which contains 400 banks in total.



(a) UK network

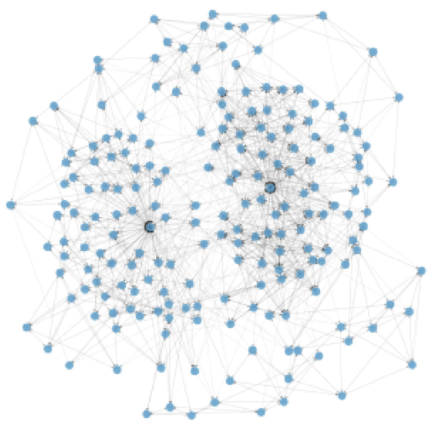


(b) Graph Embedding Results, lower dimension by TSNE method

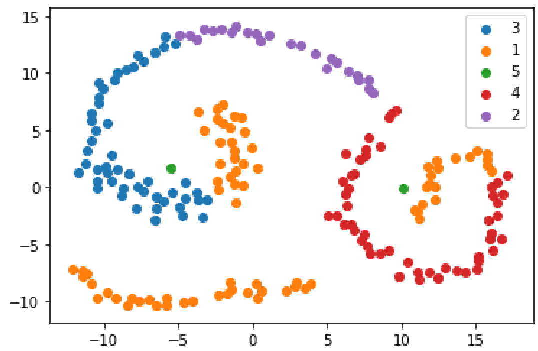
Figure 31: Embedding results of UK bank network, N=400

Similarly, the nodes have similar sequences will have same color. Note that 5 currency centers of UK market is colored by orange in 31(b)

Figure 32 shows the result of graph embedding of Swiss bank market which contains 150 banks in total.



(a) Swiss network



(b) Graph Embedding Results, lower dimension by TSNE method

Figure 32: Example: Swiss Bank Market, N=150

The embedding result 32(b) is symmetric because Swiss market network is a 2-center market that has symmetric structure.

4.2.4 GNNs

In section 4.2.3, we have generated embedding vector which is input layer in Graph Neural Network. In this section, I will introduce how to construct the whole Graph Neural network (i.e GCNs).

Figure 32 shows the process of GNNs

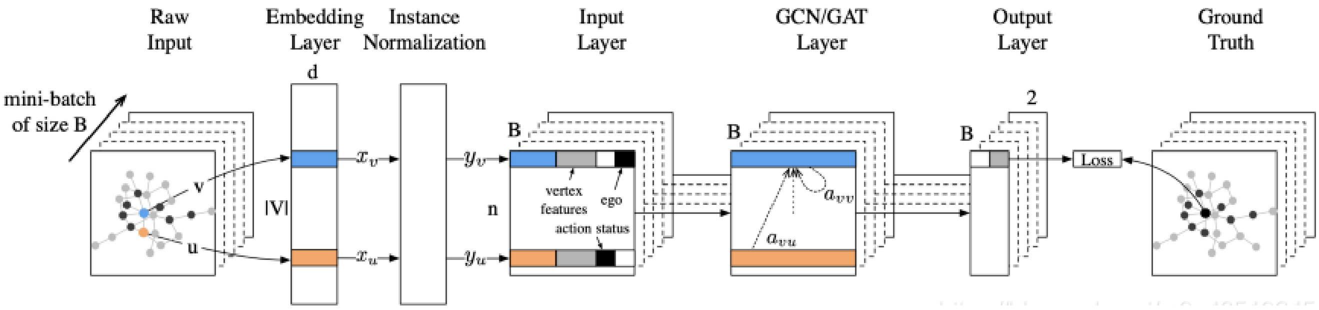


Figure 33: GNNs

As shown in Figure 33, we have already got the input vector, and GCN layer is very similar to the CNNs. Thus we write the problem in mathematical way.

As we discussed in section 4.1, the banks in interbank network is divided into three categories: dangerous, potential patient, and recovered, which is denoted by 1, 2, and 3 respectively. Then what GNN do is classifying all nodes in a network to 1, 2 or 3. It can help government with clarifying which banks' connection is not healthy in interbank network.

Firstly, we use Monte-Carlo method to get sample from 467 graphs including 2405 banks in total. Then each bank in each graph is labeled with

$$y_j = \begin{cases} 1 & \text{Infection} \\ 2 & \text{Susceptible} \\ 3 & \text{Recovered}\ , \end{cases} \tag{4.10}$$

where $j = 1, 2, \dots, p$.

Then we assume that the output of nueral network is $(f_1(X), f_2(x), f_3(x)) \in [0,1]^{p\times 3}$, where $f_i(X) = P(Y = i \mid X)$ is the p-dimensional probablity vector for all banks in one graph. Although the model is used for classification, what we actually do is estimating probablity for each of 3 classes. Then the banks are assigned to the class which has the highest probablity. The process of GCN is shown in Figure 33.

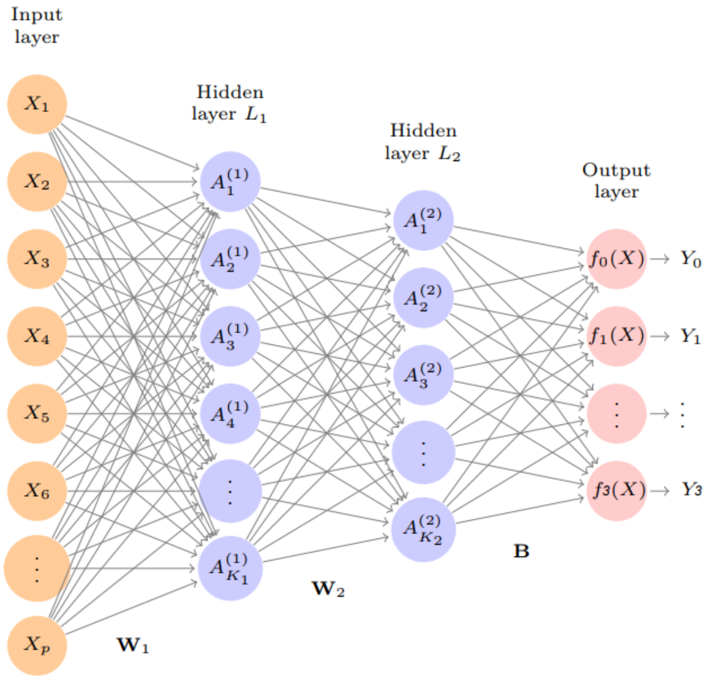


Figure 34: GCN

Thus, the loss function of the problem is given by

$$minE(W) = -\sum_{j=1}^{p}\sum_{i=1}^{3} y_{ji}\log\left(f_i\left(x_j\right)\right). \tag{4.11}$$

The first hidden layer L_1 is given by

$$A_k^{(1)} = \sigma\left(w_{k0}^{(1)} + \sum_{j=1}^{p} w_{kj}^{(1)} X_j\right), \tag{4.12}$$

where $\sigma()$ is activation function. I use sigmoid function between last hidden layer and output layer to confirm $f_i(x)$ is not zero. And in other hidden layer, I use ReLU function. Those two activation function is given by

$$sigmoid: \quad \sigma(z) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}, \qquad ReLU: \quad g(z) = (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{otherwise .} \end{cases}$$

The second hidden layer is given by

$$A_l^{(2)} = \sigma\left(w_{l0}^{(2)} + \sum_{j=1}^{p} w_{lj}^{(2)} A_k^{(1)}\right). \tag{4.13}$$

The other hidden layer is similar to equation(4.16) and (4.17). And the output layer is given by

$$Z_i = \beta_{i0} + \sum_{\ell=1}^{K_2} \beta_{i\ell} A_\ell^{(2)}. \tag{4.14}$$

At last, we get output layer Z, then the output $f_i(x)$ (i=1,2,3) is given by

$$f_i(X) = P(Y = i \mid X) = \frac{e^{Z_i}}{\sum_{\ell=1}^{3} e^{Z_\ell}}. \tag{4.15}$$

By Backpropagation (BP) algorithm and Stochastic gradient descent (SGD) algorithm, we can do the training and prediction now.

4.3 Training And Prediction

In this section, I will use GNNs introduced in section 4.2.4 to predict the role of banks in interbank market network. Moreover, I will discuss further about whether the characteristics of each graph influences the accuracy of prediction.

Sample contains banks in different countries in different years, and it is about 437 graphs and 2405 banks in total. Training set contains 70% (306) of the graphs. And validation set contains 131 graphs including 1233 banks in total.

Firstly, we use Monte-Carlo method to label each bank in each graph. Then the Graph Neural Network will train the graph in training set to get the weight matrix W. At last, model will use weight matrix W to label each bank in each graph in validation set. Note that some of the banks are not domestic banks but it do occurs in domestic graph. For example, HSBC bank is one of the currency center in UK market and it is also important in Hongkong market. And most of banks participates in USA bank market, so the result of USA is typical.

The accuracy of prediction in USA market is shown in table 1.

	3 (Recovery)	2(Susceptibles)	1(Danger)	Total(prediction)
3	697	2	37	736
2	21	11	7	39
1	75	0	113	188
Total(Truth)	793	13	157	963

Table 1: The number of banks classified correctly in USA market, N=963, T=11

Note that the data in diagonal of Table 1 are the banks predicted correctly. For example, we get 793 banks is '3' in USA by using Monte-Carlo method, and 697 of them is assigned to '3' by graph neural network. Then the accuracy rate of classification is given by

$$(697 + 11 + 113)/963 = 85.25\%$$

.

Another example is the German bank market. The resuts is shown below

	3(Recovery)	2(Susceptibles)	1(Danger)	Total(prediction)
3	311	0	11	322
2	7	3	4	14
1	58	1	32	91
Total(Truth)	376	4	47	427

Table 2: The number of banks classified correctly in German market, N=427, T=11

The accuracy rate of classification in Germany is given by

$$(311 + 3 + 32)/427 = 81.03\%$$

.

The average cluster coefficient of USA network is 0.1288570065060429 while the average cluster coefficient of German netwok is 0.5654755218490474. Thus we firstly assume that when The average cluster coefficient decreases, the GNNs maybe performs better becuase the embedding vector is more accurate in descrbing the structure aroud a specific node when the graph become more sparse.

Figure 35 shows the relationship between accuracy rate and average cluster coefficient of 131 validation graphs.

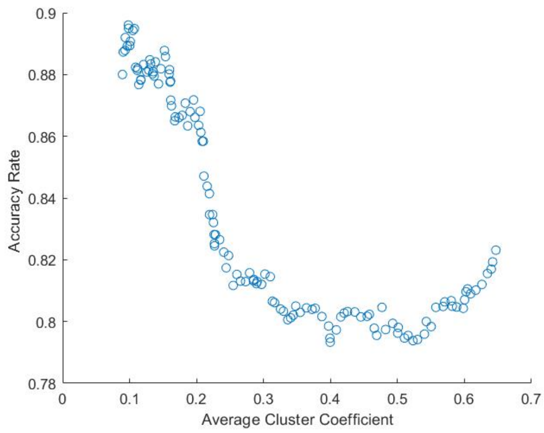


Figure 35: Accuracy rate against average cluster coefficient

Basically, the assumption is right, but the trend is not obvious when the average cluster coefficient exceeds 0.4.

Figure 34 shows the relationship between accuracy rate and average shortest path of 131 validation graphs.

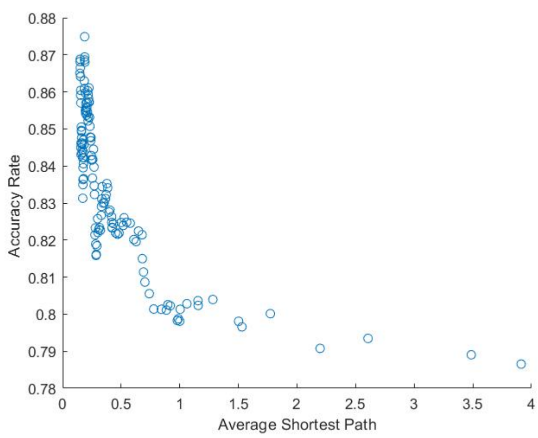


Figure 36: Accuracy rate against average shortest path

This trend is more obvious. When the average shortest path increases, the accuracy decreases. As we discussed in section 2 and 3, the interbank neteork have small-world charicteristic, which means its average shortest path is smaller than other random graph. As shown in Figure 36, most of graphs' average shortest path remain on [0,0.5]. And banks in graphs whose average shortest path is large are usually small banks and have less connection with each other. Those banks are difficult to predict even Monte-Carlo method doesn't work well on those graphs.

5 Conclusions

5.1 Conclusion

On the basis of the current research, there are some areas that need to be improved in the inter-bank market research. This paper is based on complex network and neural network. The following main conclusions are drawn through model:

(1)Construction of Dynamic interbank network

We build the interbank network baased on the Inter-bank Lending procedures. The simulation shows that the interbank network has small world character, which means it has small average shortest path and large average cluster coeffienct. In addition, a few banks have large degree centrality which is called currency center while most of banks have similar, small degree centrality. The structure of network remains dynamic stable in short term as short interest rate changing.

(2)Build the stochastic SIR model.

We build stochastic SIR model based on its original diffetential equations' (ODEs) form. The dicussian and simulation shows that both infection probablity and recovery rate has mean-reverting dynamics and their result always remains on [0,1].

(3)Applied SIR model on interbank network.

We use stochastic SIR model to simulate how the financial crisis is 'infect' through the Inter-bank Lending. The simulation shows that the graph structure with short average shortest path and small average cluster coeffienct magnifies the risk of infection during the crisis because more connections have been set up.

(4)Use graph neural network to do the prediction.

We use graph neural network to do the prediction because Monte-Carlo method wastes lots of time. Though the classification of neural network is less accurate, the simulation shows that the accuracy always remains above 75 % during the experiments. In addition, The accuracy of Graph neural network is stable but it is more accurate when average shortest path decreases.

5.2 Further Work

Since the interbank-lending is more complicated in reality, there are more problems require furthet reasearch in the future:

1) The balnce sheet of a bank includes lots of items instead of several essential parts such as deposit and investment. Moreover, deposit is not only determined by short interest rate, so the stochastic model driven by short interest rate does not descibe deposit well.

2) Random walk still wastes lots of time when network is large. Since its mechanism is a bit like Breadth First Search (BFS). Some new methods proposed recent years such as LINE (MIT 2017) and EGES (Alibaba 2018) are better, those methods are related to Depth First Search (DFS).

3) When the graph is very large, to many neurons and hidden layers may cause over-fitting problem, some classic method such as randomly drop out some neurons or regularization still work in solving over-fitting problem. The regularization means add parameter to loss function, which means the new loss function is given by

$$minimize \quad loss \quad function + \lambda \mid\mid W \mid\mid_1^2),$$

where W is parameter matrix and $\mid\mid W \mid\mid_1 = \sum w_{ij}$ is 1-norm of matrix W.

A Proof Of Important Theorem

A.1 Proof Of Theorem1

Assume that function F:$(0,1)^N \to \mathbb{R}^+$ is given by

$$F(X(t)) = -\sum_{i=1}^{N} \log\left[x_i(t)\left(1-x_i(t)\right)\right] \tag{A.1}$$

By Itô's formula we get

$$\begin{aligned} \mathrm{d}F(X(t)) = & \sum_{i=1}^{N}\left(\frac{1}{1-x_i}-\frac{1}{x_i}\right)\left[\left(\beta s_i\left(1-x_i\right)-\delta x_i\right)\mathrm{d}t+\sigma_i\left(x_i\right)s_i\left(1-x_i\right)\mathrm{d}w_i(t)\right] \\ & +\frac{1}{2}\sum_{i=1}^{N}\left(\frac{1}{\left(1-x_i\right)^2}+\frac{1}{x_i^2}\right)\sigma^2\left(x_i\right)s_i^2\left(1-x_i\right)^2\ \mathrm{d}t \end{aligned} \tag{A.2}$$

We rewrite equation(A.2) as

$$\mathrm{d}F(X(t)) = LF(X(t))\mathrm{d}t + \mathrm{d}M(t) \tag{A.3}$$

where

$$\begin{aligned} LF(X(t)) = & \sum_{i=1}^{N}\left(\frac{1}{1-x_i}-\frac{1}{x_i}\right)\left(\beta s_i\left(1-x_i\right)-\delta x_i\right) \\ & +\frac{1}{2}\sum_{i=1}^{N}\left(\frac{1}{\left(1-x_i\right)^2}+\frac{1}{x_i^2}\right)\sigma^2\left(x_i\right)s_i^2\left(1-x_i\right)^2 \end{aligned} \tag{A.4}$$

and

$$M(t) = \sum_{i=1}^{N}\int_0^t\left(\frac{1}{1-x_i(t)}-\frac{1}{x_i(t)}\right)\sigma_i\left(x_i(t)\right)s_i(t)\left(1-x_i(t)\right)\mathrm{d}w_i(t). \tag{A.5}$$

Lemma 1: If the coefficients of the equation are locally Lipschitz continuous, There will be a positive constant number K such that $LF(X) \le K\ \forall X \in \Delta$.

Firstly, the coefficients of the equation are locally Lipschitz continuous, then exists a positive constant number M such that $\sigma_i(x)$ in (3.17) satisfies

$$\sup_{x\in(0,1)} \frac{\sigma_i(x)}{x} \leq M, \quad \text{for every } i = 1, \ldots, N \tag{A.6}$$

Since

$$s_i(t) = \sum_{j\in Infected} \frac{a_{ij} w_{ij} x_j(t)}{\sum_{j\in Infected} a_{ij} w_{ij}} \leq \sum a_{ij} \leq N-1,$$

we have

$$\frac{1}{2}\left(\frac{1}{(1-x_i)^2} + \frac{1}{x_i^2}\right)\sigma^2(x_i)\, s_i^2 (1-x_i)^2 \leq \frac{1}{2}\left(\frac{x_i^2 + (1-x_i)^2}{(1-x_i)^2 x_i^2}\right) M^2 x_i^2 s_i^2 (1-x_i)^2 \leq M^2(N-1)^2 \tag{A.7}$$

Note that this is the last term of LF(X(t)). Then we have

$$\begin{aligned} LF(X(t)) =& \sum_{i=1}^{N}\left(\frac{1}{1-x_i} - \frac{1}{x_i}\right)(\beta s_i (1-x_i) - \delta x_i) \\ &+ \frac{1}{2}\sum_{i=1}^{N}\left(\frac{1}{(1-x_i)^2} + \frac{1}{x_i^2}\right)\sigma^2(x_i)\, s_i^2 (1-x_i)^2 \\ &\leq \frac{2x_i - 1}{x_i(1-x_i)}(\beta s_i (1-x_i) - \delta x_i) + M^2(N-1)^2 \end{aligned} \tag{A.8}$$

Since function $y(x) = \frac{2x-1}{x(1-x)}(\beta s(1-x) - \delta x), \quad x \in (0,1)$ reachs maximum point when $x^* = \frac{\sqrt{\beta s}}{\sqrt{\beta s}+\sqrt{\delta}}$, we have

$$LF(X) \leq N\left[\beta(N-1) + \delta + M^2(N-1)^2\right] = K \tag{A.9}$$

Theorem 1. At time zero, for any condition $X(0) = (x_1(0), \ldots, x_N(0))$ such that $x(0) \in [0,1]^N$, there exists a unique global solution to equation(3.17) and the solution remains in $[0,1]^N$.

For any $x(0) \in (0,1)^N$, there would be a unique local solution on $t \in [0,T)$. The solution does not converge when $t \geq T$.

Consequently, if the solution to (3.17) is global, $\tau_e = \infty$. $\forall i = 1, \ldots N$, let $n_0 > 0$ is large enough so that for $x_i(0) \in \left(\frac{1}{n_0}, 1-\frac{1}{n_0}\right)$. Then for each integer $n \geq n_0$, we define the T_n as follow

$$T_n = \inf\left\{t \in [0,T) : \min_{1\leq i\leq N} x_i(t) \leq 1/n \text{ or } \max_{1\leq i\leq N} x_i(t) \geq 1-1/n\right\}$$

Then we need to show $\lim_{n\to\infty} T_n = \infty$. If not, there must be a positive constant number λ and $\epsilon \in (0,1)$ such that

$$\mathbb{P}\left\{T_n \leq \lambda\right\} > \epsilon \tag{A.10}$$

which means there is an integer $n_1 \geq n_0$ such that

$$\mathbb{P}\left\{T_n \leq \lambda\right\} \geq \epsilon \quad \forall n \geq n_1 \tag{A.11}$$

By Lemma1 we have,

$$\int_0^{T_n\wedge\lambda} \mathrm{d}F(X(t)) \leq \int_0^{T_n\wedge\lambda} K \ \mathrm{d}t + M(t) \tag{A.12}$$

Take expectation of each side of (A.12), we have

$$\mathbb{E}\left[F\left(X\left(T_n\wedge\lambda\right)\right)\right] \leq \mathbb{E}[F(X(0))] + K\mathbb{E}\left(T_n\wedge\lambda\right) \leq \mathbb{E}[F(X(0))] + K\lambda \tag{A.13}$$

Since $\{T_n \leq \lambda\}$ for $n \geq n_1$, there must be a $x_j(T_n)$ such that equals to $1/n$ or $1-1/n$

$$F\left(X_j\left(Tn\right)\right) \geq -\left(\log\left(\frac{1}{n}\right) + \log\left(1-\frac{1}{n}\right)\right) \tag{A.14}$$

By equation(A.13) and (A.14) and $\mathbb{P}\left\{T_n \leq \lambda\right\} \geq \epsilon$ we have

$$F(X(0)) + K\lambda \geq \mathbb{E}\left[_{T_n<\lambda}F\left(X_j\left(Tn\right)\right)\right] \geq \epsilon(\log(n)+1) \tag{A.15}$$

Let $n \to \infty$, we have

$$\infty > F(X(0)) + KT = \infty \tag{A.16}$$

Thus assumption (A.10) is wrong so that $\lim_{n\to\infty} T_n = \infty$ holds.

References

[1] Acharya V V,Skeie D, A model of liquidity hoarding and term premia in interbank markets[J], Journal of Monetary Economics, 201 1, 58(5): 436-447

[2] Acharya V V, Yorulmazer T, Information contagion and bank herding[J]], Journal of Money, Credit and Banking, 2008, 40(1): 215-231.

[3] Acharya V V A theory of systemic risk and design of prudential bank regulationJ , Stability, 2009, 5(3): 224255

[4] Ahn J H, Bignon V, Interbank market and central bank policy[J] , 2016

[5] Ahnert T,Georg C P, Information contagion and systemic risk[R] , Working Paper, LSE, 20-12

[6] Albert R, Barabfisi A L, Statistical mechanics of complex networks[J] , Reviews of modem physics, 2002, 74(1): 47

[7] Abert R, Jeong H, Barabhsi A L, Error and attack tolerance of complex networks[J] , Nature, 2000, 406(6794): 378-382

[8] Aleksiejuk A, HOLYST J A,Kossinets G, Self-organized criticality in a model of collective bankruptcies[J] , International Journal of Modem Physics C, 2002, 13(03): 333-341

[9] Bryan Perozzi, Rami Al-Rfou, Steven Skiena, Deep Walk: Online Learning of Social Representations , [cs.SI] 27 Jun 2014

[10] Aleksiejuk A, Ho|yst J A, A simple model of bank bankruptcies[J], sica A:Statistical Mechanics and itsApplications, 2001, 299(1): 198204

[11] Allen F, Gale D, Financial contagion[J] , Journal of political economy, 2000, 108(1):1-33

[12] Antonios K. Alexandridis, Wavelet Neural Networks , WILEY

[13] Battiston S, Gatti D D, Gallegati M, Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk , Journal of Economic Dynamics and Control, 2012, 36(8): I.121.

[14] Becher C, Millard S, Soramaki K, The network topology of CHAPS Sterling[J], Physical review letters, 2003, 90(7): 078701.

[15] Bianconi G, Capocci A, Number of loops of size h in growing scalefree networks[J], 2008.

[16] Boss M, Elsinger H, Summer M, An empirical analysis of the network structure of the Austriall , Oesterreichesche National Banks Financial stability Report, 2004, 77787.

[17] Caccioli F, Farmer J D, Foti N, Overlapping portfolios, contagion, and financial stability[J], Economic Dynamics and Control, 2015, 5150-63.

[18] Cassar A, Duffy N, VerSion F, Contagion offinancial crises under local and global networks , Agent-Based Methods in Economics and Finance, (Kluwer Academic Publishers, 2001).

[19] Craig B, Fecht F, Ttimer-Alkan G, The role of interbank relationships and liquidity needs[J], Banking and Finance, 2015, 53999-111.

[20] Craig B, Von Peter G, Interbank tiering and money center banks[J], Journal of Financial Intermediation, 2014, 23(3)322-347.

[21] Cormen T H, Leiserson C E, Rivest R L, Introduction to Algorithms[J] , 2001

[22] Craig B R, Fecht F, Ttimer-Alkan G, The role of interbank relationships and liquidity needs[J, Journal of Banking and Finance, 2015, 53: 99-1 1 1

[23] Craig B, Von Peter G, Interbank tiering and money center banks[J] , Journal of Financial Intermediation, 2014, 23(3): 322-347

[24] David Romer, Advanced Macroeconomics, Forth Edition, (Mcgraw Hill).

[25] Degryse H, Elahi MA, Penas MF, Crossborder exposures and financial contagion[J], International Review Of Finance, 2010, 10(2): 209-240.

[26] Degryse H, Elahi MA, Penas MF, CrossBorder Interbank Exposures and Financial Contagion[J], 2007.

[27] Degryse H, Nguyen G, An empirical examination of systemic risk in the belgian banking System[J], 2004.

[28] Davidsen t Ebei H, Bomholdt S, Emergence of a small world from local interactions: Modeling acquaintance-networks[J], Physical Review Letters, 2002, 88(12): 128701

[29] De Castro R, Grossman J W, Famous trails to paul erdrs[J] , The Mathematical Intelligencer, 1999, 21(3): 5l53

[30] De Menezes M A, Moukarzel C F, Penna T J P, First-order transition in small-world networks[J] , 2000, 50(5): 574 (EPL Euro physics Letters), 2000

[31] DeNicolo G, Kwast M L, Systemic risk and financial consolidation: Are they related?[]] , Journal of Banking and Finance, 2002, 26(5): 86 1-880

[32] De Cries C G, The simple economics of bank fragility[J], Journal of Banking and Finance, 2005, 29(4): 803825

[33] Eric D. Kolaczyk, Gábor Csárdi, and Gábor Csárdi, Statistical Analysis of Network Data with R, First Edition, (Springer New York).

[34] Gai P, Haldane A, Kapadia S, Complexity,concentration and contagion[J], Journal of Monetary Economics, 2011, 58(5): 453-470.

[35] GUADALUPE ERIKA IBARRA KURKA, TAILANDIA EN EL 2004: DE LA TRAGEDIA A LA POLÍTICA , ASIA PACÍFICO 2005

[36] Gareth James, An introduction to statistical Learning , Springer

[37] Huang, Yajing and Liu, Taoxiong and Lien, Donald, Portfolio Homogenization and Systemic Risk of Financial Network , 13 Oct 2017

[38] loft G, Jafarey S, Padilla F, Interbank Lending and Systemic Risk Kings College[R] , Department of Mathematics Working Paper, 2003

[39] Iori G, Reno R, De Masi G, Trading strategies in the Italian interbank market[J] , Physical A: Statistical Mechanics and its Applications, 2007, 376: 467-479

[40] Ivan Nunes da Silva, Danilo Hernane Spatti, Artificial Neural Networks , Springer

[41] Jizhe Wang, Pipei HuangAlibaba Group Billion-scale Commodity Embedding for E-commerce Recommendation in Alibaba, [CS.IR] 24 May 2018

[42] Krause A, Giansante S, Interbank lending and the spread of bank failures: A network model of systemic risk, Journal of EconomicBehavior and Organization, 2012, 83(3): 583-608.

[43] Kyunghyun Cho, Learning Phrase Representations using RNN EncoderDecoder, 2011.

[44] Ladley D, Contagion and risk-sharing on the interbank market[J], Journal of Economic Dynamics and Control, 2013, 37(7): 1384-1400.

[45] LeBaron B, Arthur W B, Palmer R Time series properties of an artificial stock market[J] , Journal of Economic Dynamics and control, 1999, 23(9): 1487-1516

[46] LeBaron B, Evolution and time horizon in an agent based stock market [J], Macroeconomic Dynamics, 2001,5(2): 225-254

[47] Levy M, Levy H, Solomon S, A microscopic model of the stock market: cycles, booms, and crashes[J] , Economics Letters, 1994, 45(1): 103-1 11

[48] Lubl0y A, Domino effect in the Hungarian interbank market[J] , Hungarian Economic Review, 2005, 52(4): 377-401

[49] Lux T, Marchesi M, Scaling and criticality in a stochastic multi-agent model of a financial market[J], Nature, 1999, 397(6719): 498-500.

[50] Matthew F. Dixon, Igor Halperin, Machine Learning in Finance , Springer

[51] Moussa A, Contagion and systemic risk in financial networks[D], Columbia University, 2011.

[52] Muller J, Interbank credit lines as a channel of contagion[J], Journal of Financial Services Research, 2006, 29(1): 37-60

[53] Newman ME J, Strogatz S H, Watts D J, Random Graphhs with arbitrary degree distributions and their applications[J], Physical review E, 2001, 64(2): 036118.

Nier EYang JYorulmazer Tet a1Network models and financifl stability[J]Journal of Economic Dynamics and Control20073 1(6)2033-2060

[54] Nier E, Yang J, Yorulmazer T, Network models and financifl stability[J], Journal of Economic Dynamics and Control, 2007, 31(6): 2033-2060.

[55] Newman M E J, The structure and function of complex networks[J], SIAM review, 2003, 45(2): 167-256

[56] Silva T C, de Souza S R S, Tabak B M, Network structure analysis of the Brazilian interbank market[J], Emerging Markets Review, 2016, 26: 130152.

[57] Steyvers M, Tenenbaum J B, The Large-scale structure of semantic networks, Statistical analyses and a model of semantic growth, 2005, 29(1): 41-78.

[58] STEFANO BONACCORSI AND STEF ANIA OTTA VIANO, EPIDEMICS ON NETWORKS WITH HETEROGENEOUS POPULATION AND STOCHASTIC INFECTION RATES, 3 Feb 2016[math.PR]

[59] Subana Shanmuganathan, Artificial Neural Network Modelling , Springer

[60] Simone Bassis, Advances in Neural Networks: Computational and Theoretical Issues, Springer

[61] Thumer S, Hanel R, Pichler S, Risk trading, network topology and banking regulation , Quantitative Finance, 2003, 3(4): 306-319.

Tabak B MCajueiro D 0Serra T RTopological properties of bank networksthe case of Brazil[J] International Journal ofModemPhysics C20091 1211 143

[62] Tabak B M, Cajueiro D, Topological properties of bank networks: the case of Brazil[J], nternational Journal of ModemPhysics C, 2009: 1121-1143.

[63] Watts D J, Strogatz S H, Collective dynamics of small-world networks [J], Nature, 1998, 393(6684) 440-442.

[64] Yang zhe Kong, Path-Link Graph Neural Network for IP Network Performance Prediction
,