A Novel Complex Network based Credit Risk Management Strategy

Xiangfeng Meng\textsuperscript{1,a}, Yunhai Tong\textsuperscript{1,b}, Xinhai Liu\textsuperscript{2,3,c} and Shaohua Tan\textsuperscript{1,d}

\textsuperscript{1}Center for Information Science, Peking University, China
\textsuperscript{2}Credit Reference Center, People's Bank of China, China
\textsuperscript{3}Center of Financial Intelligence Research, Peking University, China

\textsuperscript{a}pku.ericmeng@foxmail.com, \textsuperscript{b}yhtong@pku.edu.cn, \textsuperscript{c}xinhailiu@cis.pku.edu.cn, \textsuperscript{d}tan@cis.pku.edu.cn

Keywords: financial risk management; complex network; directed \textit{k}-shell decomposition method

Abstract. Financial system exhibits a high degree of interdependence. Different types of interconnections exist among financial agents, such as loan guarantees and cross-shareholding. Financial contagion could diffuse along the interconnections and lead to widespread risk. Thus how to effectively maintain the stability of financial system and prevent financial contagions from propagating is a crucial problem in financial risk management. In this paper, we present a novel risk management strategy based on complex network theory with regard to the guarantee bank loans in China. First, we transform the records of guarantee loans into directed guarantee networks, with nodes representing companies and edges representing the guarantee relations. With regard to the guarantee networks, we define two types of risk measurements for each company. The first one is the risk spreading ability of each company when it defaults. The second is the exposure risk each company is faced with. These two measurements together decide the credit risk of companies. Second, to evaluate these two types of measurements, a novel directed \textit{k}-shell decomposition method is presented. It is an effective way to measure the node centrality in terms of both in-degrees and out-degrees. The spreading abilities of companies can be obtained when the out-degrees of the companies are considered in the novel directed \textit{k}-shell decomposition method, while the exposure risk of companies can be obtained when the in-degrees are considered. Experimental analysis shows that the directed \textit{k}-shell decomposition method could identify meaningful companies in guarantee networks. Companies with high exposure risk are more likely to be infected and thus default. Meanwhile, companies with great spreading abilities could lead to widespread financial risk. Thus with our strategy, the financial regulators are able to monitor and immunize the targeted companies to maintain the stability of financial system.

Introduction

As the financial system is not developed in China, bank loan has been a popular approach for Small and Medium Enterprises to raise money. As the cost of assessing whether these kinds of enterprises have good credit is high in the absence of collateral, banks are reluctant to extend loans without explicit backing. Thus these enterprises have to seek for other enterprises to back loans for them. It is said that around a quarter of loans in China’s bank system are backed by such guarantees. These guarantee relations have been threatening to spread financial contagion of default risk from struggling firms to more healthy parts of the economy, which has happened in many regions of China.

Thus the understanding of the guarantee relationship is crucial to managing the default risk. Complex network theory provides a powerful tool to settle the problem. To be more detailed, the guarantee networks can be constructed with nodes representing companies and edges representing the loan guarantee relations. Once the guaranteed companies default, it is likely that the default risk will diffuse from one company to another along the direct edges. Thus how to identify companies that could lead to widespread credit risk or that are susceptible to infections becomes crucial to maintain the network’s stability. Such questions can be settled by the \textit{k}-shell decomposition method.

However, one major drawback of the standard \textit{k}-shell decomposition method is that it is designed to measure the centrality of nodes in undirected networks. It fails to capture the asymmetric relationships implied by the direct edges. In this paper, we propose the directed \textit{k}-shell decomposition method and develop a novel network based credit risk strategy.
The contributions of this paper are two-fold. First, we propose a network-based risk management strategy to control the guarantee networks in China. In contrast with traditional risk management approaches, this novel strategy models the default risk based on the interconnections within financial agents. Second, we propose a generalized $k$-shell decomposition method to obtain the $k$-shell structure of the directed networks. This novel method can be applied in both unweighted and weighted networks. Experiments on the guarantee networks demonstrate that this novel $k$-shell algorithm is able to identify crucial companies in financial risk management.

Related works
Due to the severe damage of the financial contagions, scholars have paid attentions to this problem [1][12][13]. The majority of studies are based on complex network theory and information diffusion models. Ji, Wang and Zhao [2] studied a guarantee network generated by the enterprises guarantee relations from a commercial bank. They found that the guarantee network is a small-world and scale-free network, and that the enterprises show an obvious geographical agglomeration. Zhang, Li and Guo [3] analyzed the process of the guarantee chain crisis, and found that both the asset-liability ratio and the external guarantee amount of companies should be monitored carefully. Gu [4] proposed that the formation of the Small and Medium Sized Enterprises credit guarantee risk is due to the asymmetric information among companies, and demonstrated this proposal.

Besides, a special issue about the topic “Complex Network in Finance” was delivered in Nature Physics. It proposed that the development of complex network theory will enhance the understanding of financial system and help maintain financial system’s stability [14] [15] [16].

In addition, how to measure node importance in networks is crucial. Various metrics have been proposed, among which the most popular methods are the centrality measurements. Several centrality measurements have been introduced, including the degree centrality, the betweenness centrality [5], the closeness centrality [6], and the eigenvector centrality [7]. These methods, however, cannot reveal the layered structure of networks. Thus the $k$-shell decomposition method is proposed. It measures the location of each vertex within networks, and assigns $k$-shell values to vertices, indicating their centrality. Due to the excellent performance of the $k$-shell decomposition method, it has been conducted in various applications [8][9][10][11].

Methodologies
The standard $k$-shell decomposition method
The standard $k$-shell decomposition method ($S_{k-shell}$) aims at partitioning undirected unweighted networks into sub-networks based on node centrality. It assigns an integer $k$ to each node according to its location in the network. Nodes with large $k$ values lie at the center of networks, and nodes with small $k$ lie at the periphery of the network. The set of nodes with large $k$ form lie at the core of networks.

The detailed procedures of the standard $k$-shell decomposition method are as follows. First, the method removes nodes with degree $k = 1$ from the network, and assigns the integer $k = 1$ to them. This step is repeatedly conducted until the degrees of all remaining nodes are above 1. Then, nodes with degrees $k = 2$ are removed from the network and assigned the integer $k = 2$. This step will be repeated again until the degrees of remaining nodes are above 2. Similar procedures are repeated for nodes with degrees $k \geq 3$ until all the nodes are assigned with $k$ values.

Fig. 1a shows the layered structure of an undirected network obtained by the standard $k$-shell decomposition method. We can see that the set of nodes within the innermost ring occupy the largest $k$ value, with $k = 3$. The set of nodes between the two outermost rings have the smallest $k$ value, with $k = 1$.

The direct $k$-shell decomposition method
In this section, we try to identify the $k$-shell structure from directed networks. Fig. 1b depicts a directed network. The undirected network shown in Fig. 1a can be obtained by symmetrizing the directed network. In Fig. 1b, the 7th company is guaranteed by another company, and the 17th company...
guarantee for other companies. In such situation, the $7^{th}$ company should be free from potential risk exposures due to risk diffusion, while the $17^{th}$ company suffers from default risk diffusion. However, the $k_s$ values of these two companies in undirected network are the same. Thus the standard $k$-shell decomposition method is not able to identify the $k$-shell structure of directed network. And a directed $k$-shell decomposition method is required.

In this paper, we propose the directed $k$-shell decomposition method ($D_{k-shell}$). This novel method applies a similar pruning routine as the standard $k$-shell decomposition method, but is based on an alternative measurement of nodes’ degrees. Intuitively, information can be diffused from the starting node to the ending node through the edge direction. Thus, the importance of each node can be measured from two perspectives. The first one is the node’s ability to diffuse information to other nodes. In this case, the out-degrees of nodes matter. The second one is how the node is affected by the information diffused from other nodes. And the in-degrees of nodes will matter. Correspondingly, there are two types of the directed $k$-shell decomposition methods. The difference between them is the selection of the nodes’ degree measurement.

Particularly, the out-degrees of nodes will be considered when the node importance indicates the ability to diffuse information. To be noticed, it is likely that the out-degrees of certain nodes are 0. Thus, the $k_s$ values of these nodes will be no larger than 1. Compared with the standard $k$-shell decomposition method, the first step of the directed $k$-shell decomposition method is different. Nodes with their out-degrees equal to 0 will be removed. This step will be iteratively carried out until the out-degrees of all remaining nodes are larger than 0. The directed $k$-shell decomposition method based on the out-degrees is illustrated in Alg. 1.

![Illustration of the $k$-shell decomposition method](a) Illustration of the $k$-shell decomposition method directed network

![An example of the directed network](b) An example of the directed network

**Fig. 1: Comparison between undirected network and directed network**

**The network-based credit risk management**

In the guarantee networks, credit risk diffuses from companies that raise money from banks to companies that backloans for them. For companies in China are densely connected through the guarantee relations, the majority of companies in the guarantee networks act as guaranteed companies and guarantee companies. Thus most companies are contagious and may diffuse risk to other companies when they default. They also suffer from the exposure risk and may be infected by other companies. Thus two types of credit risk are defined as follows.

The spreading abilities of companies: On one hand, how to identify companies that could diffuse risk across the network is crucial to prevent the cascading of default risk. The spreading abilities of companies rely on both the locations within the networks and the out-degrees of the companies. To measure the node centrality from the perspective of risk sources, the out-degrees of companies are taken into account in $D_{k-shell}$. Intuitively, nodes with out-degrees equal to 0 cannot diffuse risk to their
neighbors. Only nodes located at the center of the networks in terms of out-degrees could cause wide spreads of risk within the networks.

Exposure risk for companies: On the other hand, the immunization of the most susceptible companies could help prevent the risk diffusion from widely spreading. Thus how to identify the most susceptible companies within the guarantee networks is crucial. Intuitively, companies that back loans to other companies are susceptible to the risk diffusions. Sointuitively the in-degrees of companies should be taken into consideration. It is expected that companies located at the center of the networks in terms of in-degrees are more likely to be infected due to their greater exposure risk.

Core companies in guarantee networks: There are companies that own both large spreading abilities and great exposure risk in the guarantee networks. These companies are highly contagious and exposed to great risk as well. Severe financial contagions would occur if these companies default or are infected by other companies. Therefore, great attention should be paid to them by regulators. Without loss of generality, we denote the set of these companies as the cores of the guarantee networks.

Overall risk management strategy: The overall risk management strategy based on complex network theory is as follows. First, to maintain the stability of the guarantee networks, the core companies should always be regulated with caution. Besides, non-core companies with large spreading abilities that could diffuse risk to the core companies should also be well monitored. Thus the core companies would not be infected by external risk. In addition, non-core companies with great exposure risk that could be infected by the core companies should be immunized when the core companies default. This is due to the fact that these companies lie on the paths that are most likely to diffuse financial risk.

Algorithm 1: The directed $k$-shell decomposition method based on out-degrees

Require: A directed network $G = (V, E)$, with $|V| = n$

1. $KS := \emptyset$
2. $k := 0$
3. while $V \neq \emptyset$ do
   4. \hspace{1em} $index := \text{find nodes in } G \text{ with out-degree } k$
   5. \hspace{1em} $KS(index) := k$
   6. \hspace{1em} while $index \neq \emptyset$ do
      7. \hspace{2em} $V := V - V(index)$
      8. \hspace{2em} $E := E - E(index)$
      9. \hspace{2em} $G := G(V, E)$
      10. \hspace{2em} $index := \text{find nodes in } G \text{ with out-degree } \leq k$
   11. \hspace{1em} $KS(index) := k$
5. \hspace{1em} end while
6. $k := k + 1$
7. end while

Alg. 1: The directed $k$-shell decomposition method based on out-degrees

The guaranteed loan data set

The bank loan data set adopted in this research is obtained from the central bank of China. It is collected from all the banks in China, and is able to present a comprehensive description about the guaranteed loans. To be more detailed, this data set contains the records of bank loans that are guaranteed by other companies within twelve adjacent months. Among each record of the bank loans, the company that raises money from bank, company that back for the bank loan and the details of the bank loan are provided.

The bank loan guarantee networks can be constructed with nodes representing companies and direct edges representing the guarantee relations. Twelve temporal guarantee networks are constructed. For simplicity, we apply our novel method to the first snapshot of the temporal
guarantee networks. And the remaining 11 snapshots of the temporal guarantee networks are applied to evaluate the experimental results. Without loss of generality, when we talk about the guarantee network in the rest of the paper, we refer to the first snapshot of the temporal guarantee networks.

Experiments

Numerical evaluation settings

As companies own two types of credit risk measurement, these two measurements are evaluated respectively. First, it is expected that companies with larger spreading abilities are able to infect more other companies once these companies default. Second, for companies with greater exposure risk, it is expected that they are more likely to be infected by other companies, and that they are more likely to default in the future.

We apply the SIR model to simulate the diffusion process in networks. It is able to measure how many companies could be infected by certain company (spreading abilities), and how many companies could infect one certain company (exposure risk). The detailed procedures are as follows. The probability $p$ of infection from infected nodes to healthy nodes is assumed to be unique, and varies in $[0.1,1]$ with the interval $0.1$. For each node in the network, the SIR model is carried out for 50 times. Two numerical measurements are calculated to evaluate the $k$-shell structure obtained by $D_{k-shell}$, which will be illustrated in corresponding parts.

Besides, as whether each company defaults or not in the following 11 months are known in advance, whether companies with greater exposure risk are more likely to default could be evaluated by the average default rate of companies with different exposure risk.

Evaluations of company’s centrality in terms of exposure risk

In this part, we try to evaluate the company’s centrality in terms of exposure risk. It is obtained by the $D_{k-shell}$ method based on in-degrees of companies, and denoted by $ks_{dest}$. It is expected that companies with greater exposure risk are more likely to default. Thus for companies with each $ks_{dest}$ value, we calculate the average default rate of these companies in both half a year and one year, and show them in Table 1.

In Table 1, we can see that the default rate of companies with $ks_{dest} = 1$ is the lowest, followed by that of companies with $ks_{dest} = 2$, and the default rate of companies with $ks_{dest} = 3$ are the largest. To be more detailed, only 2.23% of the companies with $ks_{dest} = 1$ default in the next half a year, and 2.92% of these companies default in the next one year. The $ks_{dest}$ values of 199 companies are 2, with 2.53% of these companies defaulting in the next half a year and 3.48% defaulting in the next one year. Among the 11 companies with $ks_{dest} = 3$, 45.45% of these companies default in the next half a year, and 54.55% default in the next one year. Thus, companies with larger $ks_{dest}$ values (greater exposure risk) are more likely to default in the future.

Table 1: The default rate of companies with different $ks_{dest}$ values obtained by $D_{k-shell}$

<table>
<thead>
<tr>
<th>$ks_{dest}$</th>
<th>Number of companies</th>
<th>Default rate in half a year</th>
<th>Default rate in one year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3356</td>
<td>0.0223</td>
<td>0.0292</td>
</tr>
<tr>
<td>2</td>
<td>316</td>
<td>0.0253</td>
<td>0.0348</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td><strong>0.4545</strong></td>
<td><strong>0.5455</strong></td>
</tr>
</tbody>
</table>

Besides, companies with larger exposure risk are more likely to be infected by other companies. We apply the SIR model to test whether these companies are more vulnerable to financial risk. Fig. 2 depicts how many companies on average could diffuse risk to companies with certain $ks_{dest}$ value by applying the SIR model. The x-axis stands for the probability that risk diffuses along each directed edge. The diffusion probability ranges from 0.1 to 1, with the interval 0.1. The y-axis stands for the average number of companies that could diffuse risk to companies with certain $ks_{dest}$ value.

From Fig. 2, we can see that companies with $ks_{dest} = -1$ cannot be infected by other companies, as these companies don’t back loans for any company. Thus the average number of companies that could infect these companies with $ks_{dest} = -1$ remain to be 1 (with the node itself counted in), even though the infection probability increases from 0.1 to 1.
With regard to companies with nonnegative $k_s$ values, the average number of companies that could diffuse default risk to them increases with the infection probability. Besides, with the same infection probability, companies with larger $k_s$ values tend to be more vulnerable to financial contagions. For instance, with the infection probability $p = 0.8$, 2.7 companies on average could infect companies with $k_s = 0$, more than 20 companies on average could infect companies with $k_s = 1$, and around 60 nodes, on average, could infect companies with $k_s = 2$.

Companies with $k_s = 3$ are exceptions. With the infection probability less than 0.5, the average number of companies that could infect companies with $k_s = 3$ is the largest, which is in accordance with the above conclusions. However, when the infection probability is no less than 0.5, the average number of companies that could infect companies with $k_s = 3$ is less than that with $k_s = 2$. The potential reason is that the 11 companies with $k_s = 3$ form a tightly connected group. This group has only several connections with the rest of the network. So this group is nearly isolated, and does not lie at the center of the network.

![Fig. 2: Average number of companies that could infect companies with different $k_s$ values](image)

**Fig. 2: Average number of companies that could infect companies with different $k_s$ values**

**Evaluations of company centrality in terms of spreading abilities**

In this part, we try to evaluate the company’s centrality in terms of their spreading abilities. This centrality is obtained by $D_{k-shell}$ based on out-degrees of companies, and denoted by $k_{source}$. The SIR model is carried out. Fig. 3 depicts the spreading abilities of companies with certain $k_{source}$ values by applying the SIR model.

From Fig. 3, we can see that companies with $k_{source} = -1$ are unable to diffuse risk. Thus the average number of companies that are infected from nodes with $k_{source} = -1$ remain to be 1 (with the node itself counted in), even though the infection probability increases from 0.1 to 1. With regard to the nonnegative $k_{source}$ values, the average number of infected companies increases with the infection probability.

Besides, with the same infection probability, companies with larger $k_{source}$ values tend to be more contagious. For instance, with the infection probability $p = 0.8$, companies with $k_{source} = 0$ could averagely infect 2.8 companies, companies with $k_{source} = 1$ could infect more than 20 companies, and companies with $k_{source} = 2$ could infect more than 60 companies.

**Correlations between two types of risk**

The correlations between the two $k_s$ values are presented in Fig. 4, in which the x-axis depicts the $k_{source}$ values obtained by $D_{k-shell}$ in terms of the out-degrees and the y-axis depicts the $k_{dest}$ values in terms of the in-degrees. Disturbance is added to the $k_s$ values so that nodes with the same coordination will not converge to a single point.

In Fig. 4, the majority of nodes are located at the left or the bottom of the figure. It indicates that these companies in the guarantee network are either not contagious or vulnerable to the risk diffusion. Besides, large $k_{source}$ values of nodes do not guarantee that the nodes will own large $k_{dest}$ values.
Particularly, Fig. 4 could be partitioned into four disjoint regions. Companies in region ‘A’ own the smallest $k_{source}$ values, which indicates that these nodes are vulnerable to being infected, but are not contagious. On the contrast, companies in region ‘B’ are contagious but are free from the infections of default risk. Companies in region ‘D’ are both highly contagious and exposed to great default risk. Companies in region ‘C’ are both contagious and exposed to potential default risk. But compared with those in region ‘D’, companies in region ‘C’ are less contagious and less likely to be infected.

Furthermore, we present the characteristics of all the disjoint regions in Fig. 4 in terms of both spreading abilities and exposure risk in Table 2. It is evident that companies in region ‘D’ should be highlighted and those in region ‘C’ should also be monitored carefully.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Is contagious?</th>
<th>Exposed to credit risk?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>C1</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>C2</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>D</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Fig. 3: Average number of companies infected by companies with different $k_{source}$ values

Fig. 4: Correlations between $k_{source}$ values and $k_{dest}$ values of companies

Table 2: Characteristics of different regions in Fig. 4
Conclusions

This paper presents a network-based credit risk management strategy. This strategy aims at identifying important companies that are highly contagious or are exposed to greater credit risk within the guarantee networks. The identification of these companies helps to control the stability of the guarantee networks and prevent the credit risk from widely spreading. To identify these companies, we present a novel directed $k$-shell decomposition method. Through detailed analysis, we find that the directed $k$-shell decomposition method based on the in-degrees of nodes could identify the most susceptible companies. And the method based on the out-degrees of nodes could identify the most contagious companies. Experiments show that more refined $k$-shell structure can be obtained by the directed $k$-shell decomposition method than that obtained by the classical $k$-shell decomposition method. And the identified companies are meaningful, and could be used to control the stability of the guarantee networks and prevent the financial contagions from widely spreading. As far as we know, this is the first time to settle the financial risk management problem of the guarantee loans with network-based approaches.

References


