A Novel Complex Network based Credit Risk Management Strategy

Xiangfeng Meng[1,a], Yunhai Tong[1,b], Xinhai Liu[2,3,c] and Shaohua Tan[1,d]

[1]Center for Information Science, Peking University, China

[2]Credit Reference Center, People's Bank of China, China

[3]Center of Financial Intelligence Research, Peking University, China

[a]pku.ericmeng@foxmail.com, [b]yhtong@pku.edu.cn, [c]xinhailiu@cis.pku.edu.cn, [d]tan@cis.pku.edu.cn

Abstract. Financial system exhibits a high degree of interdependence. Different types of interconnections exist among financial agents, such as loanguarantees and cross share-holding. Financial contagion could diffuse along the interconnections and lead to wide spreadof 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 complexnetwork theory with regard to the guarantee bank loans inChina. First, we transform the records of guarantee loans intodirected guarantee networks, with nodes representing companiesand edges representing the guarantee relations. With regard to theguarantee networks, we define two types of risk measurementsfor each company. The first one is the risk spreading abilityof each company when it defaults. The second is the exposurerisk each company is faced with. These two measurementstogether decide the credit risk of companies. Second, to evaluate these two types of measurements, a novel directed$k$-shell decomposition method is presented. It is an effectiveway to measure the node centrality in terms of both in-degreesand out-degrees. The spreading abilities of companiescan be obtained when the out-degrees of the companies areconsidered in the novel directed k-shell decomposition method, while the exposure risk of companies can be obtained when thein-degrees are considered. Experimental analysis shows that thedirected k-shell decomposition method could identify meaningfulcompanies in guarantee networks. Companies with high exposurerisk are more likely to be infected and thus default. Meanwhilecompanies with great spreading abilities could lead to wide spreadof financial risk. Thus with our strategy, the financial regulators are able to monitor and immunize the targeted companies tomaintain the stability of financial system.

Introduction

Asthe financial system is not developed in China, bank loan hasbeen a popular approach for Small and Medium Enterprises toraise money. As the cost of assessing whether these kinds ofenterprises 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 toback loans for them. It is said that around a quarter ofloans in China's bank system are backed by such guarantees.These guaranteerelations have been threatening to spread financial contagionof default risk from struggling firms to more healthy parts ofthe economy, which has happened in many regions of China

Thus the understanding of the guarantee relationship is crucial to managingthe 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 thedefault risk will diffuse from one company to another along the direct edges.Thus how to identifycompanies that could lead to wide spread of creditrisk or that are susceptible to infections becomes crucial tomaintain the network's stability. Such questions can be settledby the k-shell decomposition method.

However, one major drawback of the standard k-shell decomposition method is that it is designed to measure the centralityof nodes in undirected networks.It fails to capturethe asymmetric relationships implied by the direct edges.In this paper, we propose thedirected k-shell decomposition method anddevelop 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 withtraditional risk management approaches, this novel strategymodels the default risk based on the interconnectionswithin 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 inboth unweighted and weighted networks. Experiments on theguarantee networks demonstrate that this novel k-shell algorithm is able to identify crucial companiesin financial risk management.

Related works

Due to the severe damage of the financial contagions, scholars have paidattentions to this problem [1][12][13]. The majority of studies are basedon complex network theory and information diffusion models.Ji, Wang and Zhao [2] studied a guarantee network generatedby the enterprises guarantee relations from a commercial bank.They found that the guarantee network is a small-world andscale-free network, and that the enterprises show an obviousgeographical agglomeration. Zhang, Li and Guo [3] analyzedthe process of the guarantee chain crisis, and found that boththe asset-liability ratio and the external guarantee amount ofcompanies should be monitored carefully. Gu [4] proposedthat the formation of the Small and Medium Sized Enterprisescredit guarantee risk is due to the asymmetric informationamong companies, and demonstrated this proposal.

Besides, a special issue about the topic "ComplexNetwork in Finance" was delivered in Nature Physics. Itproposed that the development of complex network theorywill enhance the understanding of financial system and helpmaintain financial system's stability [14] [15] [16].

In addition, how to measure node importance in networks is crucial. Various metrics havebeen proposed, among whichthe most popular methods are the centrality measurements.Several centrality measurements have been introduced, including the degree centrality, the betweenness centrality [5], thecloseness centrality [6], and the eigenvector centrality [7].These methods, however, cannot reveal the layeredstructure of networks. Thus the k-shell decompositionmethod is proposed. It measures the location of each vertexwithin networks, and assignsksvalues to vertices,indicating their centrality. Due to the excellent performanceof the$k$-shell decomposition method, it has beenconducted 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 ks toeach node according to its location in the network. Nodes withlarge ks values lie at the center of networks, and nodes with small kslie at the periphery of the network. The set of nodes with large ks form lie at the core of networks.

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

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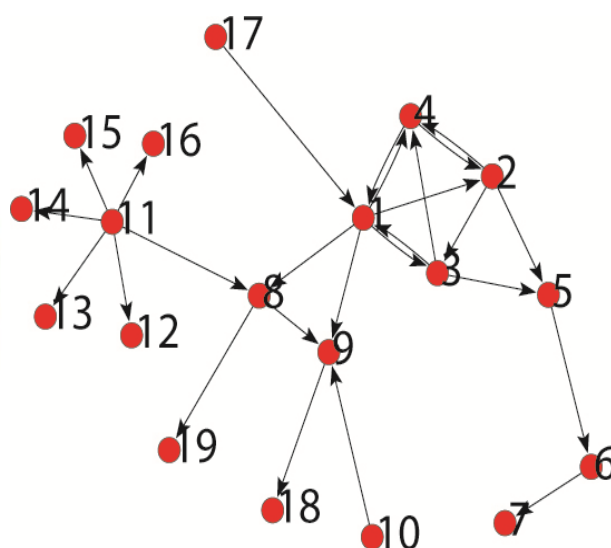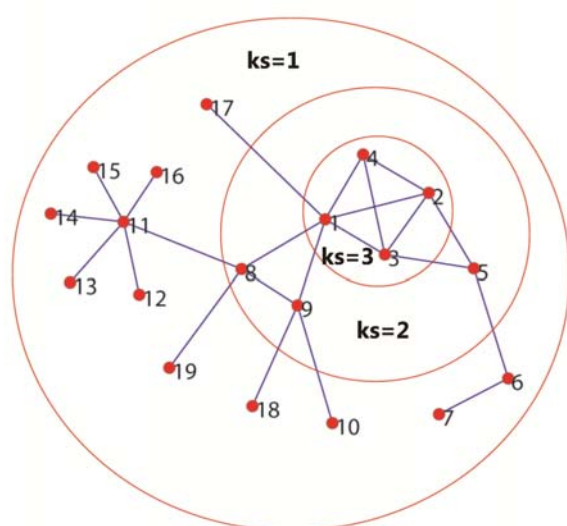
The direct k-shell decomposition method

In this section, we try to identify the k-shell structurefrom directed networks. Fig. 1b depicts a directed network.The undirected network shown in Fig. 1a can be obtainedby symmetrizing the directed network. In Fig. 1b, the 7[th]company is guaranteed by another company, and the 17[th]company

guarantee for other companies. In such situation, the7[th]company should be free from potential risk exposures dueto risk diffusion, while the 17[th]company suffers from defaultrisk diffusion. However, the ks values of these two companiesin 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*-shelldecomposition method is required.

In this paper, we propose the directed *k*-shell decomposition method ($D_{k-shell}$). This novel method applies a similarpruning 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. Thesecond one is how the node is affected by the informationdiffused 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 theability to diffuse information. To be noticed, it is likely thatthe out-degrees of certain nodes are 0. Thus, the ks valuesof these nodes will be no larger than 1. Compared with thestandard *k*-shell decomposition method, the first step of thedirected *k*-shell decomposition method is different. Nodes withtheir out-degrees equal to 0 will be removed. This step will beiteratively carried out until the out-degrees of all remainingnodesare larger than 0. The directed *k*-shell decompositionmethod based on the out-degrees is illustrated in Alg. 1.

(a) Illustration of the *k*-shell decomposition method directed network   (b) An example of the

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 mostcompanies are contagious and may diffuse risk to othercompanies when they default. They also suffer from theexposure 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, howto identify companies that could diffuse risk across thenetwork is crucial to prevent the cascading of default risk. Thespreading abilities of companies rely on both the locationswithin the networks and the out-degrees of the companies.To measure the node centrality from the perspective of risksources, the out-degrees of companies are taken into account in$D_{k-shell}$. Intuitively, nodes with out-degrees equal to 0 cannotdiffuse risk to their

neighbors. Only nodeslocated at the center of the networks in terms of out-degreescould 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 toidentify the most susceptible companies within the guaranteenetworks is crucial. Intuitively, companies that back loansto other companies are susceptible to the risk diffusions. Sointuitively the in-degrees of companies should be taken intoconsideration. It is expected that companies located at thecenter of the networks in terms of in-degrees are more likelyto be infected due to their greater exposure risk.

Core companies in guarantee networks: There arecompanies that own both large spreading abilities and greatexposure risk in the guarantee networks. These companies arehighly contagious and exposed to great risk as well. Severe financialcontagions would occur if these companies default or are infectedby other companies. Therfore, great attention should be paid to themby regulators. Without loss of generality, we denote the set ofthese companies as the cores of the guarantee networks.

Overall risk management strategy: The overall riskmanagement strategy based on complex network theory isas follows. First, to maintain the stability of the guaranteenetworks, the core companies should always be regulated withcautions. Besides, non-core companies with large spreadingabilities that could diffuse risk to the core companies shouldalso be well monitored. Thus the core companies would notbe infected by external risk. In addition, non-core companieswith great exposure risk that could be infected by the corecompanies should be immunized when the core companiesdefault. This is due to the fact that these companies lie onthe 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:      $index :=$ find nodes in G with out-degree k
5:      $KS(index) := k$
6:      **while** $index \neq \emptyset$ **do**
7:          $V := V - V(index)$
8:          $E := E - E(index)$
9:          $G := G(V, E)$
10:         $index :=$ find nodes in G with out-degree $\leq k$
11:         $KS(index) := k$
12:      **end while**
13:      $k := k + 1$
14: **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 obtainedfrom the central bank of China. It is collected from all thebanks in China, and is able to present a comprehensive description about the guaranteed loans. To be more detailed, thisdata set contains the records of bank loans that are guaranteedby other companies within twelve adjacent months. Amongeach record of the bank loans, the company that raises moneyfrom bank, company that back for the bank loan and the detailsof the bank loan are provided.

The bank loan guarantee networks can be constructedwith nodes representing companies and direct edges representing the guarantee relations. Twelve temporal guaranteenetworks are constructed. For simplicity, we apply ournovel method to the first snapshot of the temporal

guaranteenetworks. And the remaining 11 snapshots of the temporalguarantee networks are applied to evaluate the experimentalresults. Without loss of generality, when we talk about theguarantee network in the rest of the paper, we refer to the firstsnapshot 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, itis expected that companies with larger spreading abilities areable to infect more other companies once these companiesdefault. Second, for companies with greater exposure risk, itis expected that they are more likely to be infected by othercompanies, and that they are more likely to default in thefuture.

We apply the SIR model to simulate thediffusion process in networks. It is able to measure how many companies could beinfected by certain company (spreading abilities), and howmany companies could infect one certain company (exposurerisk). The detailed procedures are as follows.The probability p of infection from infected nodes to healthy nodes isassumed to be unique, and varies in [0.1,1] with the interval0.1.For each node in the network, the SIR model is carried outfor 50 times. Two numerical measurements are calculated toevaluate 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 thefollowing 11 months are known in advance, whether companies with greater exposure risk are more likely to default couldbe evaluated by the average default rate of companies withdifferent exposure risk.

Evaluations of company's centrality in terms of exposurerisk

In this part, we try to evaluate the company's centralityin 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 exposurerisk are more likely to default. Thus for companies with each ks_{dest} value, we calculate the average default rate of thesecompanies in both half a year and one year, and show themin Table 1.

In Table 1, we can see that the default rate of companieswith $ks_{dest} = 1$ is the lowest, followed by that of companieswith $ks_{dest} = 2$, and the default rate of companies with $ks_{dest} = 3$ are the largest. To be more detailed, only 2.23% ofthe 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 themdefaults in the next half a year and 3.48% default in the next oneyear. Among the 11 companies with $ks_{dest} = 3$, 45.45% of them default inthe next half a year, and 54.55% default in the next one year. Thuscompanies 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 $D_{k-shell}$

ks_{dest}	Number of companies	Default rate in half a year	Default rate in one year
1	3356	0.0223	0.0292
2	316	0.0253	0.0348
3	11	**0.4545**	**0.5455**

Besides, companies with larger exposure risk are more likely tobe infected byother companies.We apply the SIR model to test whether these companiesare more vulnerable to financial risk. Fig. 2 depicts how manycompanies on average could diffuse risk to companies with certain ks_{dest} value by applying the SIR model. The x-axis stands for theprobability that risk diffuses along each direct edge. Thediffusion probability ranges from 0.1 to 1, with the interval0.1. The y-axis stands for the average number of companiesthat 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 donot back loans for any company. Thus the average numberof companies that could infect these companies with $ks_{dest} = -1$ remain to be 1 (with the node itself counted in), eventhough the infection probability increases from 0.1 to 1.

With regard to companies with nonnegative ks 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 ksvalues tend to be more vulnerable to financial contagions. Forinstance, with the infection probability p = 0.8, 2.7 companieson average could infect companies with $ks_{dest} = 0$, morethan 20 companies on average could infect companies with$ks_{dest} = 1$, and around 60 nodes, on average, could infectcompanies with $ks_{dest} = 2$.

Companies with $ks_{dest} = 3$are exceptions. With theinfection probability less than 0.5, the average number ofcompanies that could infect companies with $ks_{dest} = 3$ is thelargest, 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 companieswith $ks_{dest} = 3$ is less than that with $ks_{dest} = 2$. The potentialreason is that the 11 companies with $ks_{dest} = 3$ form a tightlyconnected group. This group has only several connections withthe rest of the network. So this group is nearlyisolated, and does not lie at the center of the network.
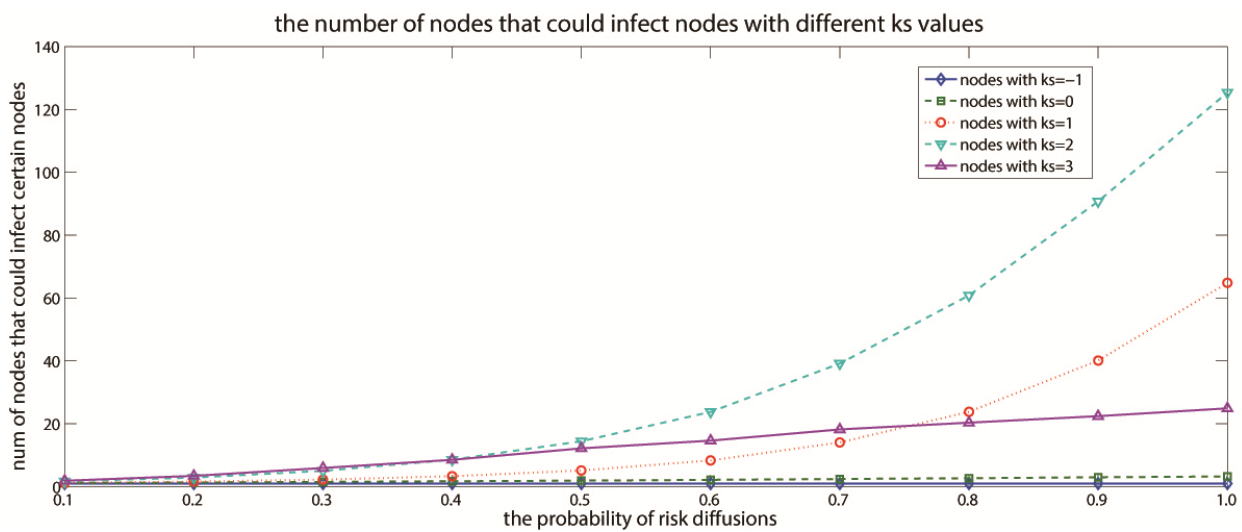

Fig. 2: Average number of companies that could infect companies with different ks_{dest} values

Evaluations of company centrality in terms of spreadingabilities
In this part, we try to evaluate the company'scentrality in terms of their spreading abilities. This centrality isobtained by $D_{k-shell}$ based on out-degrees ofcompanies, and denoted by ks_{source}. The SIR model is carriedout. Fig. 3 depicts the spreading abilities of companies withcertainks_{source} values by applying the SIR model.

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

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

Correlations between two types of risk
The correlations between the two ks values are presentedin Fig. 4, in which the x-axis depicts the ks_{source} valuesobtained by $D_{k-shell}$ in terms of the out-degrees andthe y-axis depicts theks_{dest} values in terms of the in-degrees.Disturbance is added to the ks values so that nodes with the same coordination willnot converge to a single point.

In Fig. 4, the majority of nodes are located at the left orthe bottom of the figure. It indicates that these companies inthe guarantee network are either not contagious or vulnerable to therisk diffusion. Besides, largeks_{source} values of nodes donot guarantee that the nodes will own largeks_{dest} values.

Particularly, Fig. 4 could be partitioned into four disjointregions. Companies in region 'A' own the smallestks_{source}values, which indicates that these nodes are vulnerable to beinfected, but are not contagious. On the contrast, companies inregion '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. Companiesin 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 likelyto be infected.

Furthermore, we present the characteristics of allthe disjoint regions in Fig. 4 in terms of both spreading abilitiesand exposure risk in Table 2. It is evident that companiesin region 'D' should be highlighted and those in region 'C'should also be monitored carefully.
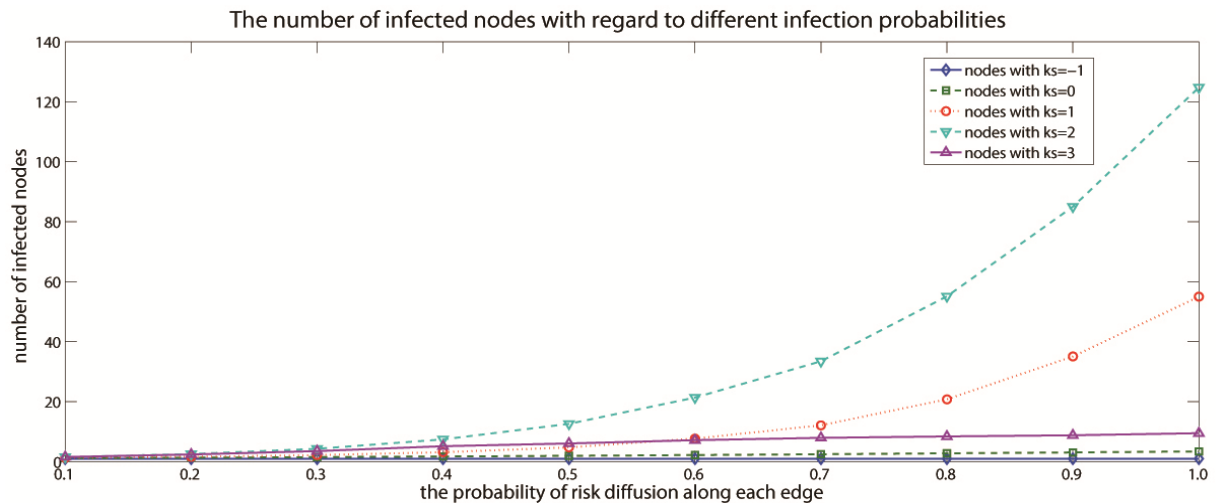

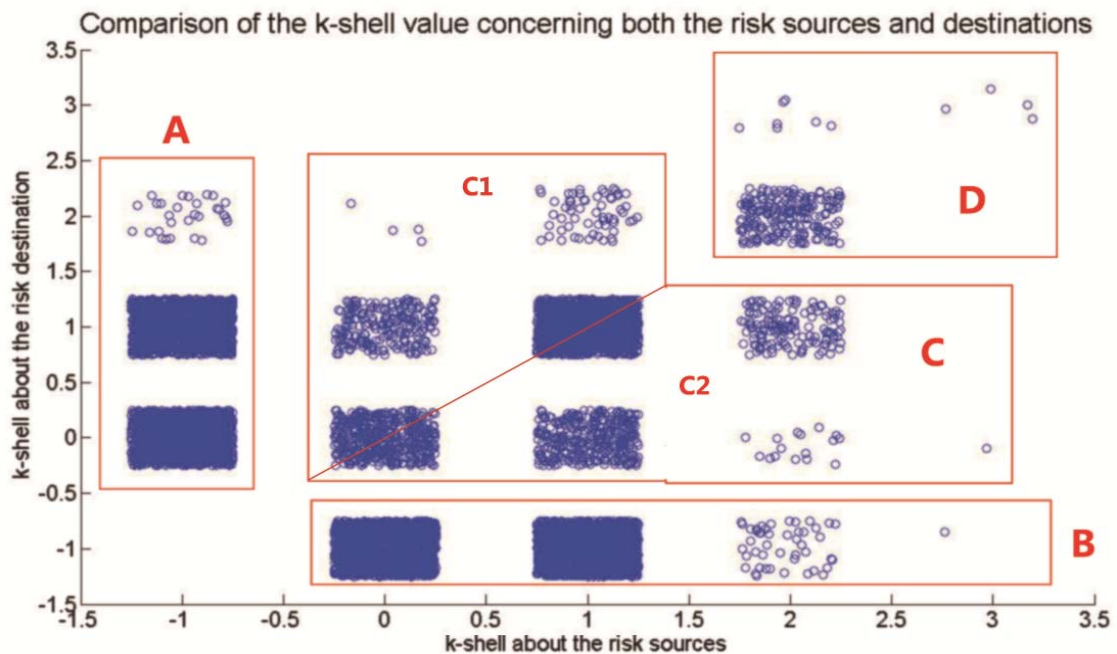Fig. 3: Average number of companies infected by companieswith different ks_{source}values

Fig. 4: Correlations between ks_{source} values and ks_{dest} valuesof companies

Table 2: Characteristics of different regions in Fig. 4

Regions	Is contagious?	Exposed to credit risk?
A	No	Yes
B	Yes	No
C1	Low	Medium
C2	Medium	Low
D	High	High

Conclusions

This paper presents a network-based credit risk management strategy. This strategy aims at identifying importantcompanies that are highly contagious or are exposed to greatcredit risk within the guarantee networks. The identification of these companies helps to control the stability of the guaranteenetworks and prevent the credit risk from widely spreading.To identify these companies, we present a novel directed k-shell decomposition method. Through detailed analysis, wefind that the directed k-shell decomposition method based onthe in-degrees of nodes could identify the most susceptiblecompanies. And the method based on the out-degrees of node could identify the most contagious companies. Experimentsshow that more refined k-shell structure can beobtained by the directed k-shell decomposition method thanthat obtained by the classical k-shell decomposition method.And the identified companies are meaningful, and could beused to control the stability of the guarantee networks andprevent the financial contagions from widely spreading.As far as we know, this is the first time to settle thefinancial risk management problem of the guarantee loanswith network based approaches.

References

[1] T. Van Gestel and B. Baesens. Credit Risk Management: Basic Concepts: Financial Risk Components, Rating Analysis, Models, Economicand Regulatory Capital. OxfordUniversity Press, 2008.

[2] J. Yanbing, W. Wei, and Z. Yawei. C0MPLEX SYSTEMS ANDCOMPLEXITY SCIENCE (In CHINESE), 11(2):17–23, 2014.

[3] Z.-x. Zhang, P.-x. Li, and J.-e. Guo. Systems Engineering, 4:005, 2012.

[4] G. Hai-feng. The Theory and Practice of Finance andEconomics, 3:001, 2007.

[5] L. C. Freeman. Sociometry, pages 35–41, 1977.

[6] M. E. J. Newman. Social networks, 27(1):39–54, 2005.

[7] P. Bonacich. Americanjournal of sociology, pages 1170–1182, 1987.

[8] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E.Stanley, and H. A. Makse. Nature Physics, 6(11):888–893, 2010.

[9] A. Garas, P. Argyrakis, C. Rozenblat, M. Tomassini, and S. Havlin.New Journal of Physics,12(11):113043, 2010.

[10] A. Garas, F. Schweitzer, and S. Havlin. New Journal of Physics, 14(8):083030, 2012.

[11] B. Wei, J. Liu, D. Wei, C. Gao, and Y. Deng. Physica A: Statistical Mechanics and its Applications, 420:277–283,2015.

[12] F. Allen, A. Babus, and E. Carletti. Financial connections and systemicrisk. Technical report, National Bureau of Economic Research, 2010.

[13] P. Gai and S. Kapadia. In Proceedings of the Royal Society of London A: Mathematical, Physical andEngineering Sciences, page rspa20090410. The Royal Society, 2010.

[14] M. Galbiati, D. Delpini, and S. Battiston. NaturePhysics, 9(3):126–128, 2013.

[15] G. Caldarelli, A. Chessa, F. Pammolli, A. Gabrielli, and M. Puliga.Nature Physics, 9(3):125–126, 2013.

[16] M. Catanzaro and M. Buchanan. Network opportunity. Nature Physics,9(3):121–123, 2013.